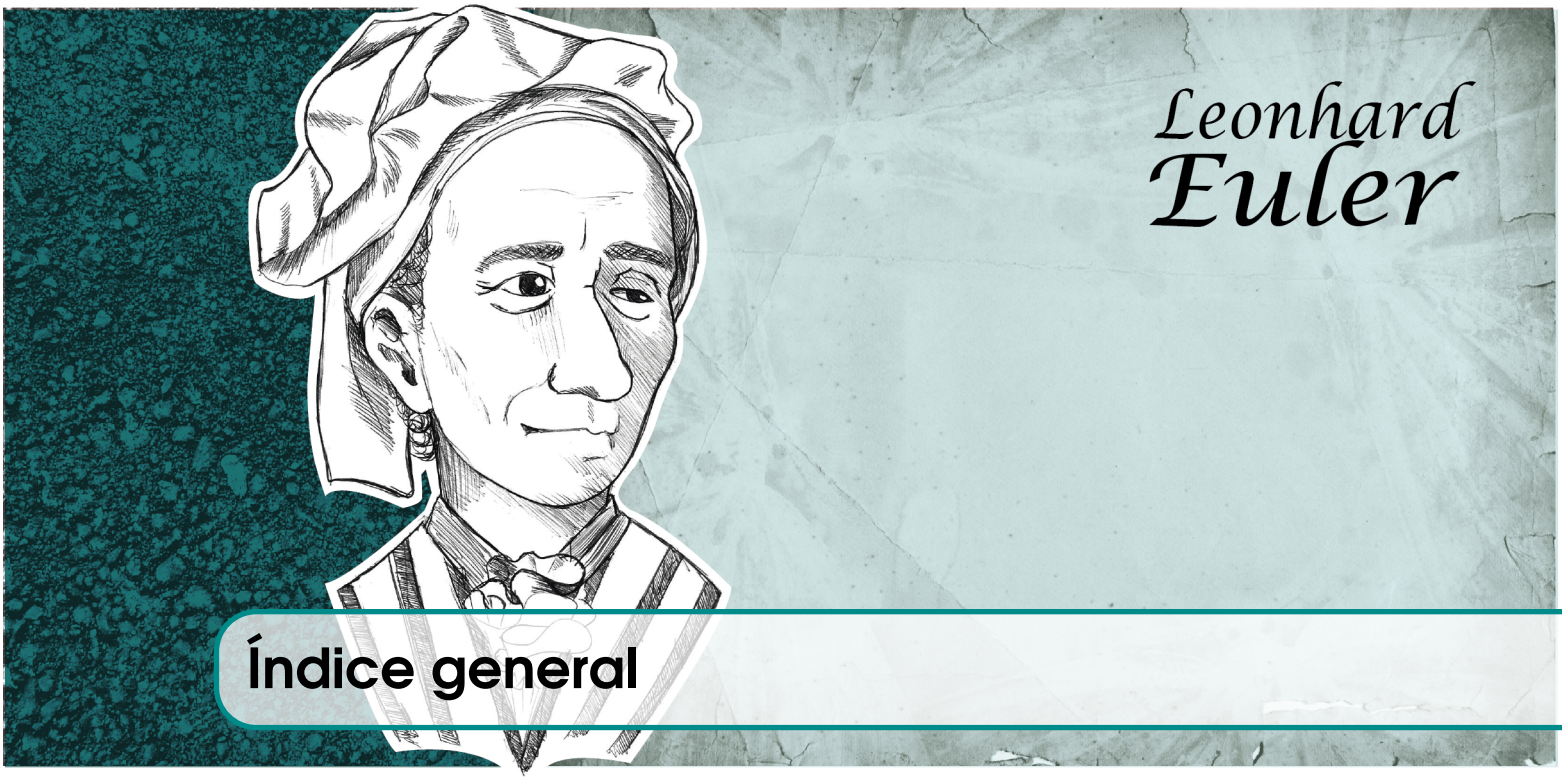
Teorema de los números primos

Trabajo para la obtención del grado de Licenciado en Ciencias UTFSM

Eduardo Núñez Olguín



Índice general

Introducción 5

1 Breve historia de un teorema 9
1.1 La infinitud analítica de los números primos 10
1.2 Conjeturando desde una tabla de números primos 12
1.3 El complejo camino a la demostración 15

2 Extensión analítica y función Gamma 25
2.1 Extensión analítica 26
2.2 Función Gamma 29
2.3 Polos y raíces de la Función Gamma 34

3 Función dseta de Riemann 37
3.1 Extensión analítica a la función dseta 38
3.2 Identidad de Euler 44
3.3 Estimaciones de su crecimiento 50
3.4 Ceros de la función dseta 56

4 Teorema de los números primos 59
4.1 Describiendo el problema 60
4.2 Reducciones de Tchebyshev 62
4.3 Relación con la función dseta 65
4.4 El último paso 70

Conclusiones 79

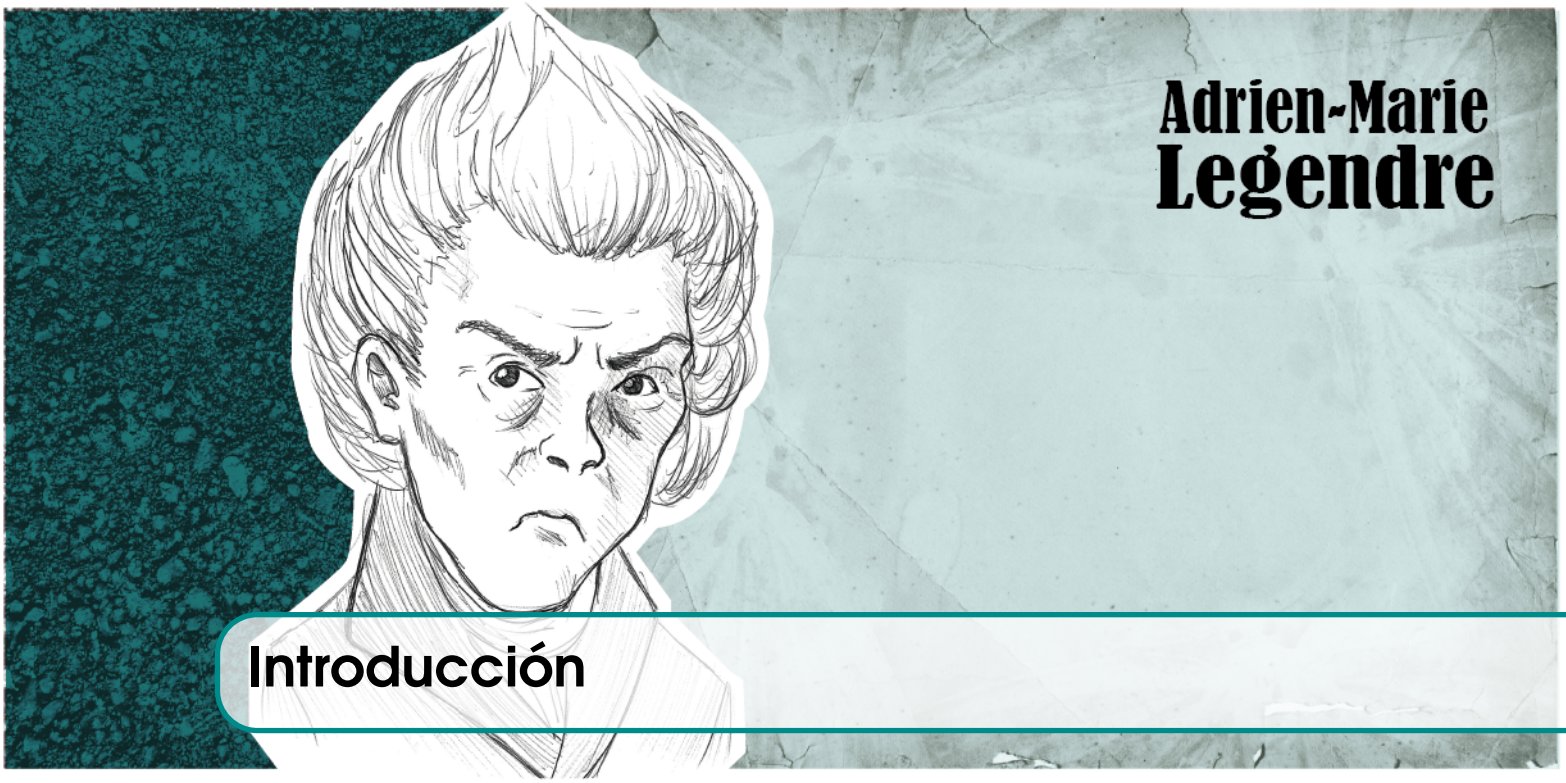
5 Apéndice: Análisis complejo 81

Apéndice: Análisis complejo 81

5.1 Teoremas sobre funciones holomorfas 81

5.2 Función Theta 83

Bibliografía 85



Introducción

Los números primos ocupan un lugar especial en el corazón de las matemáticas. Esta consideración procede, en parte, de la maravillosa propiedad de que todo número natural puede expresarse como un único producto de factores primos, cuya trascendencia le ha ganado con justo mérito el nombre de Teorema Fundamental de la Aritmética, siendo clave para la demostración de varias proposiciones tanto en la teoría numérica como en otras áreas matemáticas. Pero, la ciencia, y en particular la matemática, es una disciplina enamorada de lo todavía por conocer, y gran parte de la fascinación que sienten tanto matemáticos profesionales como aficionados hacia los números primos proviene de ese halo de misterio que surge al vincularse con varias de las conjeturas más resistentes a los esfuerzos de demostración, como son la Hipótesis de Riemann o los Problemas de Landau. Esa mezcla de enigma e importancia se vuelve un canto de sirena, irresistible para la curiosidad matemática. Varias de estas interrogantes donde los números primos juegan un papel decisivo, incluyendo las conjeturas ya mencionadas, se derivan de la distribución aparentemente aleatoria de los números primos en la sucesión de números naturales. La secuencia de números primos, veamos...

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, \ldots$$

nos es conocida desde los primeros niveles de la escuela y, sin embargo, se resiste a ser sintetizada en una breve fórmula o calculada por algún rápido algoritmo. La dificultad de este problema es tal que nuestro fracaso ha sido implementado por sistemas criptográficos (RSA y similares) con tanto éxito que hoy es el pilar fundamental de nuestra seguridad informática moderna. Leonhard Euler, el genio que dominó la matemática del siglo XVIII, se lamentó una vez:

> No hay más que darles una mirada a las tablas de números primos —que algunas personas se han dado el trabajo de seguir más allá del cien mil— para darse cuenta, primero, que ahí no reina ningún orden ni regla. [3]

Como es costumbre al espíritu de la ciencia, tales palabras no constituyeron un profético impedimento, sino que impulsaron un desafío que motivó a múltiples matemáticos a buscar patrones que otorgaran algún orden, aunque sea parcial o aproximado, a dichas caóticas tablas. Dentro de estas observaciones, posiblemente las más importantes sean las que condujeron a conjeturar el **Teorema de los Números Primos**: Adrien-Marie Legendre y Carl Friedrich Gauss observaron a finales del siglo XVIII que la cantidad de números primos menores que un valor dado es aproximable por dicho valor dividido su logaritmo natural, esto es:

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

donde $\pi(x)$ representa la cantidad de números primos menores a x. Esta aseveración constituyó un salto de genio, pues nos permite conocer de forma aproximada la esquiva distribución de los números primos, especialmente para aquellas grandes escalas que superan nuestras limitaciones de cálculo. Sin embargo, no fueron Legrendre ni Gauss quienes probaron la propiedad que habían descubierto. Habría de pasar prácticamente un siglo de esfuerzos provenientes de las mentes más brillantes para que dos jóvenes matemáticos, Jacques Hadamard y Charles-Jean de la Vallée Poussin, pudieran concluir la primera demostración, cada uno de forma independiente en aquel año de 1896. Pero si esta prueba resulta curiosa es porque no se basa en las propiedades de los números primos, sino en las características analíticas de las funciones de una variable compleja, una conexión insospechada que evolucionó lenta e inadvertivamente desde una pequeña identidad ya descubierta por Leonhard Euler, pero cuyas importantes consecuencias solo serían reveladas por otros grandes matemáticos varias décadas después. Otras demostraciones sucedieron a las originales de Hadamard y de la Vallée Poussin, todas ellas utilizando este impensado nexo entre la aritmética y el análisis. Sería recién en 1948 que Atle Selberg[1] descubrió una demostración del teorema de los números primos que no utilizase herramientas procedentes del análisis real o complejo, por lo cual esta prueba ha sido llamada *elemental*, no sin falta de cierta ironía al ser más *compleja* que las pruebas que utilizan números complejos. Por lo dicho anteriormente, más allá de la importancia que contiene su enunciado, el Teorema de los Números Primos también cumplió un papel crucial en el surgimiento histórico de la Teoría Analítica de Números, del cual hoy es su joya más preciosa, y fue un importante propulsor para la investigación en Análisis Complejo.

El presente trabajo ha sido escrito para optar al grado de Licenciado en Ciencias, mención Matemáticas, y versa sobre el Teorema de los Números Primos, su demostración analítica e historia. La demostración que se presenta no pretende ser original, sino que se basa en la prueba desarrollada en Complex Analysis de Elias Stein y Rami Shakarchi [14]. Cualquier lector de dicho recomendable libro sabrá que en él las demostraciones no están completas con todo su formalismo, quedando algunas partes significativas como un ejercicio tácito al lector. El objetivo de este documento es rellenar dichos vacíos presentando las demostraciones con la formalidad más completa que me fue posible. Sin duda estos ejercicios implícitos son necesarios para que el estudiante vaya practicando los nuevos conocimientos que va adquiriendo, pero también goza de utilidad contar con una redacción más completa para que así el iniciado en matemáticas sepa lo que se espera de él y tenga con que contrastar su avance. Uno de los mayores desafíos que se presentan en la formación de un matemático es aprender a redactar, y de paso a pensar, de forma lógicamente rigurosa, un arte al cual muchos alumnos se enfrentan como cual mundo desconocido lleno de peligros. Aunque esta maduración es ante todo personal, espero que este material pueda aportar alguna luz, alguna guía en todo ese largo y a veces accidentado transitar.

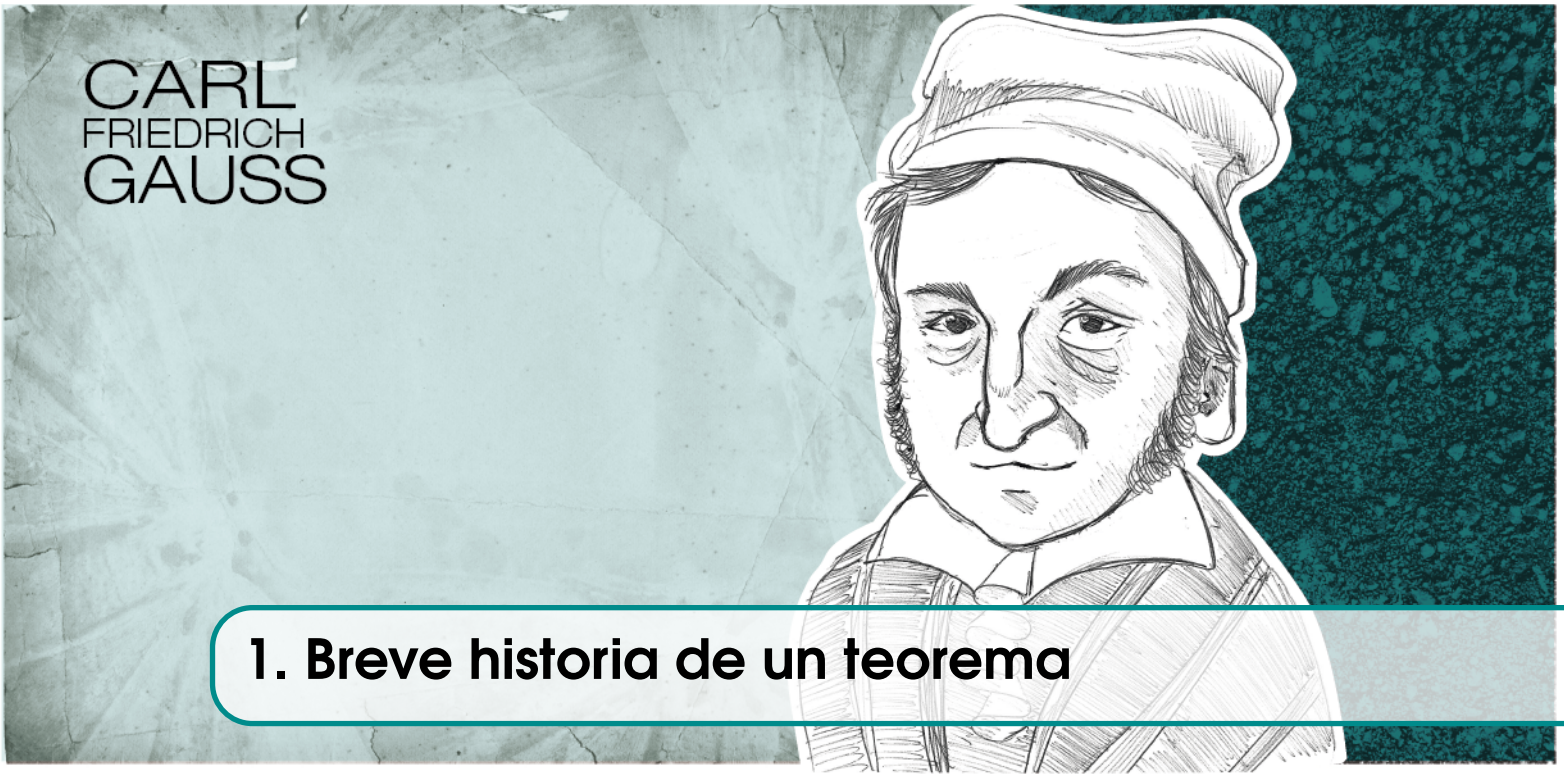
[1] Existe una lamentable disputa sobre el grado de contribución de Paul Erdös en la demostración de Atle Selberg. Ahondaremos de forma muy breve en el próximo capítulo. Al interesado, se recomienda leer [6].

Sin embargo, dicho ejercicio intelectual podría realizarse con cualquier materia, no hay diferencia, más allá de su mayor o menor complejidad, de hacerlo con un teorema u otro. Si se decidió específicamente por el Teorema de los Números Primos fue porque tiene una importancia propia que, como se explicó en párrafos anteriores y se profundizará en capítulos siguientes, proviene tanto de la elegancia y misterio del tema que trata como de su importancia histórica al haber fomentado la investigación en la teoría analítica de números y el análisis complejo. Si aún el matemático más enfocado a las aplicaciones no encuentra interés en la distribución de los primos, le será de curiosidad aprender cómo las herramientas enseñadas del análisis complejo se pueden aplicar para resolver un problema cuya interacción no se sospecha en primera vista. Nos da un entendimiento que las distintas áreas de la matemática conforman parte de un solo cuerpo que se nutre de los diversos descubrimientos procedentes desde todas las direcciones. En una época donde la matemática y la ciencia más reconocida es la interdisciplinaria, como sucede con el Programa Langlands y la Ciencia de la Complejidad, es una lección que vale que se aprenda tan pronto como sea posible. Debido a esto mismo, la búsqueda de la rigurosidad no debe perder la vista del debate apropiado de las ideas y conceptos que se están desarrollando, aunque esto exacerbe por momentos la extensión de la obra. Por lo mismo, la discusión se dará en todo el texto a través dos niveles distintos: Uno más vernacular sobre los conceptos a desarrollar y otro más formal donde se redactarán los enunciados y demostraciones. El diseño nos permitirá distinguirlos con facilidad a través del uso de enmarcaciones para el último nivel, sin que esto signifique que ambos niveles no estén en una perpetua comunicación como es debido. No hay demostración posible si se desconocen los conceptos claves, no hay debate serio sin la seguridad proporcionada por las pruebas formales.

A continuación, la obra se subdividirá en los siguientes capítulos: El primero se dedica a reseñar la historia de las ideas y esfuerzos matemáticos que condujeron al teorema y su primera demostración. El segundo versa sobre el concepto de Extensión Analítica, aplicándolo para extender la Función Gamma y concluir explorando algunas propiedades que nos serán útiles luego. El tercer capítulo, quizás el más importante, trata sobre la Función Dseta de Riemann, la principal culpable de la conexión entre números primos y valores complejos. El cuarto capítulo finalmente concluirá la demostración analítica del Teorema de los Números Primos. La obra termina con una conclusión reflexiva sobre el estudio realizado. Para la lectura apropiada de este texto es necesario tener conocimientos básicos del Análisis Complejo, como Funciones Holomorfas, Cálculo de Residuos y Teorema de Cauchy para calcular integrales. Para cuidar la extensión de la obra, se asumirán estos conceptos como conocidos, aunque se proporciona un Apéndice con teoremas importantes que son utilizados. En este sentido se acaba de perfilar la concepción última de este trabajo, como un material complementario para el estudio sobre Análisis Complejo o sobre Teoría analítica de números.

Quiero expresar mis agradecimientos a mi profesor guía Pedro Montero, cuya enorme paciencia y dedicado consejo me fueron imprescindibles para culminar este trabajo en las tumultuosos tiempos que nos tocaron. A mi novia Katherine, por brindarme su constante apoyo emocional y además ilustrar maravillosamente esta obra con retratos preciosos de los protagonistas de esta historia y con una portada que guarda la misteriosa distribución prima. Agradecer de todo corazón a mi preciosa familia, con una dedicatoria especial a mis padres y mi abuela Adriana, por haber estado a mi lado apoyándome incluso en esos momentos que ni siquiera yo hubiera estado conmigo, se los debo todo y nunca me han cobrado nada.

A la memoria de mis abuelos Domingo, Juan y Margarita, por haber creído siempre en mí.



1. Breve historia de un teorema

En la opinión del matemático universal Alexander Grothendieck, existen dos estilos de hacer matemáticas y se pueden ilustrar en cómo abrir una nuez. La primera forma, la más natural si se quiere, sería aplicar una fuerza concentrada, como con un martillo y un cincel. Se escoge el ángulo adecuado y luego se aplica el esfuerzo físico. Si la cáscara no cede, entonces se busca otro ángulo y se vuelve a insistir. Esta metáfora ilustra el procedimiento usual para demostrar teoremas: Fijar una estrategia y luego trabajar duro en deducciones y conclusiones que nos lleven a la respuesta. Si no funciona, pensar una nueva estrategia y vuelta al trabajo nuevamente. La segunda manera de hacer matemáticas se asemeja a sumergir la nuez en un recipiente con agua. El líquido circundará la nuez y lenta, pero inexorablemente, irá penetrando en la cáscara, debilitándola poco a poco hasta que, con suficiente tiempo, la nuez se abrirá hasta con un suave apretón de mano. En matemáticas esta estrategia significa ir construyendo una teoría alrededor de la conjetura, ir fabricando una inmensa maquinaria de definiciones, conceptos y relaciones que nos irán acercando paso a paso hasta que la demostración surja como una deducción evidente de esta nueva teoría.

La historia del teorema de los números primos nos instruye sobre esta segunda vía para hacer matemáticas. Resistidos todos los intentos iniciales a cincel y martillo, los matemáticos se vieron forzados a circunnavegar la conjetura para construir una teoría a su alrededor, esperando que teorema a teorema se acercaran a una demostración. Y finalmente así ocurriría: Bastó un último descubrimiento, a través de un ingenioso y elegante truco trigonométrico, para desencadenar una cascada de lemas que concluyeron en la ansiada prueba. Pero para conseguir tal cadena deductiva debió pasar todo un siglo de trabajo y colaboración matemática. Resulta un defecto un tanto frecuente en el aprendizaje matemático no atender al surgimiento histórico de las ideas estudiadas, pero la historia de las matemáticas es una fuente imprescindible para entender el quehacer de un matemático, pues necesaria es la estela de los maestros para formar el camino del aprendiz. Aunque jamás recomendable, en este caso la ausencia de reseña histórica sería impresentable, pues la historia de este teorema nos ilustra esta segunda vía de hacer matemática, el carácter colaborativo de la ciencia y la importancia de hacerse grandes preguntas, que aunque sus consecuencias no estén a la vista, pueden ser inmensas para el desarrollo del conocimiento humano.

A continuación, la historia se divide en tres partes: Una primera sobre el surgimiento de la teoría analítica de números, una segunda sobre la formulación de la conjetura y finalmente la tercera sobre el camino hacia la primera demostración.

1.1 La infinitud analítica de los números primos

Resulta desconocido desde cuando hemos logrado identificar a los números primos. Una columna del Hueso de Ishango, un utensilio congoleño de conteo que procede del Paleolítico Superior, enumera los números primos entre 10 y 20. Esta podría ser la tabla de números primos más antigua conocida, con 8 milenios de antigüedad, o solo una sorprendente coincidencia. Es probable que los matemáticos del Antiguo Egipcio y Mesopotamia conocían de los números primos, pero nuestro entendimiento formal proviene de los Antiguos Griegos y, en particular, de un libro imperecedero: Los Elementos de Euclides. Así los definía en su onceava definición del libro VII:

Un número primo es el que es medido por la unidad solamente. [9]

Cabe mencionar que los antiguos griegos fundamentaban su matemática desde la geometría, por lo que no debe extrañarnos el lenguaje geométrico presente en Los Elementos, incluso cuando se refiere a temas aritméticos. Para nuestro actual vocabulario, 'medido' significa ser 'divisible propiamente' (o sea, sin considerar al número mismo como divisor). La definición se ha mantenido a través de los milenios, aunque no sin que hayan surgido algunos desacuerdos[1]. Varias afirmaciones referidas a los números primos aparecen en los siguientes libros de Los Elementos, entre ellas algunas tan importantes como el Teorema Fundamental de la Aritmética (Proposición 14 del Libro IX). Pero, el comienzo de nuestro camino hacia el Teorema de los Números Primos está en la Proposición 20 del Libro IX: Existen infinitos números primos.

Hay más números primos que cualquier cantidad dada de números primos. [9]

La demostración de Euclides es la más conocida. Su elegante idea puede incluso reseñarse en este margen: Dese un listado finito de números primos, tan largo como usted quiera, y luego multiplique todos ellos. Puede darle un resultado tal vez pequeño, tal vez inmenso, pero siempre finito y, por tanto, tiene un sucesor. Ese sucesor no puede ser divisible por ninguno de los números primos del listado, por lo que solo hay dos posibilidades: O este sucesor es un nuevo número primo, o es divisible por un número primo no antes considerado. En cualquiera de los dos casos, hemos encontrado un número primo que no estaba en el listado. Pero ese listado que nos dimos era uno cualquiera, por lo que sin importar que tan exhaustivo seas siempre encontrarás un número primo que te falte. Si ninguna lista finita puede contener a todos los números, entonces estos son infinitos.

Sin embargo, lo que convoca nuestra atención a este teorema no es esta demostración clásica, sino por una prueba alternativa propuesta por Leonhard Euler en 1737, muchos siglos después. La demostración comienza probando una de las tantas identidades que llevan el nombre del genio suizo:

$$\sum_{n=1}^{\infty} \frac{1}{n^x} = \prod_{p \text{ primo}} \frac{1}{1-p^{-x}} \qquad \forall x > 1$$

Esta identidad se cimienta en el Teorema Fundamental de la Aritmética (TFA): Observe dicho producto, su argumento son las series geométricas $(1+p^{-x}+p^{-2x}+...)$ con p primo. Si aplicáramos la propiedad distributiva, obtendríamos una suma de todas las posibles combinaciones de productos entre potencias de primo (cada uno además elevado a $-x$). Por TFA, estos sumandos corresponderán a todos y cada uno de los números naturales (elevados a $-x$), por tanto, coincide con la serie. Por supuesto, como estamos trabajando con infinitos debe tenerse mucho cuidado con los límites, por lo que una demostración formal y su debida discusión se encuentra en el teorema 3.2.1.

[1] Al lector interesado, se le recomienda el siguiente paper [4] sobre la primalidad del número 1 a través de la historia.

La demostración de Euler sobre la infinitud de los números primos se deriva enseguida si tomamos el límite cuando $x \to 1$. El lado izquierdo diverge porque se convierte en la serie armónica, por lo que el lado derecho también debe diverger. La única manera que un producto se te vaya a infinito es que tenga infinitos términos y como hay un factor por cada primo, hay infinitos números primos.

La importancia histórica que tuvo la demostración euleriana radica en que se basó fuertemente en ideas analíticas: La infinitud de los números primos se deduce por la divergencia de la serie armónica. Esto se convirtió en toda una novedad para su época, porque hasta entonces el análisis y la aritmética parecían mundos aparte. Y sin embargo, esta demostración adolecía de un problema para constituirse en un puente entre ambas áreas: Estaba probando un hecho hace milenios conocido y demostrado bajo técnicas elementales. Si esta conexión era digna de investigación, primero requería justificar su utilidad demostrando una conjetura abierta, una que las técnicas tradicionales no se hayan mostrado capaces de resolver. Hasta que aquello sucediera, la demostración euleriana quedó limitada a ser una curiosidad matemática.

Pasaría un siglo hasta que la teoría analítica de números atravesara su bautizo de fuego. Era 1808 cuando Adrien-Marie Legendre conjeturó que existían infinitos números primos en algunas sucesiones aritméticas. Recordando: Una serie aritmética son aquellas donde los términos de la sucesión están separados por la misma cantidad. Podemos representar a todos los elementos de dicha sucesión con la fórmula $a \cdot n + b$ con a y b constantes y $n = 0, 1, \ldots$. Si tanto a como b son naturales, es suficiente para que la sucesión sea de números naturales. Resulta evidente de la fórmula que, si existe un divisor común para a y b, todos los elementos de la sucesión serán divisibles por ese divisor común, por lo que el único número primo puede ser el primero de la sucesión. Pero si a y b son coprimos, es decir, no tienen divisores comunes entre sí, no es evidente cuántos números primos puede haber en la sucesión. Adrien-Marie Legendre conjeturó que siempre que a y b sean coprimos, habría infinitos números primos en la sucesión.

En 1837 llegaría la demostración a manos del matemático alemán Peter Gustav Dirichlet. Inspirada en la demostración euleriana, Dirichlet enfocó su prueba a demostrar que, para $a, b \in \mathbb{N}$ coprimos, la suma

$$\sum_{\substack{p \text{ primo} \\ p \in \{a \cdot n + b\}}} \frac{1}{p}$$

diverge[2], lo que implicaría que hay infinitos sumandos, por tanto, infinitos números primos en la sucesión aritmética $\{a \cdot n + b\}$. No ahondaremos en esta demostración, aunque si mencionaremos su estrategia: Dirichlet busca una expresión equivalente a la serie

$$\sum_{\substack{p \text{ primo} \\ p \in \{a \cdot n + b\}}} \frac{1}{p^s} \qquad s > 1$$

que le permita demostrar que diverge a ∞ cuando $s \to 1$. Para encontrar esa expresión, estudia una generalización de las series consideradas por Leonhard Euler y que actualmente se llama función L de Dirichlet:

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

Donde χ es un carácter de Dirichlet. Al lector interesado en esta demostración, existe una prueba moderna en el clásico [1] que se invita a profundizar.

[2]En otras demostraciones actuales, la suma escogida es $\sum \log p/p$ en vez de $\sum 1/p$, como es el caso de [1], pero el enfoque de la demostración es la misma.

La demostración del hoy llamado Teorema de Dirichlet tuvo importantes repercusiones en la investigación matemática. No solo introdujo conceptos nuevos como la función L de Dirichlet o sus carácteres, que son profundamente estudiados hasta el día de hoy, sino también fue la primera vez que una conjetura que estaba en la primera línea de la investigación en teoría numérica fue demostrada gracias a técnicas analíticas. Esto inspiró a otros matemáticos a utilizar el análisis para resolver problemas aritméticos, incluyendo a Benhard Riemann en su trabajo pionero sobre el Teorema de los Números Primos. Las conexiones entre el análisis y la aritmética pasaron de ser una simple curiosidad a ser hasta hoy uno de los campos importantes de la investigación matemática. Con Dirichlet, finalmente, había nacido la Teoría Analítica de Números.

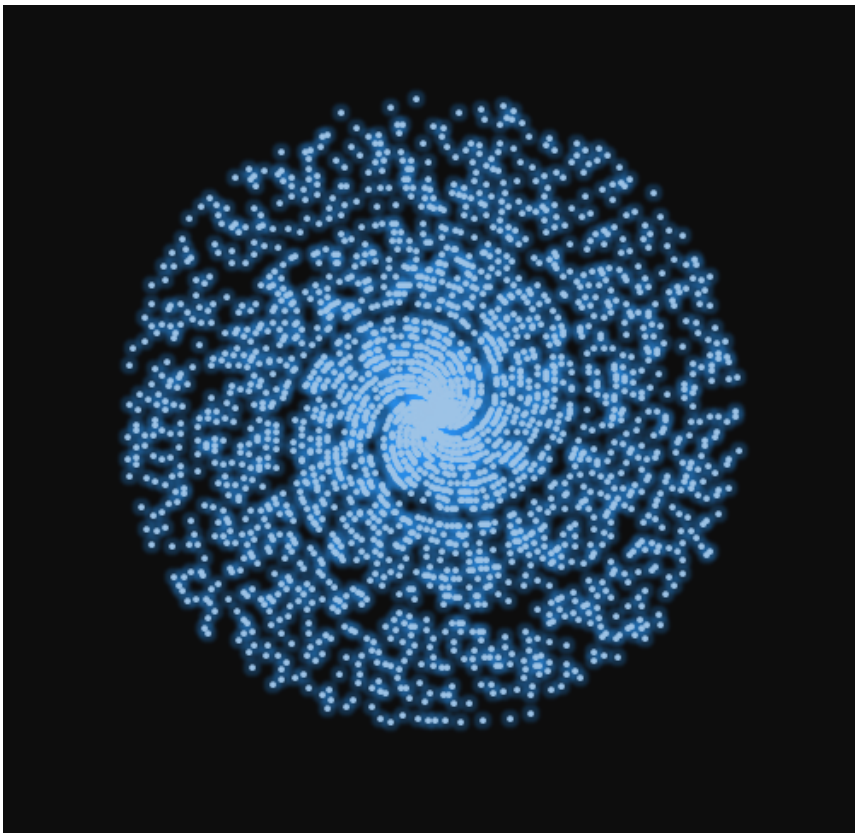


Figura 1.1: Gráfica de los puntos (p,p) en coordenadas polares, con p primo menor a 25000. Las espirales corresponden a cada una de las 20 sucesiones aritméticas $44 \cdot n + b$, donde b es coprimo de 44. Que estas espirales jamás acaben es una bella ilustración del **Teorema de Dirichlet**. Fuente: Idea original de Dwymark [5] y explicada brillantemente en 3Blue1Brown [12]. Gráfica propia.

1.2 Conjeturando desde una tabla de números primos

Era el año 1792 o 1793. Aunque con solo 15 años y de origen humilde, Carl Friedrich Gauss ya gozaba de una fama local como joven virtuoso. Recibía desde hace años instrucciones adicionales en matemáticas e incluso un estipendio de parte del Duque de Brunswick, a quien había dejado muy impresionado. Dentro de este ambiente educativo tan especial, Gauss empezó sus investigaciones personales sobre alta aritmética. Uno de sus primeros proyectos fue estudiar la disminución de la frecuencia de los números primos. En notación actual, Gauss estudiaba como decrece la función $\pi(x)/x$, donde $\pi(x)$ es la cantidad de números menores que x. Su interés no era predecir el próximo número primo, sino saber aproximadamente cuantos de los naturales menores a x son primos.

Gauss empezó a calcular dicha tasa $\pi(x)/x$ en intervalos de mil en mil. A pesar de las fluctuaciones que se le presentaban, fue capaz de observar que dicha frecuencia se iba aproximando al recíproco multiplicativo del logaritmo natural. Para nuestra moderna notación, este hallazgo se escribe:

$$\frac{\pi(x)}{x} \sim \frac{1}{\log x}$$

Donde $\sim$ significa asintóticamente equivalentes, es decir, se cumple el siguiente límite:

$$\lim_{x\to\infty} \frac{(\pi(x)/x)}{(1/\log x)} = 1$$

A nivel conceptual, esto nos dice que a medida que x va creciendo, ambas funciones empiezan a comportarse de forma similar. Esto no quiere decir que su diferencia se vaya haciendo cero, pero sí que este error se va haciendo insignificante a comparación de x.

Ya esta observación era de por sí remarcable y es a la que nos enfocaremos de demostrar en este libro. Sin embargo, Gauss fue más allá: Verá que existen dos formas de interpretar a $\pi(x)/x$. La primera, que hemos usado hasta ahora, es como la tasa de cuántos números naturales menores que x son números primos. Pero la segunda interpretación es pensar en $\pi(x)/x$ como una aproximación a una razón de cambio promedio de $\pi(x)$. Por supuesto esta no es la derivada de $\pi(x)$, pues esta función crece a saltos discontinuos, luego su derivada es 0 salvo en los puntos que no está definida y donde la función realmente cambia. Sin embargo, captura la idea de una razón de cambio global y, sin ser rigurosos, a través del teorema fundamental del cálculo podemos sospechar que la expresión anterior nos está sugiriendo lo siguiente:

$$\pi(x) \sim \int_2^x \frac{dx}{\log x}$$

Esta es una segunda formulación del Teorema de los Números Primos. Es equivalente a la anterior, pero funciona como una mejor aproximación para $\pi(x)$. El lado derecho se le llama la integral logarítmica desplazada, denotada $\mathrm{Li}(x)$.

Gauss evaluó el funcionamiento de su hipótesis utilizando las tablas de logaritmos creadas por Lambert. Años después, iría ampliando sus cálculos según mejores recursos tenía a su disposición: En 1796, tras obtener las tablas de Vega, testeó su hipótesis hasta $x = 400{,}031$. En 1811 amplió sus calculos para llegar al primer millón. Tiempo después, con la asistencia de su estudiante Carl Goldschmidt y haciendo uso de las tablas de Burkhardt, llevo sus cómputos hacia los tres millones. En la siguiente tabla se muestra los resultados obtenidos por Gauss, incluyendo sus ligeros errores al calcular $\pi(x)$ si comparamos con el valor informado por Wolfram Alpha. La perseverancia y precisión de los cálculos de Gauss es desconcertante para una época previa a los ordenadores:

x	$\pi(x) \pm$ Error	$\mathrm{Li}(x) \pm$ Error
500.000	41.556 - 18	41.606,4 $\pm$ 50,4
1.000.000	78.501 - 3	79.627,5 $\pm$ 126,5
1.500.000	114.112 + 43	114.263,1 $\pm$ 151,1
2.000.000	148.883 + 50	149.054,8 $\pm$ 171,8
2.500.000	183.016 + 56	183.245,0 $\pm$ 229,0
3.000.000	216.745 + 71	216.970,6 $\pm$ 225,6

Cuadro 1.1: Cálculos de Gauss sobre el teorema de números primos. Fuente: Goldstein [7].

Resultan impresionantes los logros conseguidos por Gauss, pero, lamentablemente, Gauss jamás publicó ninguno de estos hallazgos. Nuestro conocimiento de sus investigaciones se basa en una carta de 1849 dirigida a su amigo, el astrónomo y ex alumno suyo Johann Encke. Este hecho no es inusual en la biografía de Gauss, en varias ocasiones adelantó a los matemáticos de su época, pero no hizo públicos sus descubrimientos. Así sucedió, por ejemplo, sobre el nacimiento de las geometrías no euclidianas donde Gauss precedió secretamente en décadas a Bolyái y Lobachevsky. Son pocas las dudas sobre la veracidad de las afirmaciones de Gauss, pero para más certeza dicho trabajo privado también es citado por Benhard Riemann en su memorable *Sobre el número de primos menores que una cantidad dada*, del que hablaremos más adelante.

La primera publicación de la conjetura quedaría en manos del gran matemático francés Adrien-Marie Legendre, quien lo incluiría en su *Essai sur la Théorie des Nombres* en 1798, considerado el primer libro de texto en Teoría de Números. En éste se plantea la siguiente conjetura sobre la distribución de los números primos:

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

Con A y B algunas constantes. Diez años después, en la segunda edición de su libro[3], refinaría su hipótesis como sigue:

$$\pi(x) \sim \frac{x}{\log x - 1{,}08366}$$

A semejanza de Gauss, Legendre sustentó su hallazgo con evidencia numérica. Usando las tablas de Vega, evaluó su conjetura hasta $x = 400{,}000$ publicando estos resultados en su libro. Parte de dicha tabla se incluye a continuación, con el cálculo del error respecto a valores modernos proporcionados por Wolfram Alpha. También se incorpora una estimación adicional para $x = 1{,}000{,}000$ que también Legendre proporcionó en su libro. Nuevamente la precisión y extensión son dignas de alabanza:

x	$\pi(x) \pm$ Error	$x/(\log x - 1{,}08366) \pm$ Error
50.000	5.134 - 1	5.136 - 0,48
100.000	9.592	9.588 + 0,40
150.000	13.849 - 1	13.844 + 0,37
200.000	17.984	17.982 - 0,29
250.000	22.045 - 1	22.035 + 0,06
300.000	25.998 - 1	26.025 - 1,13
350.000	29.977	29.961 - 0,45
400.000	33.861 - 1	33.854 - 0,34
1.000.000	78.527 - 29	78.543 + 0,18

Cuadro 1.2: Cálculos de Legendre en *Théorie des Nombres*. Fuente: Klyve [11].

Resulta intrigante cómo se puede derivar este teorema desde una tabla de números primos. Si uno observa dichas tablas, rápidamente se da cuenta de lo difícil que es encontrar algún patrón, especialmente si no sabe de antemano lo que se está buscando. Lamentablemente no hay información de cómo Gauss o Legendre dedujeron su hipótesis, al fin de cuentas los estudios de Gauss eran privados y de la vida de Legendre se conserva muy poco[4]. Para nuestra suerte, Apostol [2] sugirió una buena manera de concluir la conjetura a partir de estas tablas, que se ahonda a continuación:

[3] Donde también conjeturaría el futuro Teorema de Dirichlet ya mencionado.

[4] Por no ir más lejos, no se conserva un retrato completo de Legendre. La imagen de la introducción se basa en una caricatura de 1820, la única imagen que nos queda del genio francés.

Comencemos desde el enfoque de las investigaciones del joven Gauss centrando el estudio en la tasa $\pi(x)/x$. Gauss estudió sus valores en intervalos de mil hasta los primeros tres millones, para hacer más evidente la deducción en la siguiente tabla usaremos potencias de mil extendiéndonos a los cómputos más recientes[5]. No resulta sorprendente que la razón $\pi(x)/x$ sea decreciente, por lo que es usual indagar en su inverso multiplicativo $x/\pi(x)$ en la búsqueda de posibles patrones, que incluimos en la presente tabla. Si observamos dicha columna, podemos notar que $x/\pi(x)$ incrementa de forma aritmética a medida que crecen los exponentes, lo que se confirma si observamos que los incrementos registrados en la cuarta columna son aproximadamente constantes. Esto nos insinúa que estamos en la presencia de un comportamiento logarítmico, pero ¿cómo podríamos saber que se trata del logaritmo en base natural? Observemos que el incremento en cada paso es de aproximadamente $6,9$, esto quiere decir $2,3$ veces el incremento del exponente en cada paso (como confirma la quinta columna). Aunque el $2,3$ nos parezca misterioso en tiempos modernos, en la época de Gauss era bien conocido que $2,3$ es aproximadamente el logaritmo natural de 10 ya que este es su factor de escalamiento, recordemos que por el teorema de cambio de base $\log(x) = \log(10) \cdot \log_{10}(x)$ y $\log_{10}(x)$ coincide con los exponentes que incrementamos en cada caso. **En resumen:** La idea ganadora está en el mismo enfoque de Gauss de estudiar la tasa $\pi(x)/x$, pues si se observa su inverso -algo natural que hacer ante una sucesión decreciente-, el comportamiento logarítmico se revela por las características inusuales de su naturaleza y solo queda despejar el factor de escalamiento para saber de qué logaritmo se trata. Es posible que el joven Gauss o Legendre hayan pasado por un proceso similar; por ejemplo, cuando observaron los valores para mil, diez mil y cien mil; sin embargo sin mayor evidencia histórica no podemos ni debemos asegurarlo. Posteriormente Gauss y Legendre confirmaron que habían encontrado algo al corroborar numéricamente la increíble similitud entre $\pi(x)/x$ y $1/\log(x)$ haciendo uso de las tablas logarítmicas a su disposición, lo que nosotros hacemos en las última dos columnas.

x	$\pi(x)/x$	$x/\pi(x)$	Δ	$\Delta/3$	$1/\log(x)$	Error (%)
10^3	0,168	5,9524			0,14476	14 %
10^6	0,0784986	12,7391	6,7867	2,2622	0,07238	8 %
10^9	0,050847534	19,6666	6,9276	2,3092	0,04825	5 %
10^{12}	0,037607912018	26,5901	6,9235	2,3078	0,03619	4 %
10^{15}	0,029844570422669	33,5069	6,9168	2,3056	0,02895	3 %
10^{18}	0,024739954287740860	40,4204	6,9135	2,3045	0,02413	2 %
10^{21}	0,021127269486018731928	47,3322	6,9117	2,3039	0,02068	2 %
10^{24}	0,018435599767349200867866	54,2429	6,9107	2,3036	0,01810	2 %
10^{27}	0,016352460426841680446427399	61,1529	6,9100	2,3033	0,01608	2 %

Cuadro 1.3: Tablar para deducir de la conjetura. Fuente: Weisstein. [15]

1.3 El complejo camino a la demostración

Aun cuando Gauss y Legendre hicieron increíbles esfuerzos para obtener evidencia numérica para soportar su conjetura, no hicieron avances públicos hacia la demostración. Aunque en general los ejemplos numéricos solo sirven para una refutación de una sentencia sobre un conjunto infinito, en este caso una hipótesis sobre el comportamiento al infinito de una función solo puede decidirse con argumentos lógicos incluso para su forma negativa.

[5] Todos los datos los presentamos en esta tabla al cuarto decimal de precisión salvo $\pi(x)/x$ que lo ponemos con todos sus decimales por si alguien quisiera saber los valores exactos para $\pi(x)$.

Los primeros esfuerzos matemáticos para resolver la conjetura utilizaron el primer enfoque relatado por Grothendieck: Fueron a martillar la nuez. De todos estos intentos directos, el más excelso fue realizado por el matemático ruso Pafnuti Tchebyshev[6] duante los años 1848 a 1852. El mayor problema que tenía la conjetura de los números primos es que $\pi(x)$ es una función excesivamente misteriosa, no se tiene una expresión sencilla para trabajar con ella, por lo que Tchebyshev partió por "cambiar" el problema a un enunciado equivalente, pero más abordable. Para ello, Tchebyshev definió las funciones que hoy llevan su nombre:

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

$$\psi(x) = \sum_{p^m \leq x} \log p$$

La conexión de estas funciones con el teorema de los números primos proviene de que los siguientes enunciados se pueden probar equivalentes:

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

$$\vartheta(x) \sim x$$

$$\psi(x) \sim x$$

Las demostraciones proceden de acotar apropiadamente las funciones para que los límites, si es que existen, deban coincidir por encajamiento. Una muestra de aquello se puede encontrar en el teorema 4.2.2. La motivación de Tchebyshev es que, al cambiar de π por ψ o ϑ, se pueden aprovechar las propiedades de logaritmo y aproximaciones tales como la fórmula de Sterling para investigar el comportamiento de la función. No profundizaremos en estas técnicas durante este trabajo más allá del referido teorema 4.2.2, pero cabe señalar que esta astucia permitió a Tchebyshev demostrar la siguiente desigualdad a través de métodos elementales:

$$0{,}92129 \leq \liminf_{x\to\infty} \frac{\pi(x)\log(x)}{x} \leq 1 \leq \limsup_{x\to\infty} \frac{\pi(x)\log(x)}{x} \leq 1{,}10535$$

Esto tiene una consecuencia directa: Si el límite de $\pi(x)\log(x)/x$ existe, entonces tiene que ser 1. ¡Tchebyshev ya había recorrido medio camino hacia la demostración! Aunque el genio ruso no logró probar que dicho límite exitiera, su aproximación le permitió probar el Teorema de Bertrand: Para cada $n \in \mathbb{N}$, existe un primo p entre n y $2n$. En 1881, el matemático británico James Joseph Sylvester perfeccionó los límites de esta desigualdad encontrada por Tchebyshev, colocando de cota inferior 0,95695 y de cota superior 1,04423. Así se refería Sylvester en el American Journal of Mathematics respecto al estado del problema y los logros de Tchebyshev:

> Probablemente tendremos que esperar a que nazca en el mundo alguien que supere con creces a Tchebyshev en intuición y profundidad como Tchebyshev se ha probado superior en estas cualidades al común de la humanidad. [2]

Sylvester no tenía como sospechar que, no solo el resto de nuestros protagonistas ya habían nacido, sino que solo quedaban 10 años para la primera demostración. Pero en nuestro camino debemos retroceder a mediados de siglo, a cuando aún vivía el artífice principal de la demostración: Benhard Riemann, matemático alemán discípulo del propio Gauss, no alcanzó a vivir 40 años, pero marcó para siempre la historia de las matemáticas, otorgando ideas fundacionales al análisis real y complejo, la geometría y la teoría numérica. Como diría Hawking: *"a pesar de la brevedad de su obra, o quizás debido a ella, se puede decir que todo lo que tocó revolucionó las matemáticas"* [8].

[6] Como suele suceder desde el cirílico ruso, su apellido no tiene una traducción única. En la literatura se encontrará con o sin T, con v o doble f, con s o c en el shev, etcétera. Traducir un nombre no es una función, lamentablemente.

Sin lugar a dudas, el enfoque de Riemann a la demostración del teorema de los números primos se gana el apelativo de revolucionario. Como el agua rodeando la nuez, Riemann profundizó en la identidad ya descubierta por Euler:

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ primo}} \frac{1}{1-p^{-s}}$$

Esta propiedad se mantiene incluso cuando s es un número complejo, siempre que su parte real sea mayor estrictamente que 1. Más todavía la serie define una función analítica en $\mathrm{Re}(s) > 1$, es decir, posee una derivada compleja para todo $\mathrm{Re}(s) > 1$. Una propiedad crucial de las funciones analíticas es que, si existe una extensión que también sea analítica, esta es única. Profundizaremos más en este concepto en el capítulo siguiente, pero esto permite que exista una única forma de expandir la función a dominios más grandes, siempre que conservemos la deseable propiedad de analiticidad. El primer hallazgo trascendental de Riemann fue que la serie de la identidad de Euler puede extenderse analíticamente para todo $\mathbb{C} - \{1\}$. La demostración de este hecho dedicaremos la primera sección del capítulo 3, pero esencialmente se hace encontrando una función que extiende a $\mathbb{C} - \{0,1\}$ a la expresión:

$$\pi^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) \sum_{n=1}^{\infty} \frac{1}{n^s}$$

Esto permite definir una extensión para la serie en todo $\mathbb{C} - \{1\}$, a la que actualmente se le llama Función Dseta de Riemann y se le denota con la letra griega ζ. Una propiedad importante que se deriva de esta demostración es que la extensión de $\pi^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) \zeta(s)$ es simétrica bajo el eje Re(s) = 1:

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

Despejando se puede encontrar una relación explícita entre $\zeta(s)$ y $\zeta(1-s)$. Esta se suele presentar más simplificada al usar la fórmula de reflexión de Euler para la función gamma (el teorema 2.3.3):

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

Esta expresión es llamada la fórmula funcional de Riemann. ¿Y por qué es tan importante? Pues como se verá con más notoriedad en el capítulo 3, la extensión que da pie a la función dseta contiene integrales impropias difíciles de resolver analíticamente, pero con la fórmula funcional podemos estudiar la función dseta en $\mathrm{Re}(s) < 0$ observando su comportamiento en $\mathrm{Re}(s) > 1$, donde corresponde a la serie con que comenzamos y que es a veces más sencilla de estudiar. Lamentablemente esto deja incógnita una franja entre $\mathrm{Re} = 0$ y $\mathrm{Re} = 1$, que es denominada la franja crítica y veremos que será crucial luego.

Hasta ahora todo muy bien (o todo muy mal) con las investigaciones de Riemann sobre la función dseta, pero el lector puede estar desesperándose de no saber qué utilidad guarda para demostrar el teorema de los números primos. Riemann descubrió una inesperada conexión que se esboza a continuación, aunque cabe advertir que el genio alemán formuló sus ideas de forma distinta a lo que haremos. Primero porque Riemann, fiel discípulo de Gauss, usaba la formulación de la conjetura usando $\mathrm{Li}(x)$[7]. Segundo porque, como suele pasar en los trabajos pioneros, la formulación de Riemann es más complicada y, remitiendo la opinión de Apostol [2], muchos detalles técnicos requieren de mejor justificación. El esbozo seguirá el camino establecido por el matemático también alemán Hans von Mangoldt en 1894 y que se utiliza mayormente en los tratamientos modernos.

[7] Y en un detalle aparte, Riemann también prefería usar la función Π a Γ, una notación también bastante gaussiana.

Comencemos con la identidad de Euler:

$$\zeta(s) = \prod_{p \text{ primo}} \frac{1}{1-p^{-s}} \qquad \forall s \in \mathbb{C} : \mathrm{Re}(s) > 1$$

Si se aplica el logaritmo natural queda:

$$\log \zeta(s) = -\sum_{p \text{ primo}} \log\left(1-p^{-s}\right)$$

Si se deriva y cambia signo se logra:

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{p \text{ primo}} \frac{p^{-s}\log p}{1-p^{-s}}$$

Observando que lo que acompaña al logaritmo es una serie geométrica:

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{p \text{ primo}} \sum_{m=1}^{\infty} \frac{\log p}{p^{ms}}$$

Sin abordar toda la formalidad, que ya nos ocuparemos con prolijidad en la demostración de la proposición 4.3.1, podemos ver esta doble serie como una serie en números naturales donde el denominador es n^s y el numerador es $\log p$ si n es una potencia de p primo, o es 0 en cualquier otro caso. Denotando a este numerador alternante como $\Lambda(n)$, tenemos:

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

Ahora vamos a conectar con la función ψ de Tchebyshev. Si observamos su definición, queda claro que para un h muy próximo a 0, $d\psi(x) = \psi(x) - \psi(x-h)$ vale $\log p$ si x es la potencia de p primo o 0 en otro caso. Es decir, se comporta como nuestro numerador. Por lo que podemos ver la serie como una Integral de Riemann-Stieljes:

$$-\frac{\zeta'(s)}{\zeta(s)} = \int_1^{\infty} \frac{d\psi(x)}{x^s}$$

Haciendo integración por partes y reordenando términos nos queda:

$$\frac{\zeta'(s)}{\zeta(s)} = \int_1^{\infty} \frac{s\psi(x)}{x^{s+1}}dx \Longleftrightarrow \frac{\zeta'(s)}{s\zeta(s)} = \int_1^{\infty} x^{s-1}\left(\frac{\psi(x)}{x^{2s}}\right)dx$$

Esta última integral es conocida en las matemáticas: Es una Transformada de Mellin, de la función $x^{-2s}\psi(x)$ en este caso, y tiene una transformación inversa que al aplicarla queda:

$$\frac{\psi(x)}{x^{2s}} = \int_{2-i\infty}^{2+i\infty} x^{-s}\left(\frac{\zeta'(s)}{s\zeta(s)}\right)ds \Longleftrightarrow \psi(x) = \int_{2-i\infty}^{2+i\infty} \frac{x^s}{s}\left(\frac{\zeta'(s)}{\zeta(s)}\right)ds$$

Esta relación deja a la función ψ de Tchebyshev escrita en términos de la función dseta y su derivada. Pero la genialidad de Riemann fue que, al extender la función dseta en todo $\mathbb{C}-\{1\}$, podemos utilizar las herramientas del análisis complejo para estudiar esa integral. En particular, el Teorema Integral de Cauchy considera que toda integral en un camino cerrado de una función analítica tanto en el camino como en el interior es 0. Pero tenemos dos asuntos para aplicar este teorema: Primero la integral no es sobre un camino cerrado. Esto se arregla conceptualizando que dicha integral es el lado derecho de un rectángulo compuesto por $\mathrm{Re}(s) = 2$, $\mathrm{Im}(s) = \infty$, $\mathrm{Im}(s) = -\infty$ y $\mathrm{Re}(s) = -\infty$. Esto obviamente es una imagen heurística, en realidad estamos calculando la integral sobre un rectángulo cuando los límites de integración tienden al infinito. Es posible mostrar a través de acotaciones que el argumento de la integral se hace 0 para los "otros lados de este rectángulo infinito", por lo que la integral que tenemos coincide con la integral de este rectángulo infinito:

$$\psi(x) = \lim_{R\to R\infty} \int_R \frac{x^s}{s}\left(\frac{\zeta'(s)}{\zeta(s)}\right)ds$$

El segundo problema proviene que la función que integramos no es analítica en todos los puntos del interior sino que es meromorfa, es decir, tiene unos polos donde se va a infinito. Pero de la teoría del cálculo de residuos, esto es arreglable con que el valor de la integral no sea cero, sino que corresponda a los residuos de los polos de la función. Por lo que el cálculo de la integral impropia se ha simplificado en conocer dónde están sus polos y cuáles son sus residuos.

Sobre la primera pregunta, los polos se encuentran en $s = 0$ donde se indefine $\frac{1}{s}$, en $s = 1$ donde se indefine $\zeta'(s)$ y en los ceros de la función dseta. Riemann mostró la utilidad de la fórmula funcional para buscar los ceros de la función dseta. Primero, la identidad de Euler nos permite demostrar que $\zeta(s) \neq 0$ en $\mathrm{Re}(s) > 1$. Este hecho se probará con formalidad en el teorema 3.4.1, pero en resumen la propiedad que un producto no es cero si ninguno de sus factores lo es se puede aplicar para ciertos productos infinitos como el presente en la identidad de Euler. Seguido a ello, si observamos la fórmula funcional de Riemann:

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

Para $\mathrm{Re}(s) < 0$, como $\zeta(1-s) \neq 0$ por lo anteriormente dicho, los únicos ceros de la región estarán donde $\sin\left(\frac{\pi s}{2}\right)\Gamma(1-s) = 0$. Estos ceros se dan en los pares negativos $s = -2, -4, -6, \ldots$ y son conocidos como los ceros triviales de la función dseta. Lamentablemente la fórmula funcional no nos descubre los ceros en la franja crítica $0 \leq \mathrm{Re}(s) \leq 1$, de hecho no hay forma sencilla de estudiarlos. A estos ceros se le denomina los ceros no triviales de la función dseta.

Ya descubierto donde están los polos, queda estudiar los residuos. Recordemos del cálculo de residuo que si el siguiente límite existe se trata de un polo simple y su residuo es:

$$\mathrm{Res}(f, z) = \lim_{s \to z} (s - z) f(s)$$

Donde f es la función meromorfa con polo en $s = z$. La extensión encontrada por Riemann para definir a la función dseta hace fácil el cálculo de los residuos de $\frac{x^s}{s}\left(\frac{\zeta'(s)}{\zeta(s)}\right)$, ya que la función dseta tiene la forma:

$$\zeta(s) = \frac{1}{s-1} + a_0 + a_1(s-1) + a_2(s-1)^2 + \ldots$$

Y por ende su derivada es de la forma:

$$\zeta'(s) = -\frac{1}{(s-1)^2} + a_1 + 2a_2(s-1) + \ldots$$

Esto permite saber que todos los polos son simples y sus residuos son:

$$\mathrm{Res}\left(\frac{x^s}{s}\left(\frac{\zeta'(s)}{\zeta(s)}\right), 1\right) = x$$

$$\mathrm{Res}\left(\frac{x^s}{s}\left(\frac{\zeta'(s)}{\zeta(s)}\right), 0\right) = -\frac{\zeta'(0)}{\zeta(0)}$$

$$\mathrm{Res}\left(\frac{x^s}{s}\left(\frac{\zeta'(s)}{\zeta(s)}\right), -2n\right) = \frac{x^{-2n}}{2n} \qquad n \in \mathbb{N}$$

$$\mathrm{Res}\left(\frac{x^s}{s}\left(\frac{\zeta'(s)}{\zeta(s)}\right), \rho\right) = -\frac{x^\rho}{\rho}$$

Donde ρ es un cero no trivial de la función dseta, o sea, $\zeta(\rho) = 0$ y $\mathrm{Re}(\rho) \in [0, 1]$.

Al calcular la integral impropia al final de la página 18 a través de sumar todos los residuos, las ideas de Riemann decantan en la siguiente fórmula explícita para la función *psi* de Tchebyshev:

$$\psi(x) = x - \frac{\zeta'(0)}{\zeta(0)} + \sum_{n=1}^{\infty} \frac{x^{-2n}}{2n} - \sum_{\rho} \frac{x^{\rho}}{\rho}$$

Esta es la culminación de las ideas de Riemann sobre el Teorema de los Números Primos y hay que observala con atención. Podemos hacerle algunos arreglos, primero observando que el tercer elemento de la derecha es la serie de Taylor de $-\log(1-x^{-2})/2$:

$$\psi(x) = x - \frac{\zeta'(0)}{\zeta(0)} - \frac{\log(1-x^{-2})}{2} - \sum_{\rho} \frac{x^{\rho}}{\rho}$$

Luego si dividimos por x:

$$\frac{\psi(x)}{x} = 1 - \frac{1}{x}\left(\frac{\zeta'(0)}{\zeta(0)}\right) - \frac{\log(1-x^{-2})}{2x} - \sum_{\rho} \frac{x^{\rho-1}}{\rho}$$

Uno de los enunciados equivalentes de *Tchebyshev* era que $\psi(x) \sim x$, o sea, $\psi(x)/x \to 1$ cuando $x \to \infty$. Por lo que solo nos queda tomar el límite y de hecho ya podemos ver el 1 en el lado derecho (otorgado por el residuo del polo en $s = 1$). Es fácil de ver que el segundo y el tercer elemento del lado derecho tienden a 0. Por lo que todo la demostración queda resuelta si

$$\lim_{x\to\infty} \sum_{\rho} \frac{x^{\rho-1}}{\rho} = 0$$

Riemann desconocía si la función dseta tenía infinitos ceros no triviales. Hoy día lo sabemos. Por lo que este es un doble límite y se debe tener cierto cuidado formal. Pero de forma intuitiva, aunque formalizable, podemos ver que dado:

$$\left|\frac{x^{\rho-1}}{\rho}\right| = \frac{x^{\mathrm{Re}(\rho)-1}}{|\rho|}$$

Los términos de la serie deberían irse a 0 cuando la parte real de ρ es menor que 1, o sea, Riemann ha reducido con éxito la conjetura de los números primos a demostrar que la función dseta no tiene ceros en $\mathrm{Re}(s) = 1$. Lamentablemente, a pesar del inmenso y prodigioso trabajo que hizo Benhard Riemann, no logro demostrar este último paso. Eso sí Riemann calculó a mano una gran cantidad de ceros no triviales para poder estudiar su comportamiento y llegó a la hoy famosa Hipótesis de Riemann que señala que la parte real de todo cero no trivial de la función dseta es $\frac{1}{2}$. Hoy en día esa hipótesis es considerada por muchos el problema matemático abierto más importante de nuestra época y tiene una bastedad de consecuencias si llega a probarse verdad. Riemann publicó todo este trabajo magistral en una única obra titulada *"Ueber die Anzahl der Primzahlem unter einer gegebenen Grösse"* (Sobre el número de primos menores que una cantidad dada) para el año 1859, así celebrando su nominación a la Academia de Ciencias de Berlín. 3 años después, Riemann enfermaría de pleusería gravemente y fallecería en 1866 por dicha enfermedad. Dicha magna obra fue su única contribución a la Teoría Numérica, una totalmente revolucionaria que la cambiaría para siempre, y sería el último de sus grandes trabajos. Queda solo a la especulación qué todavía más Riemann hubiera dado al mundo si la vida le hubiera dado más tiempo.

Las matemáticas, como la vida, siguen, y el último paso dejado por Riemann permanecía. Como pasa usualmente en la historia de la ciencia, grandes ideas ocurren en distintas partes casi simultáneamente y ese es el caso de la primera demostración: Los matemáticos treinteañeros Jacques Hadamard (de Francia) y Charles-Jean de la Vallée Poussin (de Bélgica) lograron finalizar la demostración del Teorema de los Números Primos de forma independiente en 1896, culminando prácticamente un siglo de brillantes esfuerzos matemáticos.

Ambas demostraciones se centraron en demostrar el último paso dejado por Riemann: Que no hay ceros no triviales de la función dseta en la recta $\mathrm{Re}(s) = 1$, pero utilizaron métodos totalmente diferentes. La demostración de Jacques Hadamard sería publicada en un paper del Boletín de la Sociedad Matemática Francesa y se basa en la relación entre $\zeta(\sigma + i\tau)$ y $\zeta(\sigma + 2i\tau)$. Una curiosidad destacada por Apostol [2] es que, en un acto de honor y diplomacia, Jacques Hadamard reconoce que el matemático holandés Thomas Stieljes tendría una demostración sin publicar de la Hipótesis de Riemann, lo que demostraría también el Teorema de los Números Primos. Sin embargo, esa demostración de la Hipótesis de Riemann no se concretó.

Por otro lado, De la Vallée Poussin publicó tres memorias sobre teoría de números: La primera contiene su demostración sobre el Teorema de Números Primos, la segunda extiende su método para demostrar un teorema de números primos para sucesiones aritméticas y la tercera trata sobre primos representables con formas cuadráticas. La demostración De la Vallée Poussin indaga más en las propiedades de la función dseta para demostrar que $\zeta(\sigma + i\tau) \neq 0$ en la región $\sigma > 1 - A/\log(\tau)$ para un τ suficientemente grande. De la Vallée Poussin también demuestra extensiones del Teorema de los Números Primos como la siguiente para $n \in \mathbb{N}$:

$$\pi(x) \sim \frac{x}{\log(x)} + \frac{1! \cdot x}{\log^2(x)} + \ldots + \frac{n! \cdot x}{\log^n(x)}$$

O la siguiente extensión que indica que el número de primos menores que x en una sucesión aritmética $a \cdot n + b$, con $a, b \in \mathbb{N}$ coprimos, se distribuye asintóticamente:

$$\sim \frac{x}{\phi(a)\log(x)}$$

donde $\phi(a)$ es la función de Euler que cuenta la cantidad de coprimos de a existen entre 1 y a.

Al final de su tercera memoria, De la Vallée Poussin reconoce que la demostración de Hadamard es más simple, pero también agrega que se puede simplificar con un truco trigonométrico. Esta versión simplificada de la demostración de Hadamard por De la Vallée Poussin es la que utilizaremos en el teorema 3.4.2, aunque el truco lo usaremos antes en el lema 3.2.4. Ahondaremos en detalle durante el capítulo 3, pero en un breve resumen podemos demostrar con una brillante manipulación de los logaritmos complejos y la identidad de Euler que:

$$\log\left|\zeta^3(x)\zeta^4(x+iy)\zeta(x+2iy)\right| = \sum_{n=1}^{\infty} c_n n^{-x}(3 + 4\cos(-y\log n) + \cos(-2y\log n))$$

Donde $x > 1$ e $y \in \mathbb{R}$ y $c_n > 0$ para todo $n \in \mathbb{N}$. El truco trigonométrico ideado por De la Vallée Poussin es el siguiente:

$$3 + 4\cos(\theta) + \cos(2\theta) = (1 + \cos(\theta))^2$$

Esto nos permite saber que la serie es de términos positivos, por ende cuál sea el valor a que converja debe ser mayor que:

$$\log\left|\zeta^3(x)\zeta^4(x+iy)\zeta(x+2iy)\right| > 0$$

Esto nos permite llegar a la siguiente conclusión importante:

$$\left|\zeta^3(x)\zeta^4(x+iy)\zeta(x+2iy)\right| > 1$$

La demostración de Hadamard de que la función dseta no tiene ceros en $\mathrm{Re}(s) = 1$ queda simplificada en una reducción al absurdo: Suponga que para algún $y_0 \in \mathbb{R}$, se tiene que $\zeta(1 + y_0) = 0$. Veamos que sucede con la expresión:

$$\left|\zeta^3(x)\zeta^4(x+iy_0)\zeta(x+2iy_0)\right|$$

cuando $x \to 1$ por la derecha.

Amplificando por $(x-1)^4/(x-1)^4$ y reordenando se tiene:

$$\left|(x-1)^3\zeta^3(x)\right| \cdot \left|\frac{\zeta(x+iy_0)-0}{(x-1)}\right|^4 \cdot |\zeta(x+2iy_0)| \cdot (x-1)$$

Es fácil de ver que el límite de los primeros tres elementos de la derecha tienden a algo finito: El primero es el residuo de la función dseta en $s=1$, el otro es el valor de la derivada en el hipotético cero, el otro es simplemente un límite de una función continua. Nótese que en el segundo factor no podríamos haber construido la definición de la derivada sin usar el supuesto que $\zeta(1+y_0)=0$. Pero el último término tiende a 0, por lo que todo el producto se irá a 0, pero eso contradice lo que se había obtenido con el truco trigonométrico:

$$\left|\zeta^3(x)\zeta^4(x+iy)\zeta(x+2iy)\right| > 1$$

Por lo que, por llevar a una contradicción, no pueden haber ceros de la función dseta en la recta $\mathrm{Re}(s)=1$.

El ingenio simultáneo de Hadamard y De la Vallée Poussin coronaron un siglo de avances hacia una deseada primera demostración. Para concebir el tiempo transcurrido, la Francia de Legendre era una nación agitada por la revolución y Alemania no era más que un sueño. Al finalizar Hadamard su demostración, Francia y Alemania eran imperios globales cuyas tensiones llevarían en un par de décadas a la Primera Guerra Mundial. En el caso de nuestro país, Chile pasó de ser una pobre colonia española a la época de la desigual riqueza salitrera. Eso sí la primera demostración obtenida fue inesperada: El problema desde un comienzo fue complicado al tratar del comportamiento al infinito de una función π, que además de ser discontinua en forma de escalera, daba sus saltos de forma aparentemente caprichosa. Tanto Tchebyshev como Riemann acertaron en que debía cambiarse el enunciado a uno análogo que permitiera aplicar mejores técnicas: Tchebyshev fue a por las propiedades del logaritmo, Riemann fue a por las herramientas del análisis complejo. Pero sus enfoques fueron distintos, mientras Tchebyshev usó todo su genio en intentar quebrar la nuez con las mejores herramientas que tenía, Riemann inspirado en Dirichlet dio la sorpresa al elaborar toda una teoría desde una identidad aislada ya antes descubierta por Euler. La demostración por la vía explorada por Riemann se concretó en la demostración de Hadamard y De la Vallée Poussin, lo que inició todo un nuevo campo matemático para explorar las aplicaciones del análisis real y complejo en la teoría numérica: El teorema de los números primos, cuyo enunciado nos da una información crucial sobre los bloques que forman los números enteros, también es famoso por ser la joya inspiradora para la Teoría Analítica de Números. Y esta conexión no solo brindó de resultados a la Aritmética, sino también potenció el interés en el Análisis Complejo al verse en este nuevas e interesantes aplicaciones. Cuando ramas distintas de las matemáticas se conectan, el intercambio suele ser beneficioso para ambas áreas, tanto en resultados como incentivos a investigar. Hoy el análisis complejo y la teoría numérica tienen aplicaciones prácticas en la Física de Partículas y la Encriptación, que posiblemente suenen tan insospechadas a los matemáticos del tiempo de Hadamard como sonaba de aplicar el Análisis Complejo para estudiar números primos en el tiempo de Gauss.

¿Y qué hay con el enfoque más directo de Tchebyshev? ¿Se puede demostrar el teorema de números primos sin recurrir al análisis complejo, solo recurriendo a propiedades elementales de números, series y logaritmos? Durante años las demostraciones que siguieron a la primera de Hadamard y De la Vallée Poussin también usaban del análisis complejo y la conexión con la función dseta. Se encontraron simplificaciones, formalizaciones y vías alternativas a las pruebas originales, pero nada en la vía explorada por Tchebyshev y Sylvester. Al paso de las décadas, muchos matemáticos empezaron a dudar que tal demostración elemental pudiera existir. Así diría el gran matemático británico Godfrey Hardy en una conferencia dada a la Sociedad Matemática de Copenhage en 1921:

> Ninguna prueba elemental del teorema de números primos es conocida y uno puede preguntarse si es razonable esperar una. Ahora sabemos que el teorema es más o menos equivalente a un teorema sobre una función analítica, el teorema que la función dseta de Riemann no tiene ceros en una cierta línea. Una prueba de dicho teorema, que no dependa fundamentalmente de la teoría de funciones, me parece extraordinariamente improbable. Es imprudente asegurar que un teorema matemático no puede ser probado de una manera en particular; pero una cosa me parece bastante clara. Tenemos ciertas visiones sobre la lógica de una teoría; pensamos que algunos teoremas se sitúan en lo profundo y otros están más en lo superficial. Si alguien produce una demostración elemental del teorema de los números primos, él desvelará que dichas visiones son incorrectas y el tema no está conectado de la manera que suponemos, y será el tiempo de que algunos libros sean dejados de lado y la teoría sea reescrita. [6]

Como suele suceder en la historia de las matemáticas, su desarrollo suele sorprender y contravenir lo que parece intuitivo incluso para los mejores matemáticos del momento. En 1949, el matemático noruego Atle Selberg publicó su paper *"An elementary proof on prime-number theorem"* (Una demostración elemental de teorema de los números primos) en la prestigiosa Annals of Mathematics. Aunque era la primera publicada, fue la segunda demostración elemental del Teorema de los Números Primos descubierta por Atle Selberg. La razón de publicar la segunda prueba en vez de la primera se debió a una lamentable controversia con el matemático húngaro Paul Erdős. El paper de Dorian Goldfeld ahonda lo sucedido en detalle, pero en resumen Atle Selberg comunicó resultados parciales no publicados a Paul Erdős, quien de estos derivó un lema. Al comunicarlo a Selberg, éste se dio cuenta que sumado a otros resultados preliminares no comunicados a Erdős podía demostrar el Teorema de los Números Primos de forma elemental. Selberg y Erdős se disputaron por el grado de autoría sobre el santo grial, mientras Selberg limitaba el papel de Erdős al lema que probó, Erdős reclamaba la coautoría sobre el teorema. Finalmente Selberg no demoró en encontrar otra demostración que no usaba el lema de Erdős. No lográndose un acuerdo, Selberg procedió a publicar su segunda demostración, con un esbozo de la primera en la introducción incluyendo una mención a Erdős y su resultado. Por su parte, Erdős intentó publicar la primera demostración de Selberg, pero fue rechazada por la revista.

No profundizaremos en las técnicas usadas en las dos demostraciones de Selberg, pero en forma de reseña, ambas pruebas se basan en la siguiente fórmula asintótica fundamental:

$$\vartheta(x)\log(x)+\sum_{p\leq x}\log(p)\vartheta\left(\frac{x}{p}\right)=2x\log(x)+O(x)$$

Donde $O(x)$ es la notación O-grande de Landau, es decir, existe una constante $C>0$ tal que $O(x)<Cx$ para x suficientemente grande. Esta fórmula con un diligente uso de propiedades de series permite demostrar uno de los enunciados equivalentes de Tchebyshev: $\vartheta\sim x$. A quien esté interesado en profundizar, se recomienda visitar el paper original de Selberg [13].

Con esto concluimos el apartado sobre la historia del teorema. En los capítulos siguientes demostraremos el teorema por la vía analítica. Esta demostración no pretende ser idéntica a la original, en parte nos restringiremos a los resultados netamente necesarios. Durante este reporte histórico se descuido a veces la formalidad en búsqueda de enfocarnos en la historia y sus lecciones sobre el quehacer matemático, la interconexión de las matemáticas y la importancia de las grandes preguntas. Obviamente quien desee indagar en la demostración original y sus formalidades se le invita a recurrir a las fuentes originales o a las numerosas fuentes secundarias que rescatan la fascinante historia de un teorema centenario y su tan desconcertante como emocionante demostración.

2. Extensión analítica y función Gamma

Nuestro objetivo es demostrar el **Teorema de los Números Primos**, que nos otorga un conocimiento aproximado de la distribución de los números primos a grandes escalas. Sin embargo, al lector puede extrañarle que solo en el cuarto capítulo enunciaremos formalmente el teorema y concluiremos su demostración. Esta decisión obedece a que la vía acostumbrada e histórica para demostrar el teorema proviene de conceptos ubicados en el otro confín de las matemáticas, por lo que como en las odiseas de los cuentos fantásticos deberemos iniciar una paciente andada, adquiriendo lo que vayamos descubriendo por el camino, para que cuando llegue el último desafío podamos vencer con la ayuda de todo lo aprendido en el largo viaje.

El punto de partida se encuentra en la teoría de funciones de variable compleja. Como se anunció en la introducción, se dará por conocidos los elementos básicos de esta área, aunque el lector menos acostumbrado podrá encontrar en el apéndice un apoyo valioso sobre algunos teoremas claves que recurriremos sin demostrar[1]. De todas formas, esto no quita que no podamos dedicar algunas palabras sobre el origen de nuestro camino: Esta teoría estudia principalmente el concepto de holomorfidad, que es el equivalente complejo a ser localmente[2] derivable. Algo más formal, se dice que una función f es holomorfa en un abierto $\Omega \subset \mathbb{C}$ si para cada $z \in \Omega$ existe el límite

$$\lim_{h\to 0} \frac{f(z+h)-f(z)}{h} \quad h \in \mathbb{C}$$

La única diferencia apreciable con la definición de derivada es que h puede tomar valores complejos. Esta modificación a primera vista puede parecer pequeña, pero genera las enormes repercusiones que separan al análisis complejo del análisis real. Por nombrar algunas, las funciones holomorfas son infinitamente diferenciables y para cada punto pueden ser expresadas localmente como una serie de potencias, lo que no podemos afirmar de las funciones diferenciables en los reales.

[1] Al lector que desee reforzar sus conocimientos de Análisis Complejo, se les recomienda [14].

[2] Aunque se puede definir tener derivada compleja en un punto, el término holomorfo se reserva al menos para un abierto, para que así estén presentes las propiedades que más se asocian a esta idea, por ejemplo la analiticidad local. Para la discreción del lector, esta tradición no está presente en el comienzo de [14].

Aunque no ahondaremos en todas estas propiedades, lo que de hecho requeriría todo un nuevo libro o varios más dependiendo de la profundidad de estudio deseada, sí iniciaremos con una de estas cualidades: La extensión analítica. Esta idea nos permite extender el dominio de las funciones holomorfas sin ambigüedad con prácticamente solo exigir que se mantenga la holomorfidad. Este concepto resulta fundamental para el análisis complejo y, en el rumbo hacia nuestro objetivo, fue utilizado por Benhard Riemann para definir la Función dseta, que resultará ser el puente clave entre los números primos y el análisis complejo.

2.1 Extensión analítica

Extender una función consiste sencillamente en definir una nueva función en un dominio más amplio y que para cada punto del dominio original ambas funciones coincidan. De una forma más informal, pero más intuitiva, uno agrega puntos al dominio y luego asigna los valores que la función tomará en estos nuevos puntos, sin alterar donde la función ya estaba definida. Como se pueden 'agregar' tantos puntos como se quiera al dominio y se puede 'asignar' cualquier imagen a cada punto, existen infinitas formas de extender una función. Sin embargo, raramente en matemáticas estaremos interesados en todas las posibles extensiones de una función. Más comúnmente, cuando se extiende una función, se va a preferir que conserve ciertas propiedades interesantes para su estudio. En análisis real, resulta razonable solicitar que se conserven cualidades como la continuidad o la diferenciabilidad de una función. Sin embargo, salvo algunos casos particulares, exigir que se mantengan estos atributos es insuficiente para que la extensión fuese única. Más allá de los puntos de acumulación del dominio, la continuidad y la diferenciabilidad en los números reales no hace ninguna exigencia de qué valores pueda tomar una extensión de la función.

Pero esta situación difiere bastante en Análisis Complejo. Si se expande el dominio hacia un abierto conexo, el exigir que se conserve la holomorfidad hace que haya una única extensión posible. Esto es consecuencia directa de un hecho todavía más gravitante: Si dos funciones holomorfas coinciden en una sucesión convergente de puntos distintos, entonces coinciden en cualquier abierto conexo que contenga dicha sucesión y su límite[3]. Resultará fácil al lector considerar algún ejemplo en variable real donde estas propiedades no se cumplan, basta con recordar los típicos ejercicios del Cálculo Diferencial sobre funciones diferenciables por tramos.

Esta última propiedad será demostrada en el teorema 2.1.1 y el corolario 2.1.2 que vienen a continuación. Como se puede apreciar en su demostración, este hecho proviene de que toda función holomorfa es localmente analítica, o sea, que para cada punto existe un disco donde la función puede representarse como una serie de potencias centrada en ese punto.

$$f(s) = a_0 + a_1(s-z) + a_2(s-z)^2 + a_3(s-z)^3 + \ldots \qquad \forall s \in D(z, r_z)$$

El teorema 2.1.1 demostrará para el caso particular donde una de las funciones coincidentes es 0. En el corolario 2.1.2 se prueba el caso general, para lo que no hace falta más que aplicar el teorema 2.1.1 a la diferencia entre ambas funciones.

¿Qué nos dice la demostración del teorema 2.1.1? Una idea intuitiva es ver a una función holomorfa como una especie de polinomio de grado infinito. Tenemos el siguiente patrón: Si un polinomio lineal tiene 2 ceros distintos, entonces es igual a 0. Si un polinomio cuadrático tiene 3 ceros distintos, entonces es igual a 0. Y así sucesivamente. ¿Podríamos sospechar que la función holomorfa, vista como este polinomio de grado infinito, se hace 0 porque tiene una sucesión de infinitos ceros distintos?

[3] Por supuesto, las funciones deben estar definidas y ser holomorfas en dichos abiertos.

Aunque es una idea intuitiva sobre lo que está sucediendo, esta idea tiene más de un problema. El primero es que la función holomorfa es analítica solo localmente, es decir, aunque en cada punto la función puede ser vista como una serie de potencias, esta expresión solo es válida en un disco alrededor del punto. En general, una función holomorfa sí puede tener infinitos ceros estando estos "convenientemente separados" aprovechando así lo local de localmente analítica. Ejemplos son tanto la función gamma como la función dseta que veremos en los siguientes capítulos. Para afrontar la localidad, la demostración se separa básicamente en dos partes: Primero estudiamos un disco conveniente donde la función holomorfa resultará ser 0 y segundo nos aprovecharemos de la propiedad de conexidad para extender el resultado a todo el abierto conexo.

¿Qué disco nos conviene utilizar? En base a la idea intuitiva, se requiere que la función holomorfa pueda verse como una única serie de potencias en ese disco y que el disco contenga infinitos ceros de la función. La convergencia nos viene en ayuda: Si tomamos la expansión en serie de potencias de la función holomorfa en el límite de la sucesión, no solo garantizamos poder ver la función como un polinomio de grado infinito, sino también al ser el límite de la sucesión se asegura que infinitos elementos estarán ahí y recuerde que la función se hace cero en cada uno de los elementos de la sucesión. Ya tenemos un lugar donde nuestra idea intuitiva puede trabajarse.

El siguiente problema es la formalidad. Aunque la idea intuitiva es una sospecha interesante, no está correctamente justificada cuando se pasa del grado finito al infinito. Incluso se sabe que en general no es válida: La función seno es analítica con infinitos ceros. Pero la convergencia nos viene a auxiliar nuevamente: No es solo que hay infinitos ceros en el disco, sino que hay ceros distintos al propio centro tan cercanos a éste como se quiera, ya que este centro es también el límite de la sucesión. Para entender mejor como esto afecta a la función, veamos la expansión analítica en este disco bajo una ligera factorización:

$$f(s) = \underbrace{a_m(s-z)^m}_{\text{Parte 1}} + \underbrace{a_{m+1}(s-z)^{m+1} + ...}_{\text{Parte 2}} = (s-z)^m(\underbrace{a_m}_{\text{Parte 1}} + \underbrace{a_{m+1}(s-z) + ...}_{\text{Parte 2}})$$

donde z es el límite de la sucesión. Que la función se anule en el disco es equivalente a que todos los coeficientes de la expansión sean ceros. Supongamos por momento que esto no es así, veamos a que contradicción nos conduce. Llamemos a_m al primer coeficiente no nulo de la expansión. Para cada punto de la sucesión, la parte 1 nunca se anula, por lo que debe ser compensado por la parte 2 para que la función valga cero en ese punto. Pero a medida que nos acercamos al centro, la parte 2 cae en módulo mucho más rápido que la parte 1. Esto es evidente tras la factorización, donde la parte 1 es una constante y la parte 2 es dependiente de la distancia al centro. Por lo que habrá un punto de la sucesión tan cerca del centro que la compensación ya no va a ser posible, pero como es distinto al centro exige que la compensación exista. Reducción al absurdo que culmina esta primera parte de la demostración: La función debe anularse en el disco.

La segunda parte de la prueba utiliza la conexidad en una reflexión certera sobre el resultado anterior. Se ha probado que cada límite de una sucesión de ceros tiene una vecindad donde la función se anula. Piense en el conjunto de los puntos con esta propiedad, o sea, los ceros que tienen una vecindad de ceros. Es trivial que es un abierto, basta ir a la definición de abierto en espacio métrico. Pero también es cerrado, porque el resultado anterior dice que todos los límites de sus sucesiones están en este conjunto. La conexidad tiene una relación especial con los conjuntos cerrados-abierto, porque estos y sus complementos hacen un par de abiertos disjuntos que contienen todo. Por conexidad, el abierto conexo debe estar dentro o del conjunto de ceros o de su complemento, y como sabemos que hay una sucesión que converge en ese abierto conexo ya sabemos que es parte de un conjunto de ceros, concluyendo la demostración.

Aclaradas las ideas, pasemos al teorema y su corolario con sus demostraciones formales:

Teorema 2.1.1 Sea f una función holomorfa en un abierto conexo $\Omega \subset \mathbb{C}$. Si existe una sucesión de puntos distintos convergente en Ω donde f se anule punto a punto, entonces $f = 0$ en Ω.

Demostración. Sea $\{z_n\}_{n\in\mathbb{N}}$ una sucesión cualquiera de puntos distintos que sea convergente en Ω y que $f(z_n) = 0$ para todo n natural. Llamese $z \in \Omega$ al límite de esta sucesión.

Primero se demostrará por contradicción que existe una vecindad de z donde $f = 0$. Como f es analítica, existe una vecindad V de z donde se puede representar como una serie de potencias:

$$f(s) = \sum_{k=0}^{\infty} a_k (s-z)^k$$

Como f es continua, $f(z) = \lim_{n\to\infty} f(z_n) = 0$. Esto implica que $a_0 = f(z) = 0$. Suponga que $f \neq 0$ en V, entonces existe algún coeficiente a_k distinto de 0. Denótese como a_m al primer coeficiente no nulo, entonces se puede factorizar la serie de potencias como sigue:

$$f(s) = \sum_{k=m}^{\infty} a_k (s-z)^k = a_m (s-z)^m \left(1 + g(s)\right)$$

donde g es de la forma:

$$g(s) = \sum_{k=m+1}^{\infty} \frac{a_k}{a_m} (s-z)^{k-m} = \sum_{j=1}^{\infty} b_j (s-z)^j$$

Por lo que se deduce que $g(z) = 0$. Como g es continua por ser una serie de potencias, $g(z_n) \to 0$. Luego tenemos dos propiedades que deben cumplirse para la sucesión en un $N \in \mathbb{N}$ suficientemente grande: Primero $z_N \in V$ por ser una vecindad del límite de la sucesión y segundo que $1 + g(z_N) \neq 0$ ya que $1 + g(z_n) \to 1$. Además como es una sucesión de puntos distintos, se puede escoger $z_N \neq z$. Pero esto nos lleva a una contradicción porque se tendría tres elementos no nulos cuyo producto da 0:

$$0 = f(z_N) = a_m (z_N - z)^m (1 + g(z_N))$$

Lo que concluye por reducción al absurdo la primera parte de la prueba.

Segundo, se concluye la demostración usando la propiedad de conexidad de Ω.[a] Para eso fijese en el interior del conjunto de los elementos de Ω donde f se anula, que nombraremos como A. Por construcción es abierto. Más interesante, es cerrado también, pues si se toma una sucesión convergente de puntos en A hay dos posibles situaciones para su límite. La primera posibilidad es que la sucesión sea constante de algún punto en adelante, en tal caso su límite está en A. La segunda posibilidad es que la sucesión sea de puntos distintos desde algún punto en adelante, en tal caso aplica el resultado anterior y existe una vecindad del límite donde $f = 0$. Esto quiere decir que el límite está en A para cualquiera sucesión de puntos en A, concluyendo que A es cerrado. Por lo tanto el complemento de A en Ω es cerrado y se tiene que Ω es la unión de dos subconjuntos abiertos disjuntos. Por conexidad se tiene que $\Omega = A$ o $A = \emptyset$. De la parte primera se sabe que z pertenece a A, concluyendo que $f = 0$ en Ω. ∎

[a] **Recuerdo:** La definición de un conjunto conexo es que si dicho conjunto es contenido por la unión de dos abiertos disjuntos, entonces está completamente contenido por uno de los abiertos.

Corolario 2.1.2 Sean f y g funciones holomorfas en un abierto conexo $\Omega \subset \mathbb{C}$. Si existe una sucesión convergente de puntos distintos en Ω donde $f = g$ punto a punto, entonces $f = g$ en Ω.

Demostración. Aplicar el teorema anterior a la función holomorfa $f - g$.

Como se ha dicho, la aplicación principal de este teorema y su corolario es la extensión analítica. Los requerimientos son pocos y generales: Necesitamos tener una función definida en un dominio, subconjunto de $\mathbb{C}$, que contenga alguna sucesión de puntos distintos que converja en el dominio. Aquello es suficiente para saber que existe una única extensión holomorfa para cada nuevo dominio abierto conexo que contenga al original. ¿Cómo se sabe la unicidad? Pues si hubiera dos extensiones F y G de la función f definidas en el mismo nuevo dominio, se cumple para todo punto del dominio original que:

$$F(s) = f(s) = G(s)$$

Como el dominio original contiene al menos una sucesión convergente de puntos distintos, basta utilizar esta para aplicar el corolario 2.1.2 y obtener que $F = G$.

Como se verá a lo largo de esta obra, la utilidad de esta idea poderosa trasciende los intereses puros del análisis complejo: Al extender funciones reales o de otros subconjuntos numéricos, podemos usar las potentes herramientas del análisis complejo para estudiar estas extensiones y así obtener resultados de interés sobre las funciones originales. Por el momento, avancemos con un ejemplo muy importante de función extendida analíticamente: **La Función Gamma**.

2.2 Función Gamma

La Función Gamma es una de las funciones más importantes de las matemáticas. Introducida por el gran matemático francés Adrien-Marie Legendre, es considerada la extensión natural de los factoriales a los números reales positivos y se define como sigue:

Definición 2.2.1 — Función Gamma. La función gamma en los reales $\Gamma_{\mathbb{R}} : \mathbb{R}^+ \to \mathbb{R}$ está dada por:

$$\Gamma_{\mathbb{R}}(s) = \int_0^\infty e^{-t} t^{s-1} dt$$

El objetivo para la siguiente sección se tratará de extender analíticamente esta función lo mayor posible en los complejos. Más en específico, el problema a resolver será el siguiente:

Problema 2.1 Encontrar una extensión analítica para la función $\Gamma_{\mathbb{R}}$ cuyo dominio sea todo $\mathbb{C}$ salvo los enteros no positivos $0, -1, -2, \ldots$

La razón por la cual excluimos los enteros no negativos es que, como se verá más adelante, estos puntos constituirán polos de la extensión analítica que buscamos, es decir, puntos donde la función extendida se irá irremediablemente a infinito. Note que este problema tiene sentido porque la función gamma está definida en un conjunto que al menos contiene una sucesión convergente de puntos distintos y al conjunto que queremos extenderla es un abierto conexo, por lo que habrá una única respuesta. Ningún requerimiento de que la función gamma sea previamente analítica es necesario (de hecho, así definida no lo es ya que su dominio no contiene abiertos en $\mathbb{C}$). Aunque las exigencias para extender analíticamente una función son exiguas, lamentablemente existen las excepciones: No podemos plantearnos el problema de extender analíticamente el factorial de forma única porque su dominio $\mathbb{N}$ no contiene una sucesión convergente de puntos distintos, de ahí la necesidad de comenzar con la función gamma y no con el factorial.

Un buen punto de partida para abordar el problema es ir a la extensión más natural que se nos puede ocurrir: Sustituir el s real por un s complejo en la integral que define la función gamma. Para que esta extensión sirva a nuestros propósitos debe cumplir con dos propiedades: Primero, existir. Sin que la integral impropia converja, no tiene sentido siquiera plantearla como una extensión. Lamentablemente la integral no converge para cualquier $s \in \mathbb{C}$, lo que era esperable pues tampoco lo hace en todo $\mathbb{R}$. Pero sí converge para la mitad del plano complejo donde las partes reales son positivas, que llamaremos el semiplano H. La prueba es relativamente sencilla y se deriva de la siguiente propiedad elemental que, de paso, nos ahorrará muchos cálculos futuros:

Proposición 2.2.1 Sea $a \in \mathbb{R}^+$ y $b \in \mathbb{C}$, entonces se tiene que:

$$\left|a^b\right| = a^{\mathrm{Re}(b)}$$

Demostración. Sea $x = \mathrm{Re}(b)$ e $y = \mathrm{Im}(b)$, entonces:

$$\left|a^b\right| = \left|a^{x+iy}\right| = \left|e^{(x+iy)\log(a)}\right| = e^{x\log(a)}\left|e^{iy\log(a)}\right| = e^{x\log(a)} = a^{\mathrm{Re}(b)}$$

Usando la proposición 2.2.1, es sencillo ver que la integral impropia converge absolutamente, ya que la integral de términos absolutos está acotada por la función gamma con valores en los reales. Desafortunadamente, también nos señala que esta extensión no es la que buscamos al estar en un dominio más pequeño, sin embargo sirve como un buen punto de partida.

El segundo requisito para que esta sea una extensión analítica es que la función extendida sea holomorfa con dominio abierto conexo. Estamos en suerte que H es un abierto conexo, por lo que nos gustaría demostrar que la integral impropia converge en una función holomorfa en el semiplano H. Para esta prueba, se puede comenzar desde el teorema 5.1.1 del apéndice que nos proporciona información sobre la holomorfidad de funciones definidas a través de una integral. Resulta evidente corroborar que nuestra integral cumple con los requerimientos del teorema, exceptuando por uno: El intervalo de integración en el teorema es $[0,1]$, mientras en nuestro caso es $(0,\infty)$. Un truco interesante es que podemos cambiar el intervalo de integración con solo componer el argumento de la integral por una transformación lineal, de tal forma que si hacemos un cambio de variable nos queda el mismo argumento con diferente intervalo de integración. Lamentablemente esta idea nos deja con dos problemas: Primero, esta manera no nos permite que uno de los límites del intervalo se vaya hasta infinito. Segundo, el teorema exige que el argumento sea continuo en 0, pero $e^{-t}t^{s-1}$ ni siquiera está definido en $t = 0$, lo que no es arreglable con un cambio de variable conveniente.

De todas formas, nada está perdido: El teorema 5.1.1 nos dice que las funciones definidas por $\int_c^d e^{-t}t^{s-1}dt$ son holomorfas para cualquier intervalo con $0 < c < d < \infty$. La idea natural es intentar aproximar a la integral impropia con una sucesión de estas funciones, donde el intervalo de integración se vaya incrementando más y más. Pero ¿el límite de una sucesión de funciones holomorfas es una función holomorfa? El teorema 5.1.2 nos dice que sí siempre que la sucesión converja uniforme. Para probar la convergencia uniforme se procede como en los ejercicios típicos de Análisis, utilizando desigualdad triangular y otros acotamientos. Pero será necesario una última idea para que estas cotas funcionen: En vez de probar que la función gamma es holomorfa para cualquier s en H, se probará que es holomorfa para todo s con parte real entre $a, b > 0$ fijos cualquiera. Este cambio no altera la demostración, ya que todo punto en H cumple con estar en alguna de estas franjas, solo basta con tomar a y b convenientes. Pero este cambio nos brindará cotas uniformes (es decir, que no dependen de s) en cada franja, lo que nos permite probar la convergencia uniforme en estas franjas y con ello deducir la holomorfidad en todo H. Sin mayor dilación, pasemos a la demostración formal:

Proposición 2.2.2 — Una primera extensión analítica para la función gamma. Sea $H := \{s \in \mathbb{C} \mid \mathrm{Re}(s) > 0\}$. La función gamma en el semiplano $\Gamma_H : H \to \mathbb{C}$ dada por la expresión:

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

es una extensión analítica de $\Gamma_\mathbb{R}$.

Demostración. Primero, usando la proposición 2.2.1, de un breve cálculo se tiene:

$$|\Gamma_H(s)| = \left| \int_0^\infty e^{-t} t^{s-1} dt \right| \leq \int_0^\infty |e^{-t} t^{s-1}| dt = \int_0^\infty e^{-t} t^{\sigma-1} dt = \Gamma_\mathbb{R}(\sigma)$$

Para cualquier $s \in H$, denotando $\sigma > 0$ a su parte real. Luego la integral impropia es absolutamente convergente, por ende Γ_H está bien definida en su dominio.

Por construcción, es evidente que $\Gamma_H(s) = \Gamma_\mathbb{R}(s)$ para todo $s > 0$ real, luego es una extensión. Para concluir que es la extensión analítica en H basta probar que Γ_H es holomorfa en H. Sea $z \in H$ cualquiera con parte real $x > 0$. Considere $a = x/2$ y $b = 2x$ reales positivos, se define la franja abierta $H_{a,b} := \{s \in \mathbb{C} \mid a < \mathrm{Re}(s) < b\}$. Evidentemente, $z \in H_{a,b}$ y $H_{a,b}$ es abierto. Probaremos que Γ_H es holomorfa en $H_{a,b}$.

Para esta parte de la demostración se usará las funciones auxiliares $F_\varepsilon : H_{a,b} \to \mathbb{C}$ definidas para cada $\varepsilon \in (0,1)$ por la integral:

$$F_\varepsilon(s) = \int_\varepsilon^{\frac{1}{\varepsilon}} e^{-t} t^{s-1} dt$$

Estas funciones son holomorfas. Para demostrarlo, considere para cada ε la función $G : \mathbb{C} \times [0,1] \to \mathbb{C}$ dada por

$$G(s,u) = \left(\frac{1}{\varepsilon} - \varepsilon\right)^{-1} \left(\varepsilon + u\left(\frac{1}{\varepsilon} - \varepsilon\right)\right)^{s-1} \exp\left\{-\varepsilon - u\left(\frac{1}{\varepsilon} - \varepsilon\right)\right\}$$

Resulta claro que esta función es continua y además es entera para cada $u \in [0,1]$ fijo. Aplicando el cambio de variable $t = \varepsilon + u(1/\varepsilon - \varepsilon)$, se logra la siguiente igualdad:

$$\int_0^1 G(s,u) du = \int_\varepsilon^{\frac{1}{\varepsilon}} e^{-t} t^{s-1} dt = F_\varepsilon(s)$$

Por el teorema 5.1.1 del apéndice, $s \to \int_0^1 G(s,u) du$ es holomorfa y por ende $F_\varepsilon(s)$ también.

A continuación, se probará que F_ε converge uniformemente a Γ_H en la franja $H_{a,b}$ cuando $\varepsilon \to 0$. De desigualdad triangular y proposición 2.2.1 se tiene:

$$\begin{aligned}
|\Gamma_H(s) - F_\varepsilon(s)| &\leq \left| \int_0^\varepsilon e^{-t} t^{s-1} dt \right| + \left| \int_\varepsilon^\infty e^{-t} t^{s-1} dt \right| && (s \in H_{a,b}) \\
&\leq \int_0^\varepsilon \left| e^{-t} t^{s-1} \right| dt + \int_\varepsilon^\infty \left| e^{-t} t^{s-1} \right| dt \\
&= \int_0^\varepsilon e^{-t} t^{\sigma-1} dt + \int_\varepsilon^\infty e^{-t} t^{\sigma-1} dt && (\sigma = \mathrm{Re}(s))
\end{aligned}$$

La primera integral se puede acotar de la siguiente forma:

$$\begin{aligned}\int_0^{\varepsilon} e^{-t}t^{\sigma-1}dt &\leq \int_0^{\varepsilon} t^{\sigma-1} && \left(t \geq 0 \Rightarrow e^{-t} \leq 1\right)\\ &= \frac{t^{\sigma}}{\sigma}\Big|_0^{\varepsilon} = \frac{\varepsilon^{\sigma}}{\sigma} \leq \frac{\varepsilon^{\sigma}}{a} && (a \leq \sigma)\\ &\leq \frac{\varepsilon^{a}}{a} \to 0 && (\varepsilon < 1)\end{aligned}$$

Mientras, para valores de ε suficientemente pequeños, la segunda integral se puede acotar así:

$$\begin{aligned}\int_{\varepsilon}^{\infty} e^{-t}t^{\sigma-1}dt &\leq \int_{\varepsilon}^{\infty} e^{-t}t^{b-1}dt && \left(\sigma \leq b \wedge t \geq \frac{1}{\varepsilon} \gg 1\right)\\ &\leq K\int_{\varepsilon}^{\infty} e^{-t}e^{t/2}dt && \left(\frac{t^{b-1}}{e^{t/2}} \to 0\right)\\ &\leq K\int_{\varepsilon}^{\infty} e^{-t/2}dt = 2e^{-\varepsilon/2} \to 0\end{aligned}$$

Luego tenemos que F_ε converge uniforme a Γ_H en $H_{a,b}$. Como F_ε es holomorfa, Γ_H es holomorfa en $H_{a,b}$ por teorema 5.1.2 del apéndice. Por lo tanto, Γ_H es holomorfa en z y como z era un elemento cualquiera de H, Γ_H es holomorfa en z.

Desde aquí, el camino hacia la extensión definitiva es expedito. Primero, necesitamos ver que la función gamma extendida al semiplano H cumple con la propiedad que hace la función gamma interesante: Generalizar el factorial al hacer $\Gamma_H(s+1) = s\Gamma_H(s)$. La prueba es simplemente una integración por partes. Adicionalmente se puede demostrar que generaliza a factorial por inducción:

Proposición 2.2.3 Para todo $s \in H$, se cumple que:

$$\Gamma_H(s+1) = s\Gamma_H(s)$$

En particular se deduce que $\Gamma(n+1) = n!$ para todo $n \in \mathbb{N}$.

Demostración. Aplicando integración por partes se obtiene:

$$\Gamma_H(s+1) = \int_0^{\infty} e^{-t}t^{s}dt = -e^{-t}t^{s}\big|_0^{\infty} + \int_0^{\infty} e^{-t}st^{s-1}dt = s\Gamma_H(s)$$

El caso particular puede probarse por inducción. **Caso inicial:** Para $n = 0$.

$$\Gamma_H(1) = \int_0^{\infty} e^{-t}dt = -e^{-t}\big|_0^{\infty} = 1 = 0!$$

Caso inductivo: Si se cumple para n, entonces para $n+1$.

$$\Gamma_H(n+2) = (n+1)\Gamma(n+1) = (n+1)n! = n+1!$$

Esta propiedad no es interesante solamente porque generaliza al factorial, sino porque nos da la técnica para extender sin dificultad Γ_H hacia todo $\mathbb{C}$, salvo por los enteros no positivos. Si reordenamos términos, se tiene que $\Gamma_H(s) = \Gamma_H(s+1)/s$. Pero el lado de la derecha se puede calcular más allá de H, en específico se puede calcular para todos los complejos con parte real mayor a -1, salvo 0. Por álgebra de funciones holomorfas, $\Gamma_H(s+1)/s$ es holomorfa en su dominio. La gracia es que esta función, llamemos Γ_1, también cumple con la propiedad:

$$\Gamma_1(s) = \Gamma_H(s+1)/s = \Gamma_1(s+1)/s$$

Ya que $s+1 \in H$ y Γ_1 la construimos como extensión. Luego el lado derecho podemos usarlo para hacer una nueva extensión para todo s con parte real ahora mayor a -2, salvo el 0 y el -1 (ya que Γ_1 no está definida en 0). Todo sucede igual, por lo que podemos ir extendiendo una y otra vez (es decir, una inducción) hasta definir una extensión analítica en casi todo $\mathbb{C}$, salvo por los enteros no positivos, logrando así resolver el problema objetivo de esta sección:

Teorema 2.2.4 — Extensión analítica de la función gamma. Existe una función meromorfa Γ con dominio en todo $\mathbb{C}$, salvo en sus polos $\{0,-1,-2,\ldots\}$, que extiende a Γ_H.

Demostración. La extensión se construye a través de un proceso inductivo.

Caso inicial: Sea $A_1 = \{s \in \mathbb{C} \mid \mathrm{Re}(s) > -1\}$. Se define $\Gamma_1 : A_1 - \{0\} \to \mathbb{C}$ dada por:

$$\Gamma_1(s) = \frac{\Gamma_H(s+1)}{s}$$

Dado que Γ_H es holomorfa, Γ_1 es meromorfa en A_1 salvo en el polo $s=0$ por álgebra de funciones holomorfas. Adicionalmente, por la proposición 2.2.3 se tiene para $s \in H$:

$$\Gamma_1(s) = \frac{\Gamma_H(s+1)}{s} = \frac{s\Gamma_H(s)}{s} = \Gamma_H(s)$$

Luego tenemos que Γ_1 es una extensión meromorfa de Γ_H en A_1[a]. Además Γ_1 cumple con una propiedad símil a la de la proposición 2.2.3, ya que para todo $s \in A_1$, $s+1 \in H$, lo que implica que $\Gamma_H(s+1) = \Gamma_1(s+1)$, que aplicado a la definición de Γ_1 se obtiene la propiedad deseada.

Caso inductivo: Sea $n \in \mathbb{N}$. Definamos A_n y A_{n+1} como el conjunto de los números complejos con parte real estrictamente mayor que $-n$ y $-n-1$ respectivamente. Suponga que existe una función meromorfa Γ_n que extiende a Γ_H en A_n salvo los polos en $\{0,-1,\ldots,1-n\}$ y que cumple con la propiedad 2.2.3, pero en todo su dominio. Entonces sea $\Gamma_{n+1} : A_{n+1} - \{0,-1,\ldots,-n\} \to \mathbb{C}$ definida por

$$\Gamma_{n+1}(s) = \frac{\Gamma_n(s+1)}{s}$$

Como Γ_n es holomorfa salvo en los polos $\{0,-1,\ldots,1-n\}$, entonces Γ_{n+1} es meromorfa con polos en $\{0,-1,\ldots,-n\}$. Además para todo $s \in A_n$ se tiene que:

$$\Gamma_{n+1}(s) = \frac{\Gamma_n(s+1)}{s} = \frac{s\Gamma_n(s)}{s} = \Gamma_n(s)$$

Luego Γ_{n+1} es una extensión meromorfa de Γ_n y por ende de Γ_H. Adicionalmente bajo un argumento anólogo al caso inicial, Γ_{n+1} también obedece la propiedad 2.2.3, en todo su dominio.

Luego se puede definir $\Gamma : \mathbb{C} - \{0,-1,\ldots\} \to \mathbb{C}$ como la función que, para cada s en el dominio, toma el valor $\Gamma(s) = \Gamma_k(s)$ para algún Γ_k que esté definido en s. Como la sucesión de los $A_k - \{0,-1,\ldots\}$ cubren el dominio, entonces Γ tiene valor para cada s en el dominio. Como cada Γ_k es una extensión de las anteriores, sus valores siempre coinciden y, por ende, el valor de $\Gamma(s)$ es único. Más encima como es idéntica a cada Γ_k en sus respectivos dominios, es holomorfa como lo son estas en sus dominios, concluyendo Γ es la extensión meromorfa Γ_H con dominio en $\mathbb{C} - \{0,-1,\ldots\}$.

[a] **Recuerdo:** Una función meromorfa es una función holomorfa salvo en ciertos puntos que son sus polos.

2.3 Polos y raíces de la Función Gamma

Además de ser un útil ejemplo de la construcción de extensiones analíticas, la función gamma será crucial para poder crear la extensión analítica que define a la función dseta, el mayor protagonista de la demostración del teorema de los números primos. Por lo que antes de proseguir hacia el estudio de la función dseta, es necesario demostrar algunas propiedades de la función gamma que serán útiles en el estudio futuro. Se limitará solo a las propiedades que serán usadas en las siguientes demostraciones, evitan introducirnos en el fascinante mundo de las importantes propiedades de la función gamma. Al lector interesante, el libro base [14] puede ser un recomendable punto de partida.

La primera propiedad que demostraremos es símil a la proposición 2.2.3, siendo esta la propiedad más conocida de la función gamma. La demostración es directa desde que las extensiones intermedias de la proposición anterior cumplían esta propiedad en sus dominios:

Teorema 2.3.1 Para todo $s \in \mathbb{C} - \{1, 0, -1...\}$ se cumple que:

$$\Gamma(s+1) = s\Gamma(s)$$

Demostración. Sea $s \in \mathbb{C} - \{1, 0, -1...\}$ cualesquiera. Por principio arquimideano, existe $n \in \mathbb{C}$ tal que $-n < \mathrm{Re}(s)$. Por ende, $-n < \mathrm{Re}(s+1)$ también. Luego:

$$\Gamma(s+1) = \Gamma_n(s+1) = s\Gamma_n(s) = s\Gamma_n(s)$$

Siendo Γ_n las extensiones intermedias usadas en la demostración 2.2.4 y que cumplían con la segunda igualdad.

La segunda propiedad es el cálculo de residuos de los polos de la función gamma, que nos será de utilidad más adelante. La prueba también es directa usando la propiedad anterior iteradamente en el límite del cálculo del residuo para polo simple.

Teorema 2.3.2 La función Γ tiene polos simples en $\{0, -1, \ldots\}$, siguiendo sus residuos la siguiente fórmula:

$$\mathrm{Res}_\Gamma(-n) = \frac{(-1)^n}{n!} \quad \forall n \in \{0, 1, ...\}$$

Demostración. Basta con calcular el residuo para polo simple usando $n+1$-veces la proposición 2.3.1 para $n = 1, 2, \ldots$:

$$\begin{aligned} Res(\Gamma, -n) &= \lim_{s \to -n} (s+n)\Gamma(s) \\ &= \lim_{s \to -n} \frac{(s+n)\Gamma(s+1)}{s} \\ &= \lim_{s \to -n} \frac{(s+n)\Gamma(s+n+1)}{s(s+1)\ldots(s+n)} \\ &= \frac{\Gamma(1)}{(-n)(-n+1)\ldots(-1)} = \frac{(-1)^n}{n!} \end{aligned}$$

Para $n = 0$, haciendo un cálculo análogo, para la segunda línea del cálculo anterior se desprende que el límite dará $\Gamma(0+1) = 1$, así cumpliendo también con la fórmula. Como los límites alcanzan valores definidos, se corrobora que son polos simples.

La siguiente propiedad de la función gamma es la fórmula de reflexión de Euler. Esta nos presenta una especie de simetría imperfecta de la función gamma entorno a la recta $\mathrm{Re}(s) = 1/2$, en el sentido que, conocido un valor de la función gamma para algún z, podemos obtener el valor para $1-z$ con una expresión sencilla respecto a $\Gamma(z)$. Siendo más explícitos, la fórmula es $\Gamma(s)\Gamma(1-s) = \pi/\sin(\pi s)$. Esta expresión tan sencilla es extremadamente poderosa, por ejemplo, nos permite una forma alternativa de extender analíticamente la función gamma. Evidentemente, por el corolario 2.1.2 llevan al mismo resultado, y, aunque es más corta, esta vía resulta menos intuitiva ya que requiere saber a priori hacia dónde te diriges.

La demostración de la fórmula de reflexión es en dos sencillas partes. Primero se prueba para $0 < s < 1$ real. Partiendo del producto $\Gamma(s)\Gamma(1-s)$, se las expresa como integrales y haciendo un ingenioso cambio de variable se llega a una integral conocida cuyo resultado es justamente $\pi/\sin(\pi s)$ (véase la proposición 5.1.3 del apéndice). Probada la igualdad en $(0,1)$, hemos obtenido que dos funciones holomorfas en $\mathbb{C}-\mathbb{Z}$ que coinciden en un conjunto que contiene sucesiones convergentes de puntos distintos, luego por el corolario 2.1.2 deben ser iguales en todo $\mathbb{C}-\mathbb{Z}$, extendiendo así el resultado para todo el conjunto deseado en el enunciado. Procedamos a la demostración formal:

Teorema 2.3.3 — Fórmula de reflexión de Euler. Para todo $s \in \mathbb{C}-\mathbb{Z}$, se cumple:

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

Demostración. Para $s \in (0,1)$ se puede realizar el siguiente cálculo:

$$\begin{aligned}
\Gamma(s)\Gamma(1-s) &= \int_0^\infty e^{-t}t^{s-1}\Gamma(1-s)dt \\
&= \int_0^\infty e^{-t}t^{s-1}\left(\int_0^\infty e^{-u}u^{-s}du\right)dt \\
&= \int_0^\infty e^{-t}t^{s-1}\left(\int_0^\infty e^{-vt}(vt)^{-s}tdv\right)dt \\
&= \int_0^\infty \int_0^\infty e^{-(1+v)t}v^{-s}dvdt \\
&= \int_0^\infty \int_0^\infty e^{-(1+v)t}v^{-s}dtdv \\
&= \int_0^\infty \frac{v^{-s}}{v+1}dv
\end{aligned}$$

Donde se aplicó el cambio de variable $u = vt$ con v como nueva variable. De la proposición 5.1.3, sabemos que para $s \in (0,1)$, esta integral es igual a $\frac{\pi}{\sin \pi s}$.

Como $\Gamma(s)\Gamma(1-s)$ y $\frac{\pi}{\sin \pi s}$ son dos funciones holomorfas en $\mathbb{C}-\mathbb{Z}$ que coinciden en $(0,1)$ que contiene sucesiones convergentes de puntos distintos. Luego por el corolario 2.1.2, la igualdad se cumple para todo $s \in \mathbb{C}-\mathbb{Z}$.

Para concluir este apartado, otros resultados interesantes que son fáciles de obtener gracias a la fórmula de reflexión son referentes a los ceros de la función gamma, o más bien a su ausencia. En efecto, si observamos con atención la fórmula, el lado derecho $\pi/\sin(\pi s)$ jamás se hace cero. Siendo ese el producto de dos imágenes de la función gamma, se hace imposible que la función gamma se haga cero (o el producto se haría cero en alguna parte). La única forma sería evitar la fórmula de reflexión usando un s tal que $1-s$ sea un polo. Pero esto tampoco funciona, porque si $1-s$ es un entero no positivo, s es un natural y ya sabemos que en esos casos $\Gamma(n) = \Gamma_H(n) = n-1!$. Esto se encuentra en el corolario 2.3.4.

El corolario con que cierra este capítulo es una consecuencia natural del anterior: A falta de ceros de la función gamma, la función $1/\Gamma$ se puede extender a una función entera, es decir, holomorfa en todo $\mathbb{C}$. Esto se puede entender, al no haber ceros de Γ, $1/\Gamma$ solo se indefine en los polos simples de Γ, pero esas indefiniciones son reparables de forma continua y terminan siendo ceros "simples". Una forma más explícita de probarlo es usar la fórmula de reflexión para construir una extensión analítica y reparar las indefiniciones calculando los límites en esos puntos. Dado este corolario es que en próximos capítulos se usará $1/\Gamma$ para referirse a su extensión entera.

Corolario 2.3.4 La función Γ no tiene raíces.

Demostración. Separemos en dos casos: Si $s \in \mathbb{Z}$, o es un polo de Γ, o es un entero positivo por lo que $\Gamma(s) = s-1! \neq 0$. Si $s \notin \mathbb{Z}$, cumple la fórmula de reflexión de Euler. Como $\frac{\pi}{\sin \pi s}$ no puede ser cero, $\Gamma(s)\Gamma(1-s)$ tampoco lo es y, por ende, $\Gamma(s) \neq 0$.

Corolario 2.3.5 La función $1/\Gamma$ se puede extender a una función entera con ceros en $\{0,-1,\ldots\}$.

Demostración. De la formula de reflexión se tiene que:

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

La expresión del lado derecho permite una extensión continua en todo $\mathbb{C}$, ya que en los únicos puntos donde no está definida que son los enteros positivos, el límite existe:

$$\begin{aligned}
\lim_{s\to n} \frac{1}{\pi}\Gamma(1-s)\sin(\pi s) &= \lim_{s\to n} \frac{1}{\pi}(n-s)\Gamma(1-s)\frac{\sin(\pi s)}{n-s} \\
&= \frac{1}{\pi}\lim_{s\to n}(n-s)\Gamma(1-s)\lim_{s\to n}\frac{\sin(\pi s)}{n-s} \\
&= \frac{1}{\pi}\lim_{s\to n}(n-s)\Gamma(1-s)\lim_{s\to n}\frac{\sin(\pi s)}{n-s} \\
&= \frac{1}{\pi}\lim_{u\to 1-n}(u-1+n)\Gamma(u)\lim_{s\to n}\frac{\pi\cos(\pi s)}{(-1)} \\
&= -\mathrm{Res}(\Gamma, 1-n)\cos(\pi n) = \frac{-1}{n!}
\end{aligned}$$

La extensión continua del lado derecho es holomorfa, ya que fuera de los enteros positivos es el producto de dos holomorfas y en los enteros positivos los polos fueron cancelados por ceros simples. Por lo tanto esto permite una extensión entera para $1/\Gamma$. Además, del corolario anterior, sus ceros son los de la función sinusoidal ya que Γ no tiene ceros, luego sus ceros son $\{0,1,\ldots\}$.

3. Función dseta de Riemann

El presente capítulo se dedicará en indagar en las propiedades de la protagonista de la demostración del teorema de los números primos y una de las funciones más importantes de las matemáticas: La Función Dseta de Riemann. Sus orígenes están en una función, definida a través de series, que fue considerada por Leonhard Euler para su demostración de infinitos números primos:

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

Pero sería el gran matemático Benhard Riemann, en su famoso, que consideraría extender analíticamente esta función hacia casi todos los complejos, con la excepción de 1. Esta idea fue un salto de genio que permitió utilizar las herramientas del análisis complejo para estudiar los números primos que están fuertemente vinculados con la función dseta. Esto es la clave para la demostración del teorema de los números primos y ha sido un núcleo fundamental para la teoría analítica desde entonces, al punto que hoy el problema abierto más famoso de dicha teoría sea probablemente la hipótesis de Riemann sobre algunos ceros de la función dseta.

Las secciones de este capítulo se dedicarán a 4 cuestiones principales: Primero a construir la función dseta como una extensión analítica, luego indagaremos en la identidad de Euler que relaciona esta función con los números primos, siguiente probaremos algunos teoremas que permiten estimar el crecimiento de la función dseta en ciertos contextos y finalmente en la última sección discutiremos sobre los ceros de esta función. Todos estos conocimientos serán cruciales para la demostración del teorema de los números primos de la cual versará el capítulo siguiente.

Antes de comenzar el estudio lleno de estas materias, una aclaración de interés sobre el nombre de la función. Este nombre proviene de la letra griega ζ, que en español se llama Dseta. En cambio, en el inglés a esa letra griega se le llama Zeta, de ahí viene que la mayoría de la literatura matemática se refiera a esta función como la Función Zeta de Riemann. La convención está muy extendida para ser cambiada en mi humilde opinión y tampoco es que sea un error porque el lenguaje vivo se adapta a los usos y costumbres de la gente. Pero en este texto se usará el nombre de Dseta, como una forma de remarcar un hecho curioso y a veces olvidado de esta función. El lenguaje es libre.

3.1 Extensión analítica a la función dseta

Similar a la sección 2.2 sobre la función gamma, el objetivo de este capítulo será crear la extensión analítica con mayor dominio posible en $\mathbb{C}$ para la siguiente función:

Definición 3.1.1 La función dseta en los reales $\zeta_{\mathbb{R}} : D \to \mathbb{R}$ está dada por:

$$\zeta_{\mathbb{R}}(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

donde $s \in D := \{x \in \mathbb{R} \mid x > 1\}$.

Una aclaración es necesaria: Aunque llamaremos función dseta en los reales a esta función en falta de un mejor nombre, en general el nombre de función dseta se reserva exclusivamente para la extensión que construiremos en esta sección. Dicho aquello, formulemos más formalmente nuestro problema objetivo:

Problema 3.1 Encontrar una extensión analítica para la función $\zeta_{\mathbb{R}}$ cuyo dominio sea $\mathbb{C} - \{1\}$.

Es fácil de ver que este problema está bien planteado, es decir, el problema ofrece una solución única gracias al corolario 2.1.2, que podemos aplicar ya que el dominio original $\mathbb{R}^+$ contiene sucesiones convergentes de puntos distintos y el dominio final $\mathbb{C} - \{1\}$ es un abierto conexo.

El camino hacia esta extensión analítica será en dos pasos de forma semejante al caso de la función gamma: Primero comenzaremos con la extensión más natural posible, que resulta de permitir que la variable s en la expresión que define a $\zeta_{\mathbb{R}}$ tome valores complejos. Al igual que en el caso de la función gamma, para que este cambio construya una extensión analítica, la función resultante debe cumplir con dos criterios: Primero, existir para algún abierto conexo que contenga a $\mathbb{R}^+$, esto es que la serie converja para cada s en ese conjunto. Segundo, debe ser holomorfa en dicho abierto. Es claro que ese abierto conexo no puede ser $\mathbb{C}$, ni $\mathbb{C} - \{1\}$, basta con tomar $s = 0$ y tenemos una suma de infinitos 1s que claramente diverge[1].

La demostración de ambos requisitos viene en un solo paso dado por una útil consecuencia de la proposición 2.2.1 en abiertos conexos de la forma $\mathrm{Re}(s) > a$:

$$\left|n^{-s}\right| \leq n^{-\mathrm{Re}(s)} \leq n^{-a} \Rightarrow \sum_{n=1}^{\infty} \left|n^{-s}\right| \leq \sum_{n=1}^{\infty} n^{-a} = \zeta_{\mathbb{R}}(a) < \infty$$

Vea que el lado derecho nos asegura una convergencia absoluta (¡Se tiene la existencia!), pero además se cumplen las condiciones para usar el Criterio de Weierstrass (Teorema 5.1.4 del apéndice): Los elementos de la serie están acotados por una sucesión sumable. Este criterio dice que las sumas parciales convergen uniforme y como éstas son holomorfas, la serie lo será. ¿Cuánto podemos extender este razonamiento? Entre menor el a más elementos se abarcará, pero $\zeta_{\mathbb{R}}(a)$ solo converge si $a > 1$. De forma ingeniosa, se puede extender hasta el abierto $H_1 : \mathrm{Re}(s) > 1$. Más formalmente:

Proposición 3.1.1 — Una primera extensión analítica de ζ. Sea $H_1 := \{s \in \mathbb{C} \mid \mathrm{Re}(s) > 1\}$. La función dseta en el hiperplano $\zeta_{H_1} : H_1 \to \mathbb{C}$ definida por:

$$\zeta_{H_1}(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

es una extensión analítica de $\zeta_{\mathbb{R}}$.

[1] El lector curioso notará que $s = 1$ también diverge porque es la serie armónica. Más interesante, $s = 2$ converge y es la famosa serie del Problema de Basilea, que dicho sea de paso, también fue resuelto por Leonhard Euler

Demostración. Sea $z \in H_1$ cualquiera, se demostrará que ζ_{H_1} converge y es holomorfa en z. Sea $a = (\mathrm{Re}(z)+1)/2$ fijo, evidentemente se tiene que $z > a > 1$. Para todo $s \in H_1$ tal que $\sigma = \mathrm{Re}(s) > a$ se cumple por la proposición 2.2.1 que:

$$\left|\frac{1}{n^s}\right| \leq \left|\frac{1}{n^\sigma}\right| \leq \left|\frac{1}{n^a}\right|$$

Adicionalmente como $a > 1$:

$$\sum_{n=1}^{\infty} \frac{1}{n^a} = \zeta_{\mathbb{R}}(a) < \infty$$

Al cumplirse las hipótesis, la sucesión de sumas parciales $\sum_{n=1}^{N} n^{-s}$ converge uniforme para la sección del plano complejo $\mathrm{Re}(s) > a$ por el criterio de weierstrass (Teorema 5.1.4). Luego ζ_{H_1} existe en dicha región. Además como las sumas parciales son holomorfas al ser una adición de funciones exponenciales, ζ_{H_1} es holomorfa por teorema 5.1.2 en dicha región. Ya que z pertenece a la región, ζ_{H_1} existe y es holomorfa en z. Y como z era un elemento cualquiera en H_1, ζ_{H_1} está bien definida y es holomorfa.

Finalmente es evidente que $\zeta_{H_1}(s) = \zeta_{\mathbb{R}}(s)$ para $s \in \mathbb{R}$ porque tienen la misma regla que las define, luego ζ_{H_1} es una extensión analítica de $\zeta_{\mathbb{R}}$.

El siguiente paso para extender hacia la función dseta en $\mathbb{C} - \{1\}$ requiere de un ingenio sutil y una formalidad meticulosa. Fue concebido por Benhard Riemann en su magna disertación "Sobre el número de primos menores a una cantidad dada" de 1859. La idea es extender analíticamente la serie de forma implícita, es decir, extender analítica otra función que contenga a la función dseta en su expresión y de ahí derivar la extensión para la función dseta que deseamos. Más en específico, la función con que trabajaremos es:

Definición 3.1.2 La función $\xi_{H_1} : H_1 \to \mathbb{C}$ se define por la siguiente expresión:

$$\xi_{H_1}(s) = \pi^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) \zeta_{H_1}(s)$$

Resulta claro que, como $1/\Gamma$ y $\pi^{s/2}$ son enteras, si extendemos analíticamente la función ξ_{H_1} encontraremos una extensión analítica para la función ζ_{H_1} solo despejando en la expresión. Por lo que el problema ahora es cómo extender la función ξ_{H_1}. La clave está en una expresión integral encontrada por Riemann para la función ξ_{H_1}, que es la siguiente:

$$\xi_{H_1}(s) = \frac{1}{2} \int_0^{\infty} t^{\frac{s}{2}-1} (\theta(t) - 1) dt$$

Donde la función θ está dada a través de la siguiente serie:

$$\theta(t) = \sum_{n=-\infty}^{\infty} e^{-\pi n^2 t} \qquad t > 0$$

Esta función proviene del estudio de las transformadas de Fourier, que no siendo el tópico de esta obra nos limitaremos a poner la definición y las propiedades de la función theta en la sección 5.2 del apéndice.

¿De dónde proviene esta relación entre la función ξ_{H_1} y la función θ? Lo mejor es fijarse del lado que ofrece más información, que es el lado derecho. Vea que el 1 que resta a theta elimina el sumando con $n = 0$ y que los sumandos con $-n$ son iguales a los sumandos con n, por lo que la integral queda:

$$\frac{1}{2}\int_0^\infty t^{\frac{s}{2}-1}(\theta(t)-1)dt = \int_0^\infty \sum_{n=1}^\infty t^{\frac{s}{2}-1}e^{-\pi n^2 t}dt$$

Si observamos con detención, el argumento dentro de la integral asemeja a la expresión integral de la función gamma (que podemos usar ya que estamos en $H_1 \subset H$). Para reconstruir la función gamma, necesitamos hacer dos cosas: Primero hay que intercambiar el orden entre la integral y la serie. Esto se logra con el Teorema de Fubini-Tonelli, interpretando la serie como una integral de Stieljes. Para utilizar este teorema hay que probar que para algún orden la integral-suma sea absolutamente convergente, lo que no resulta complicado de evidenciar. Segundo hay que realizar el cambio de variable $v = u\pi n^2$ para así hacer aparecer la función gamma. El resultado de hacer estos cálculos es que aparecerá la función ξ_{H_1}, como se puede apreciar:

Lema 3.1.2 La función ξ_{H_1} es holomorfa y cumple con la siguiente fórmula para todo $s \in H_1$:

$$\xi_{H_1}(s) = \frac{1}{2}\int_0^\infty t^{\frac{s}{2}-1}(\theta(t)-1)dt$$

Demostración. Observe que la siguiente integral impropia con $n \in \mathbb{N}$ fijo:

$$\int_0^\infty e^{-\pi n^2 u}u^{-1+s/2}du$$

haciendo el cambio de variable $v = u\pi n^2$ se transforma en:

$$\int_0^\infty e^{-v}\left(\frac{v}{\pi n^2}\right)^{\frac{s}{2}-1}\left(\pi n^2\right)^{-1}dv = \left(\pi n^2\right)^{\frac{-s}{2}}\int_0^\infty e^{-v}v^{\frac{s}{2}-1}dv = \pi^{\frac{-s}{2}}\Gamma\left(\frac{s}{2}\right)n^{-s}$$

De la definición 5.2.1 del apéndice, la función θ está dada por la regla:

$$\theta(u) = \sum_{n\in\mathbb{Z}} e^{-\pi n^2 u} = 2\sum_{n=1}^\infty e^{-\pi n^2 u} + 1$$

Por lo que podemos calcular que:

$$\begin{aligned}\frac{1}{2}\int_0^\infty u^{\frac{s}{2}-1}(\theta(u)-1)du &= \int_0^\infty \sum_{n=1}^\infty u^{\frac{s}{2}-1}e^{-\pi n^2 u}du = \sum_{n=1}^\infty \int_0^\infty u^{\frac{s}{2}-1}e^{-\pi n^2 u}du \\ &= \pi^{\frac{-s}{2}}\Gamma\left(\frac{s}{2}\right)\sum_{n=1}^\infty n^{-s} = \pi^{\frac{-s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \xi_{H_1}(s)\end{aligned}$$

La inversión del orden de la serie y la integral impropia se puede debido al Teorema de Fubini-Tonelli, interpretando la serie como una integral de Stieljes, usable porque:

$$\begin{aligned}\sum_{n=1}^\infty \int_0^\infty \left|u^{s/2-1}e^{-\pi n^2 u}\right|du &= \sum_{n=1}^\infty \int_0^\infty \left|u^{s/2-1}\right|e^{-\pi n^2 u}du = \sum_{n=1}^\infty \int_0^\infty u^{\sigma/2-1}e^{-\pi n^2 u}du \\ &= \pi^{-\sigma/2}\Gamma(\sigma/2)\zeta(\sigma) < \infty\end{aligned}$$

Usando en este cálculo la propiedad 2.2.1 con $\sigma = \mathrm{Re}(s) > 1$.

Resulta lógico el siguiente paso: Usar la integral del lema anterior como extensión analítica para ξ_{H_1}. Lo que no ha sido natural hasta este momento es de donde proviene toda esta estrategia, la función ξ_{H_1} y su expresión integral parecen salidas sin explicación. Este planteamiento es original de Benhard Riemann, que en su obra explica que consideró otras expresiones integrales similares para extender implícitamente la función dseta. La razón porque se decantó por la función ξ_{H_1} es una propiedad muy curiosa de su extensión: es perfectamente simétrica respecto al eje $\mathrm{Re}(s) = 1/2$, hecho que tiene consecuencias cruciales para la función dseta. El siguiente teorema 3.1.3 va a explorar hasta dónde se puede definir la integral, su holomorfidad y su simetría.

La demostración formal debe ser cuidadosa y usaremos propiedades de la función theta que se encuentran en la sección 5.2 del apéndice. Los pasos de la demostración a grandes rasgos son los siguientes: Primero separamos la integral en dos integrales, una que integra desde 0 a 1 y otra de 1 al infinito. A la integral de $(1,\infty)$ no le haremos nada más allá de cambios de notación. A la integral de $(0,1)$ aplicaremos la proposición 5.2.3, que nos señala que $\theta(t) = t^{-1/2}\theta(1/t)$. Calculando se puede separar a esta integral por la suma en otras tres, dos resolubles al no contener términos de θ y otra integral que contiene el término $(\theta(1/t)-1)/2$. A esta se le aplica el cambio de variable $u = 1/t$. El resultado de todos estos cálculos es que la integral del lema 3.1.2 es igual a:

$$-\frac{1}{s} + \frac{1}{s-1} + \int_1^\infty (u^{-(s+1)/2} + u^{s/2-1})\vartheta(u)du$$

Donde $\vartheta(t) = (\theta(t)-1)/2$ es un cambio de notación para facilitar los cálculos. Esta nueva expresión resulta decidora. Primero, resulta claro que no está definida para $s = 0$ y $s = 1$. Segundo, si se realiza el cambio de variable $s' = 1 - s$ la expresión permanece inalterada, es decir, esta expresión es simétrica respecto al eje $\mathrm{Re}(s) = 1/2$. Tercero, y más trascendental, la expresión define una función ξ que es holomorfa en $\mathbb{C} - \{0,1\}$ y que tiene polos en $s = 0$ y $s = 1$. Del lema 3.1.2, esta es la extensión que buscamos, ¿pero cómo se demuestra esto último? Pues basta con probar que la integral impropia que aparece en la expresión converge en una función entera. Para probar eso, hay que probar que tanto la integral impropia como su derivada convergen para cada $s \in \mathbb{C}$. Esto se logra principalmente con dos razonamientos:

- Obtener la derivada de la integral a través de la regla de Leibniz de derivar bajo el signo de la integral. Esto es aplicable porque, gracias a que θ es continua (proposición 5.2.1 del apéndice), el argumento de la integral es continuo, derivable parcial por s y su derivada parcial es continua, además tanto esta como su derivada parcial son integrables según el segundo razonamiento.
- Ambas expresiones integrales, tanto para la función como para su derivada, convergen porque sus argumentos son exponencialmente dominados. Los "polinomios de exponente complejo" claramente lo son y θ lo es según la proposición 5.2.2 del apéndice.

Con esto se concluye una extensión analítica en casi todo $\mathbb{C}$ para la función auxiliar ξ, lo que nos deja a un paso de extender analíticamente la función dseta como es deseado en el problema objetivo de esta sección.

Teorema 3.1.3 Existe una función meromorfa ξ con dominio en todo $\mathbb{C}$ salvo en los polos simples $\{0,1\}$ que extiende a ξ_{H_1} y cumple con la fórmula:

$$\xi(s) = \xi(1-s) \quad \forall s \in \mathbb{C} - \{0,1\}$$

Demostración. De la proposición 5.2.3 sel apéndice, se sabe que: $\theta(t) = t^{-1/2}\theta(\frac{1}{t})$. Definiendo $\vartheta(t) = (\theta(t)-1)/2$, tenemos que esa identidad toma la siguiente forma:

$$\begin{aligned}\vartheta(t) &= \frac{\theta(t)-1}{2} = \frac{t^{-1/2}\theta(1/t)-1}{2} = \frac{t^{-1/2}\theta(1/t)-t^{-1/2}+t^{-1/2}-1}{2} \\ &= t^{-1/2}\vartheta(1/t) + \frac{1}{2t^{1/2}} - \frac{1}{2}\end{aligned}$$

Desde el lema 3.1.2 y aplicando la propiedad anterior, se puede hacer el siguiente cálculo:

$$\begin{aligned}\xi_{H_1}(s) &= \frac{1}{2}\int_0^\infty t^{s/2-1}(\theta(t)-1)dt = \int_0^\infty t^{s/2-1}\vartheta(t)dt \\ &= \int_0^1 t^{s/2-1}\vartheta(t)dt + \int_1^\infty t^{s/2-1}\vartheta(t)dt \\ &= \int_0^1 t^{s/2-1}\left(t^{-1/2}\vartheta(1/t) + \frac{1}{2t^{1/2}} - \frac{1}{2}\right)dt + \int_1^\infty t^{s/2-1}\vartheta(t)dt \\ &= -\frac{1}{2}\int_0^1 t^{s/2-1}dt + \frac{1}{2}\int_0^1 t^{(s-3)/2}dt + \int_0^1 t^{(s-3)/2}\vartheta(1/t)dt + \int_1^\infty t^{s/2-1}\vartheta(t)dt \\ &= -\frac{t^{s/2}}{s}\Big|_0^1 + \frac{t^{(s-1)/2}}{s-1}\Big|_0^1 + \int_1^\infty u^{-(s+1)/2}\vartheta(u)du + \int_1^\infty t^{s/2-1}\vartheta(t)dt \\ &= -\frac{1}{s} + \frac{1}{s-1} + \int_1^\infty (u^{-(s+1)/2} + u^{s/2-1})\vartheta(u)du\end{aligned}$$

Donde a la integral de 0 a 1 se aplicó el cambio de variable $u = 1/t$ y para la integral de 1 a infinito se aplicó el cambio de variable $u = t$.

Lo siguiente es comprobar que la integral impropia converge a una función entera. Esto es claro por las siguientes dos razones:

- **La integral impropia converge:** De la proposición 5.2.2 del apéndice sabemos que existe $C_1 > 0$ tal que $\vartheta(u) \le C_1 e^{-\pi u}$. Para cada $s \in \mathbb{C}$ fijo, $u^{-(s+1)/2} + u^{s/2-1}$ es exponencialmente dominado por ser prácticamente un polinomio (solo con exponentes complejos en vez de enteros), o sea, existe $C_2 > 0$ tal que $u^{-(s+1)/2} + u^{s/2-1} \le C_2 e^t$. Esto nos permite concluir que para todo $s \in \mathbb{C}$:

  $$\int_1^\infty |u^{-(s+1)/2} + u^{s/2-1}|\vartheta(u)du \le C_1C_2\int_1^\infty e^{(1-\pi)u}\vartheta(u)du = C_1C_2\frac{e^{1-\pi}}{\pi-1} < \infty$$

  Luego hay convergencia absoluta.
- **Su derivada compleja existe en todo s:** El argumento de la integral propia se puede derivar parcialmente en s ya que, para cada $u \in (1,\infty)$ fijo, es una combinación lineal de funciones exponenciales. Su derivada parcial es:

  $$\left(u^{s/2-1} - u^{-(s+1)/2}\right)\frac{\vartheta(u)\log u}{2}$$

  De la proposición 5.2.1 se deduce que ϑ es continua, luego se tiene que el argumento de la integral es continuo para $(s,u) \in \mathbb{C}\times(1,\infty)$ al ser producto de funciones continuas conocidas. Luego aplicando la regla de Leibniz de derivar bajo el signo de la integral da:

$$\frac{\partial}{\partial s}\int_1^\infty \left(u^{-(s+1)/2}+u^{s/2-1}\right)\vartheta(u)du = \int_1^\infty \left(u^{s/2-1}-u^{-(s+1)/2}\right)\frac{\vartheta(u)\log u}{2}du$$

Dado lo que acompaña a ϑ en el argumento de esta última integral impropia también es exponencialmente dominado, con un razonamiento idéntico al del paso anterior se comprueba que la integral impropia converge para todo $s \in \mathbb{C}$. Luego la integral original tiene derivada en cada punto del plano complejo, ergo es una función entera.

En consecuencia de lo anterior, la función dada por la regla:

$$\xi(s) = -\frac{1}{s}+\frac{1}{s-1}+\int_1^\infty (u^{-(s+1)/2}+u^{s/2-1})\vartheta(u)du$$

define una función meromorfa definida en todo $\mathbb{C}$ salvo los polos en 0 y 1, extendiendo la función ξ_{H_1} por el lema 3.1.2. Además se tiene que:

$$\begin{aligned}\xi(1-s) &= -\frac{1}{1-s}+\frac{1}{(1-s)-1}+\int_1^\infty (u^{-(1-s+1)/2}+u^{(1-s)/2-1})\vartheta(u)du\\ &= \frac{1}{s-1}-\frac{1}{s}+\int_1^\infty (u^{s/2-1}+u^{-(s+1)/2})\vartheta(u)du\\ &= \xi(s)\end{aligned}$$

Con la función ξ extendida para $\mathbb{C}-\{0,1\}$, solo queda concluir la extensión analítica que buscamos para la función dseta. La única pregunta abierta sería por qué la extensión de dseta estará definida para $s=1$ si ξ no lo está, pero esto se debe a que el cero de $1/\Gamma$ se compensa con el polo de la función ξ en ese mismo punto. Dicho aquello, pasemos al esperado resultado final de esta sección:

Corolario 3.1.4 Existe una función meromorfa ζ, con dominio en todo $\mathbb{C}$ salvo en el polo simple 1, que extiende a ζ_H.

Demostración. Se define $\zeta : \mathbb{C}-\{1\} \to \mathbb{C}$ como:

$$\zeta(s) = \frac{\pi^{\frac{s}{2}}\xi(s)}{\Gamma(\frac{s}{2})}$$

Salvo para $s=0$ donde se define por el límite de esta exprensión a dicho punto.

Es conocido que $\pi^{s/2}$ es una función entera sin ceros. Del teorema 3.1.3, ξ es meromorfa con polos simples en 0 y 1. Del corolario 2.3.5, $\frac{1}{\Gamma(s/2)}$ es una función entera con ceros simples en $\{0,-2,-4,...\}$. Luego se tiene que para $s=0$, el polo y el cero se cancelan, por lo que ζ está bien definido, no se anula y es holomorfa en $s=0$. Para $s=1$, ζ hereda el polo simple de ξ. Para el resto de los puntos, es el producto de funciones holomorfas, luego ζ es holomorfa. Finalmente, del teorema 3.1.3 y la definición 3.1.2, se tiene que es una extensión meromorfa de ζ_{H_1}.

3.2 Identidad de Euler

La importancia de la función dseta en las matemáticas es su conexión íntima con los números primos. Este nexo viene dado por una identidad descubierta por Leonhard Euler, que permite expresar la función dseta como un producto en términos de todos los números primos, más en específico:

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

Esto se cumple solamente para la sección del plano complejo H_1, donde la función dseta puede escribirse como una serie. Esta identidad es crucial y tiene muchas consecuencias que se estudiarán no solo en esta sección, sino durante gran parte del resto de la obra.

La idea intuitiva detrás de la identidad se encuentra en el Teorema Fundamental de la Aritmética: Cada número natural es el producto de una combinación única de números primos, sin importar el orden. Mientras que la función dseta en H_1 es la suma de todos los naturales elevados a $-s$, el argumento del producto en la identidad de Euler son series geométricas de p^{-s} con p primo, luego intuitivamente podemos ver la identidad de Euler como la siguiente expresión:

$$1+\frac{1}{2^s}+\frac{1}{3^s}+\frac{1}{4^s}+\frac{1}{5^s}+\ldots = \left(1+\frac{1}{2^s}+\frac{1}{2^{2s}}+\ldots\right)\left(1+\frac{1}{3^s}+\frac{1}{3^{2s}}+\ldots\right)\ldots$$

Resulta sencillo observar de dónde viene la igualdad: Si el lado derecho lo expandimos, tendremos que el lado derecho sumará todas las combinaciones de productos de p^{-ks}, con p primo y k natural. Cuando en una de esas combinaciones haya infinitos factores distintos a 1, el término se hará cero y no contribuirá a la suma. Pero cuando solo haya finitos factores distintos de 1 corresponderá a un único natural elevado a $-s$ y estándo de forma única todas las combinaciones, estarán una única vez cada natural elevado a $-s$. Es decir, cuando se expande el lado derecho estarán todos los sumandos del lado izquierdo más otros sumandos que serán despreciables. Por ejemplo, el sumando $28^{-}s$ del lado izquierdo aparecerá del producto de 2^{-2s}, 7^{-s} y los demás 1s.

Todo muy bien hasta aquí. ¿Pero cómo demostrar esta identidad formalmente? La demostración se dividirá en dos partes, aunque la primera parte también se subdividirá en otras dos subpartes. La primera parte consiste en demostrar la identidad para los reales mayores que 1. La segunda parte consiste en demostrar que el producto define una función holomorfa en H_1, luego dos funciones holomorfas que son iguales en un conjunto que contiene al menos una sucesión convergente de puntos distintos (como los reales mayores que 1) tienen que ser iguales en todo H_1 según el corolario 2.1.2. Cabe destacar que este paso demuestra la potencia del corolario 2.1.2, que nos reduce toda la tarea de demostrar que dos funciones holomorfas son iguales en H_1 a demostrarlo para un conjunto más pequeño.

Para demostrar la identidad de Euler en los reales se procede con la tradicional táctica de probar la desigualdad en ambos sentidos. Esto porque para probar que un límite obedece una desigualdad basta con probar que la sucesión la cumple. Por lo que se trata de demostrar que la forma parcial de un lado de la identidad es menor o igual que el otro lado de la identidad. Esto no requiere dificultad si se sigue la idea intuitiva. Veamos que la forma parcial del producto

$$\prod_{\substack{p \text{ primo} \\ p \leq P}} \sum_{k=0}^{K} p^{-ks}$$

se le puede aplicar la propiedad distributiva quedando una suma de naturales distintos elevados a $-s$, y como todo es positivo, tiene que ser menor que la serie.

Por otro lado, para la forma parcial de la serie:

$$\sum_{n=0}^{N} \frac{1}{n^s}$$

Siguiendo la idea intuitiva, todos los sumandos de esta suma parcial deben aparecer para alguna forma parcial del producto infinito, porque el teorema fundamental de la aritmética asegura que los números del 1 al N sean generados por una secuencia finita de números primos, que evidentemente alcanzan hasta un exponente máximo finito. Poniendo P al primo mayor de la secuencia y K el exponente máximo, se tiene la forma parcial que incluye todos los sumandos de la suma parcial. Como estamos sumando cosas positivas y multiplicando términos mayores que 1, se logra que la suma parcial es menor que algún producto parcial que es menor que el producto infinito.

Finalmente la prueba de la segunda parte, es decir, que el producto define una función holomorfa, es sencillamente la utilización de una versión modificada del Criterio de Weierstrass para productos infinitos (Teorema 5.1.6 del Apéndice), similar a la demostración de la proposición 3.1.1. Por lo que podemos proceder a la prueba formal:

Teorema 3.2.1 — Identidad de Euler. Si $\mathrm{Re}(s) > 1$, entonces:

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

Demostración. Primero, demostraremos la igualdad para $s > 1$ real.

Sea $s > 1$ real y $P, K \in \mathbb{N}$ cualesquiera. Considere el siguiente factor parcial:

$$\prod_{\substack{p \text{ primo} \\ p \leq P}} \sum_{k=0}^{K} p^{-ks}$$

Aplicando la propiedad distributiva, corresponderá a la suma de todas las combinaciones de producto entre primos menores a P elevados a un valor entero entre 0 y K, estando cada término de esta suma elevado a su vez a $-s$. Por teorema fundamental de la aritmética, cada término corresponderá a un natural distinto elevado a $-s$, por lo que como son sumas de términos positivo se puede concluir que:

$$\prod_{\substack{p \text{ primo} \\ p \leq P}} \sum_{k=0}^{K} p^{-ks} \leq \sum_{n=1}^{\infty} \frac{1}{n^s}$$

Tomando $P, K \to \infty$ se logra:

$$\prod_{p \text{ primo}} \frac{1}{1-p^{-s}} = \prod_{p \text{ primo}} \sum_{k=0}^{\infty} p^{-ks} \leq \sum_{n=1}^{\infty} \frac{1}{n^s}$$

Por otro lado, sea $N \in \mathbb{N}$. Por el teorema fundamental de la aritmética, cada natural corresponde a un producto de potencias de primos. Sea P' un natural mayor al mayor primo que es factor de algún natural entre 1 a N. Sea K' un natural mayor que el máximo exponente que tenga algún factor primo al dividir a algún natural entre 1 a N. Luego los números naturales de 1 a N son completamente generados por los productos de números primos menores que P' elevados por un valor entre 0 y K'. Por tanto, todos los sumandos de la suma parcial están en la forma parcial del producto infinito con P' y K', y siendo todas las sumas de términos positivos y productos de valores mayores que 1, se concluye:

$$\sum_{n=1}^{N} \frac{1}{n^s} \leq \prod_{\substack{p \text{ primo} \\ p \leq P'}} \sum_{k=0}^{K'} p^{-ks}$$

Tendiendo $P', K' \to \infty$:

$$\sum_{n=1}^{N} \frac{1}{n^s} \leq \prod_{p \text{ primo}} \frac{1}{1-p^{-s}}$$

Finalmente de $N \to \infty$:

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ primo}} \frac{1}{1-p^{-s}}$$

Esto también comprueba la convergencia del productorio en este dominio a través del teorema de acotamiento gracias a que la serie converge en este dominio.

Segundo, se demostrará que el productorio define una función holomorfa en H_1. Sea $z \in H_1$ cualquiera y denotemos como $\sigma = \mathrm{Re}(z)$ y $\tau = \mathrm{Im}(z)$. Sea $a = \frac{\sigma+1}{2}$. Vea que:

$$\left| \frac{1}{1-p^{-s}} - 1 \right| = \left| \frac{p^{-s}}{1-p^{-s}} \right| = \left| \frac{1}{p^s - 1} \right| \leq \frac{1}{|p^s| - 1} = \frac{1}{p^\sigma - 1} \leq \frac{1}{p^a - 1}$$

Y como son sumas de términos positivos se cumple:

$$\sum_{p \text{ primo}} \frac{1}{p^a - 1} < \sum_{n=1}^{\infty} \frac{1}{n^a - 1} < \infty$$

Y la última serie converge por criterio de comparación al límite del cociente con $\zeta(a)$ que es convergente al ser $a > 1$:

$$\lim_{n\to\infty} \frac{\frac{1}{n^a-1}}{\frac{1}{n^a}} = \lim_{n\to\infty} \frac{n^a}{n^a - 1} = \lim_{n\to\infty} \frac{1}{1 - \frac{1}{n^a}} = 1$$

De esto se deduce que los productos parciales $\prod_{\substack{p \text{ primo} \\ p \leq P}} \frac{1}{1-p^{-s}}$ convergen uniformemente al producto infinito por el teorema 5.1.6. Como $1/\left(1-p^{-s}\right)$ son holomorfas en $\mathrm{Re}(s) > a$, entonces el producto infinito es holomorfo en esa región por el teorema 5.1.2, luego es holomorfo en z y siendo $z \in H_1$ cualquiera, se tiene que el producto define una función holomorfa en H_1. Finalmente el producto infinito y la función dseta son dos funciones holomorfas que coinciden en un conjunto que contiene sucesiones convergentes de puntos distintos (los reales mayores que 1), por lo que la identidad de Euler se extiende a todo el abierto conexo H_1 por el corolario 2.1.2.

Por interés histórico, agregamos un corolario que es la demostración analítica de Leonhard Euler sobre la infinidad de números primos. Este fue el objetivo inicial con que fue descubierta esta identidad y fue uno de las primeras demostraciones analíticas sobre un hecho de teoría de números, aunque fuese uno ya conocido.

Antes de entrar en este corolario, nótese primero que, en la demostración anterior, para probar la identidad de Euler en los reales no necesitamos nunca el supuesto que los números primos era infinitos. Sin embargo, sí asumimos implícitamente aquello en nuestro tratamiento de la segunda parte, pero si los números primos no fueran infinitos el resultado se mantendría ya que el productorio daría una función holomorfa por solo ser una multiplicación finita de funciones holomorfas.

La demostración es muy directa desde la identidad de Euler y el hecho que ζ tiene un polo en $s = 1$.

Corolario 3.2.2 Existen una cantidad infinita de números primos.

Demostración. Suponga que existe una cantidad finita de números primos. Como el productorio de la identidad de Euler define una función continua, se tiene que:

$$\lim_{s\to 1} \prod_{p \text{ primo}} \frac{1}{1-p^{-s}} = \prod_{p \text{ primo}} \frac{1}{1-p^{-1}} < \infty$$

Pero de la identidad de Euler también es cierto que:

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

Ya que la función dseta tiene un polo en $s = 1$. Hay una contradicción, por reducción al absurdo existe una cantidad infinita de números primos.

En el resto de la sección, se demostrarán dos lemas fundamentales para la prueba del Teorema de los Números Primos. El primero son dos formas equivalentes de expresar $\log \zeta$ en H_1, que son el equivalente para esa función de la Identidad de Euler. La demostración de la primera igualdad es la más interesante:

$$\log \zeta(s) = \sum_{\substack{k\in\mathbb{N} \\ p \text{ primo}}} \frac{p^{-ks}}{k}$$

Es una doble serie que surge de aplicar el logaritmo a la Identidad de Euler. Por continuidad, el logaritmo entra en el límite y con sus propiedades convierte el producto en una suma. La segunda serie viene usar la expansión de serie de Taylor de $\log(1/(1-x))$. Pero esta Serie de Taylor solo puede usarse para $x \in (0,1)$, luego solo podemos probar de esta forma para $s > 1$ real. Para los demás $s \in H_1$, tenemos que usar el corolario 2.1.2 como en la demostración anterior: Probar que la doble serie define una función holomorfa en H_1 vía Criterio de Weierstrass aplicado 2 veces, luego como coincide con $\log \zeta$ en los reales mayores que 1, tiene que coincidir en todo H_1 tomando la rama principal del logaritmo. En otro punto, la notación señala que el orden en que se apliquen las series no importa, para demostrar este hecho hay que usar el Teorema de Fubini-Tonelli.

Para la segunda igualdad no hay complejidad. Sencillamente hay que escoger una sucesión de coeficientes c_n conveniente para generar la igualdad con la doble serie. De todas formas, la consecuencia de esta última forma es importante, porque permite simplificar a $\log \zeta$ en una serie simple, lo que será usado más adelante.

Lema 3.2.3 Si $\mathrm{Re}(s) > 1$, para algunos $c_n \geq 0$ entonces para la rama principal del logaritmo:

$$\log \zeta(s) = \sum_{\substack{k \in \mathbb{N} \\ p \text{ primo}}} \frac{p^{-ks}}{k} = \sum_{n=1}^{\infty} c_n n^{-s}$$

Demostración. La segunda igualdad se deduce de forma directa, escogiendo $c_n = 1/k$ para cuando $n = p^k$ con p primo y $c_n = 0$ en cualquier otro caso. Por lo que basta con probar la primera igualdad:

Supongamos que $s > 1$ es real. Entonces se calcula:

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

Donde la primera igualdad proviene de la Identidad de Euler, la segunda por continuidad del logaritmo, y la última proviene de que $p^{-s} \in (0,1)$ que permite usar la serie de Taylor:

$$\log \left(\frac{1}{1-x} \right) = \sum_{k=1}^{\infty} \frac{x^k}{k} \qquad \forall x \in (0,1)$$

Sea $s \in H_1$ cualquiera y sea $\sigma = \mathrm{Re}(s) > 1$. Aplicando lo anterior y la proposición 2.2.1:

$$\sum_{p \text{ primo}} \sum_{k=1}^{\infty} \left| \frac{p^{-ks}}{k} \right| = \sum_{p \text{ primo}} \sum_{k=1}^{\infty} \frac{p^{-k\sigma}}{k} = \log \zeta(\sigma) < \infty$$

Interpretando las series como Integrales de Riemann-Stieljes, el Teorema de Fubini-Tonelli permite conmutar el orden de las series.

Sea $z \in H_1$ cualquiera y $a = (\mathrm{Re}(z) + 1)/2$. Se tiene en todo $\sigma = \mathrm{Re}(s) > a$ que:

$$\left| \frac{p^{-ks}}{k} \right| = \frac{p^{-k\sigma}}{k} \leq \frac{p^{-ka}}{k} \qquad \left| \sum_{k=1}^{\infty} \frac{p^{-ks}}{k} \right| \leq \sum_{k=1}^{\infty} \left| \frac{p^{-ks}}{k} \right| \leq \sum_{k=1}^{\infty} \frac{p^{-ka}}{k}$$

Ambas cotas superiores son sumables ya que:

$$\sum_{k=1}^{\infty} \frac{p^{-ka}}{k} = \log \left(\frac{1}{1 - p^{-a}} \right) < \infty \qquad \sum_{p \text{ primo}} \sum_{k=1}^{\infty} \frac{p^{-ka}}{k} = \log \zeta(a) < \infty$$

Las desigualdades de la izquierda válidas para todo p primo fijo. Luego por criterio de Weierstrass, las sumas parciales $\sum_{k=1}^{K} \frac{p^{-ks}}{k}$ convergen uniforme en $\mathrm{Re}(s) > a$. Como $\frac{p^{-ks}}{k}$ es holomorfo, por teorema 5.1.2 la serie $\sum_{k=1}^{\infty} \frac{p^{-ks}}{k}$. Por argumento análogo, pero usando las desigualdades de la derecha, se llega que la serie doble es holomorfa en $\mathrm{Re}(s) > a$, en particular en z y como este es cualquiera en H_1, la serie doble es holomorfa en H_1.

Finalmente, tomando la rama principal del logaritmo complejo, $\log \zeta$ es holomorfa en H_1 al ser composición de funciones holomorfas y coincide con la doble serie en el segmento $s > 1$ como se probó más arriba, luego por el corolario 2.1.2 coinciden en todo H_1.

El último lema de la sección es una desigualdad ideada por Charles-Jean de la Vallée Poussin con que logró reducir las dos primeras demostraciones del teorema de los números primos. Este lema es crucial para demostrar que la función dseta no tiene ceros en $\mathrm{Re}(s) = 1$, que a su vez es la clave del teorema. La demostración es un cálculo ingenioso que se basa en las siguientes ideas:

- El logaritmo del módulo es igual a la parte real del logaritmo complejo (para cualquier rama), esto permitirá deshacerse de los módulos.
- Usar la segunda serie del lema anterior para cambiar $\log \zeta$.
- Como la parte real es una función continua, se puede introducir dentro de la serie.
- La parte real de $e^{i\Theta}$ es $\cos(\Theta)$, luego todo se convierte en una identidad trigonométrica.

A continuación, veamos cómo funcionan estas ideas en acción consecutiva:

Lema 3.2.4 Si $z = x + iy$ con $x > 1$ e $y \in \mathbb{R}$, entonces:

$$\log\left|\zeta^3(x)\zeta^4(x+iy)\zeta(x+2iy)\right| \geq 0$$

Demostración. Un resultado preliminar: Sea $s \in \mathbb{C}$ expresado de forma polar como $|s|e^{i\theta}$. Aplicando logaritmo se tiene que:

$$\log(s) = i\theta + \log|s|$$

Por lo tanto, para cualquier rama del logaritmo complejo, su parte real coincide con el logaritmo del módulo de s.

Dicho aquello, se puede calcular lo siguiente:

$$\begin{aligned}
&\log|\zeta^3(x)\zeta^4(x+iy)\zeta(x+2iy)| \\
&= 3\log|\zeta(x)| + 4\log|\zeta(x+iy)| + \log|\zeta(x+2iy)| \\
&= 3\mathrm{Re}\left(\log\zeta(x)\right) + 4\mathrm{Re}\left(\log\zeta(x+iy)\right) + \mathrm{Re}\left(\log\zeta(x+2iy)\right) \\
&= 3\mathrm{Re}\left(\sum_{n=1}^{\infty} c_n n^{-x}\right) + 4\mathrm{Re}\left(\sum_{n=1}^{\infty} c_n n^{-x-iy}\right)| + \mathrm{Re}\left(\sum_{n=1}^{\infty} c_n n^{-x-2iy}\right) \\
&= \sum_{n=1}^{\infty} 3c_n n^{-x} + 4c_n\mathrm{Re}(n^{-x-iy}) + c_n\mathrm{Re}(n^{-x-2iy}) \\
&= \sum_{n=1}^{\infty} 3c_n n^{-x} + 4c_n\mathrm{Re}(e^{-x\log n - iy\log n}) + c_n\mathrm{Re}(e^{-x\log n - 2iy\log n}) \\
&= \sum_{n=1}^{\infty} 3c_n n^{-x} + 4c_n\mathrm{Re}(n^{-x}\mathrm{cis}(-y\log n)) + c_n\mathrm{Re}(n^{-x}\mathrm{cis}(-2y\log n)) \\
&= \sum_{n=1}^{\infty} c_n n^{-x}(3 + 4\cos(-y\log n) + \cos(-2y\log n)) \\
&= \sum_{n=1}^{\infty} c_n n^{-x}(2 + 4\cos(-y\log n) + 2\cos^2(-y\log n)) \\
&= \sum_{n=1}^{\infty} 2c_n n^{-x}(1 + \cos(-y\log n))^2
\end{aligned}$$

Donde en la tercera línea se aplicó el lema 3.2.3, que nos asegura también que $c_n \geq 0$. Luego todos los términos del argumento de la serie son no negativos, luego el resultado de la serie debe ser no negativo, demostrando lo que se quería.

3.3 Estimaciones de su crecimiento

En los pasos finales de la demostración del teorema de los números primos estaremos trabajando con integrales que contienen dentro la función dseta y su derivada. Calcular algunos valores de la función dseta ya es una odisea, calcular una integral sobre una expresión complicada que involucra la función dseta es una tarea titánica. Pero no requerimos calcular el valor exacto de la función, ni de dichas integrales, sino que para la demostración nos bastará una estimación de cómo se comportan en ciertas circunstancias. La presente sección tratará sobre dos teoremas que nos permitirán generar estimaciones de la función dseta, su derivada y su inverso multiplicativo.

Pero antes de ir a esos teoremas, será necesario indagar en otra forma de expresar la función dseta en el casi semiplano complejo $H-\{1\}$ en el lema 3.3.1 y su corolario 3.3.2. Esta forma de expresar la función dseta es como una serie de funciones enteras (más un término de corrección) que tienen tres propiedades especiales detalladas en el lema 3.3.1. ¿Cómo son estas funciones enteras? En la demostración del lema se caracterizan explícitamente:

$$f_n(s)=\int_n^{n+1}\left(\frac{1}{n^s}-\frac{1}{x^s}\right)dx$$

Si se recuerda del cálculo, la demostración que la serie $\zeta_{\mathbb{R}}(s)$ converge en $s>1$ y diverge para $s<1$ viene por el criterio de la integral, cuya convergencia es más fácil de estudiar. Estas funciones pueden pensarse como que van midiendo la separación entre esa integral de $1/x^s$ y la suma de $1/n^s$ que ocurre el intervalo de $(n,n+1)$. De esto trata la tercera propiedad: La suma parcial hasta N de $f_n(s)$ da la separación entre la integral y la suma parcial en el intervalo $(1,N)$. Las otras dos propiedades son cotas de estas funciones $f_n(s)$, que serán fundamentales para cuando estudiemos como acotar o estimar el crecimiento de la función dseta en los teoremas siguientes de esta sección.

La demostración es directa. Las funciones f_n son enteras por el teorema 5.1.1. La primera estimación viene del teorema del valor medio aplicado a $g(z)=z^{-}s$ para el intervalo $[n,x]$. La segunda viene directamente de acotar la integral que define a f_n. Finalmente, la última propiedad sale de sumar directamente. Veamos la prueba formal:

Lema 3.3.1 Existe una sucesión de funciones enteras $\{f_n\}_{n\in\mathbb{N}}$ con las siguientes tres propiedades

1) $|f_n(s)|\leq |s|/n^{\mathrm{Re}(s)+1}$
2) $|f_n(s)|\leq 2/n^{\mathrm{Re}(s)}$
3) $\sum_{1\leq n<N}\frac{1}{n^s}-\int_1^N\frac{dx}{x^s}=\sum_{1\leq n<N}f_n(s)\quad\forall s\in\mathbb{C}$

Con $N\in\mathbb{N}-\{1\}$ cualquiera.

Demostración. Considere la sucesión de funciones $f_n:\mathbb{C}\to\mathbb{C}$ dadas por la regla:

$$f_n(s)=\int_n^{n+1}\left(\frac{1}{n^s}-\frac{1}{x^s}\right)dx=\int_0^1\left(\frac{1}{n^s}-\frac{1}{(n+y)^s}\right)dy=\int_0^1 F(s,y)dy$$

Comprobando que $F(s,y)$ es claramente continua en $\mathbb{C}\times[0,1]$ y para cada $y_0\in\mathbb{C}$, $F(s,y_0)$ es holomorfa siendo una exponencial, tenemos por el teorema del apéndice que $f_n(s)$ es entera para todo $n\in\mathbb{N}$.

Para obtener la primera propiedad, sea $x \in [n, n+1]$. Aplicando el teorema de valor medio para $g(z) = z^{-s}$ entre x^{-s} y n^{-s}, sabemos que existe $y \in [n, x]$ tal que:

$$\frac{n^{-s} - x^{-s}}{n - x} = \frac{-s}{y^{s+1}}$$

Lo que nos lleva a que:

$$\left| \frac{1}{n^s} - \frac{1}{n^s} \right| = \frac{|s|}{|y^{-1-s}|} |n - x| \leq \frac{|s|}{|y^{-1-s}|} = \frac{|s|}{y^{1+\sigma}} \leq \frac{|s|}{n^{1+\sigma}}$$

Con $\sigma = \mathrm{Re}(s) > 1$. Aplicando a f_n se logra el acotamiento deseado:

$$|f_n(s)| \leq \int_n^{n+1} \left| \frac{1}{n^s} - \frac{1}{x^s} \right| dx \leq \int_n^{n+1} \frac{|s|}{n^{1+\sigma}} dx = \frac{|s|}{n^{1+\sigma}}$$

La segunda propiedad sale de forma directa:

$$|f_n(s)| \leq \int_n^{n+1} \left| \frac{1}{n^s} - \frac{1}{x^s} \right| dx \leq \int_n^{n+1} \left| \frac{1}{n^s} \right| + \left| \frac{1}{x^s} \right| dx \leq \left| \frac{2}{n^s} \right| \int_n^{n+1} dx = \frac{2}{n^\sigma}$$

con $\sigma = \mathrm{Re}(s)$.

Para finalizar la demostración demostrando la tercera propiedad, basta aplicar la regla de f_n a la suma:

$$\sum_{n=1}^{N} f_n(s) = \sum_{n=1}^{N} \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s} \right) dx = \sum_{n=1}^{N} \left(\frac{1}{n^s} - \int_n^{n+1} \frac{dx}{x^s} \right) = \sum_{n=1}^{N} \frac{1}{n^s} - \int_1^{N+1} \frac{dx}{x^s}$$

El corolario concluye el objetivo de construir esta forma alternativa de expresar la función dseta en $H - \{1\}$. Su enunciado señala que la serie de funciones f_n es holomorfa en H y da la función dseta menos $1/(s-1)$. Demostrarlo no reviste complicación: Usando la primera propiedad del lema 3.3.1, se puede probar que la serie es una función holomorfa en H usando el criterio de Weierstrass (teorema 5.1.4 del apéndice) en la forma usual que se ha hecho en demostraciones anteriores. Para la relación entre la serie y la función dseta, basta con tomar el límite $N \to \infty$ en la tercera propiedad del lema 3.3.1. Pero esto solo servirá para H_1, donde la función dseta se puede expresar como una serie. Para extender al resto de H basta usar el corolario 2.1.2 aprovechando que tanto la serie como $\zeta(s) - 1/(s-1)$ sn holomorfas en H (ya que el $-1/(s-1)$ cancela el polo de la función dseta en $s = 1$).

Antes de introducirnos a la prueba formal del corolario 3.3.2, un comentario interesante: Tanto este corolario como el lema 3.3.1 pueden escribirse como una vía alternativa para extender analíticamente la función dseta a H. Con algunos agregados, esta vía también puede usarse para ir extendiendo meromorfa la función dseta para conjuntos de la forma $\mathrm{Re}(s) > -n$, llegando paso a paso a extender a todo $\mathbb{C} - \{1\}$, similar a lo que sucede con la extensión analítica de la función gamma. Proseguir con esta vía alternativa no nos otorgará información necesaria para demostrar el teorema de los números primos, por lo que no se ahondará más en este camino. Pero el lector interesado puede consultar [14], especialmente los problemas 2 y 3 del capítulo 6.

Corolario 3.3.2 Para $s \in H - \{1\}$, se tiene que:

$$\zeta(s) - \frac{1}{s-1} = f(s)$$

Donde $f(s) = \sum_{n=1}^{\infty} f_n(s)$ define una función holomorfa en H.

Demostración. Comenzamos con probar que $f(s)$ es una función holomorfa bien definida en H. Sea $z \in H$, defina $a = \mathrm{Re}(s)/2$ y $b = 2|z|$. De la primera propiedad del lema 3.3.1, para todo $s \in H$ tal que $\sigma = \mathrm{Re}(s) > a$ y $|s| < b$ se tiene que:

$$|f_n(s)| \leq \frac{|s|}{n^{\sigma+1}} \leq \frac{b}{n^{a+1}}$$

La serie $\sum_{n=1}^{\infty} bn^{-a-1}$ converge a $b\zeta(a+1)$, luego por Criterio de Weierstrass (teorema 5.1.4) las sumas parciales de $f(s)$ convergen uniformente en la intersección de $\mathrm{Re}(s) > a$ y $B(0,b)$. Como f_n son enteras para todo $n \in \mathbb{N}$, se concluye que la serie converge a una función holomorfa f en esta intersección. En particular, f converge y es holomorfa en z. Como $z \in H$ cualquiera, se tiene f es una función holomorfa bien definida en H.

De la segunda propiedad del teorema anterior, se obtiene para H_1 cuando $N \to \infty$:

$$f(s) = \sum_{n=1}^{\infty} f_n(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} - \int_1^{\infty} \frac{dx}{x^s} = \zeta(s) - \frac{1}{s-1} \quad \forall s \in H_1$$

$\zeta - 1/(s-1)$ es una función holomorfa en H, luego su coincidencia con f en H_1 se extiende a todo el abierto conexo $H - \{1\}$ por el corolario 2.1.2.

El próximo teorema 3.3.3 ofrece estimaciones para la función dseta y su derivada. Usando la notación O grande de Landau[2], el teorema señala que la función dseta y su derivada son $O(|\mathrm{Im}(s)|^{\varepsilon})$ para cualquier $\varepsilon > 0$ y s con $\mathrm{Re}(s) \geq 1$ y $|\mathrm{Im}(s)| \geq 1$. Para la franja entre $\mathrm{Re}(s) \in (0,1)$, se puede ampliar el resultado de la función dseta a $O(|\mathrm{Im}(s)|^{1-\gamma+\varepsilon})$ para $0 \leq \gamma \leq \mathrm{Re}(s)$. Este resultado será de gran importancia en el último paso de la demostración.

¿Y cómo se demuestra? Usando de forma cuidadosa los acotamientos del lema 3,3,1 y aplicando los resultados al corolario 3,3,2, se puede llegar a la siguiente desigualdad:

$$|\zeta(s)| \leq \left|\frac{1}{s-1}\right| + 2\,|s|^{1-\gamma+\varepsilon} \sum_{n=1}^{\infty} \frac{1}{n^{1+\varepsilon}}$$

La idea siguiente es acotar los términos en función de s por $|\mathrm{Im}(s)|^{1-\gamma+\varepsilon}$. Para $\left|\frac{1}{s-1}\right|$ es directo considerando $|\mathrm{Im}(s)| \geq 1$. Para el otro término hay que separar dos casos. Si $\mathrm{Re}(s) \leq 1+\varepsilon$, entonces es bien directo encontrar $|s|^{1-\gamma+\varepsilon} \leq c_{\varepsilon}^{(1)} |s|^{1-\gamma+\varepsilon}$. En cambio, si $\mathrm{Re}(s) > 1+\varepsilon$, la proposición 2.2.1 implica que $|\zeta(s)| \leq \zeta(\mathrm{Re}(s))$. Y en los reales, es fácil ver que ζ es decreciente. Luego $|\zeta(s)| \leq \zeta(1+\varepsilon)$ y podemos aplicar lo concluido para el otro caso.

Para demostrar el acotamiento a la derivada de dseta, la prueba consiste en aplicar la Fórmula Integral de Cauchy para caracterizar ζ', aplicar la parametrización usual de un círculo y luego utilizar la cota encontrada para dseta. Con un cálculo cuidadoso se logra el resultado deseado.

[2]Una función f es $O(g)$ si existe una constante $c > 0$ tal que $|f(s)| \leq c|g(s)|$. Esta útil notación es original del gran matemático judío alemán Edmund Landau.

Como última observación, en el enunciado del teorema se exige que la constante sea la misma para ambos acotamientos. Esto no es necesario, pero es conseguible fácilmente tomando la mayor de las dos constantes. Dicho aquello, pasemos a la demostración formal:

Teorema 3.3.3 Sea $s = \sigma + i\tau$ con $\sigma, \tau \in \mathbb{R}$. Entonces para cada $\gamma \in [0,1]$ y $\varepsilon > 0$, existe una constante c_ε tal que:

1) $|\zeta(s)| \leq c_\varepsilon |\tau|^{1-\gamma+\varepsilon}$, para $\sigma \geq \gamma$ e $|\tau| \geq 1$
2) $|\zeta'(s)| \leq c_\varepsilon |\tau|^{\varepsilon}$ para $\sigma \geq 1$ al igual que $|\tau| \geq 1$

Demostración. Sea $\varepsilon > 0$ y $\gamma \in [0,1]$ cualesquiera. Sea $s = \sigma + i\tau$ con $\sigma, \tau \in \mathbb{R}$ cualesquiera mientras que $\sigma \geq \gamma$ y $|\tau| \geq 1$. Defina $\delta = 1 - \gamma + \varepsilon > 0$. Luego, utilizando las funciones del lema 3.3.1 y sus cotas, se sabe que:

$$|f_n(s)| = |f_n(s)|^\delta |f_n(s)|^{1-\delta} \leq \left(\frac{|s|}{n^{\sigma+1}}\right)^\delta \left(\frac{2}{n^\sigma}\right)^{1-\delta} = \frac{2^{1-\delta}|s|^\delta}{n^{\sigma+\delta}} \leq \frac{2|s|^\delta}{n^{\gamma+\delta}} = \frac{2|s|^\delta}{n^{1+\varepsilon}}$$

Aplicando tanto esta como la desigualdad triangular al corolario 3.3.2, se logra la expresión principal:

$$|\zeta(s)| \leq \left|\frac{1}{s-1}\right| + \sum_{n=1}^{\infty} |f_n(s)| \leq \left|\frac{1}{s-1}\right| + 2|s|^\delta \sum_{n=1}^{\infty} \frac{1}{n^{1+\varepsilon}}$$

De que $\sigma \geq \gamma \in [0,1]$ y $|\tau| \geq 1$:

$$\left|\frac{1}{s-1}\right| = \frac{1}{\sqrt{(\sigma-1)^2+\tau^2}} \leq |\tau|^{-1} \leq |\tau|^\delta$$

Considere $\sigma \leq 1+\varepsilon$, entonces existe $c_\varepsilon^{(1)} > 1$:

$$|s|^\delta \leq |1+\varepsilon+i\tau|^\delta \leq c_\varepsilon^{(1)} |\tau|^\delta$$

Aplicando a nuestra expresión principal se logra finalmente:

$$|\zeta(s)| \leq \left(1 + 2c_\varepsilon^{(1)} \zeta(1+\varepsilon)\right) |\tau|^\delta \qquad \mathrm{Re}(s) \leq 1+\varepsilon$$

Como n^{-x} es decreciente con x variable real, es claro que $\zeta(x)$ es decreciente para x variable real mayor que 1. Luego, para $\sigma > 1+\varepsilon$, como $|\tau| \geq 1$ y $\delta > 0$ se cumple que:

$$|\zeta(s)| \leq \zeta(\sigma) \leq \zeta(1+\varepsilon) \leq 2c_\varepsilon^{(1)} |\tau|^\delta \zeta(1+\varepsilon) + |\tau|^\delta \qquad \mathrm{Re}(s) > 1+\varepsilon$$

Por lo que para $c_\varepsilon^{(2)} = 1 + 2c_\varepsilon^{(1)} \zeta(1+\varepsilon)$ se cumple la primera cota deseada.

Para la segunda desigualdad, considere $\sigma \geq 1$ y $\gamma = 0$. Sea $r > 0$ tal que $r \leq \min\{|\tau| - 1, \sigma\}$. De la fórmula integral de Cauchy en un círculo C de centro s y radio r, con la parametrización usual, se tiene:

$$\zeta'(s) = \frac{1}{2\pi i} \int_C \frac{\zeta(z)}{(z-s)^2} dz = \frac{1}{2\pi i} \int_0^{2\pi} \frac{\zeta(s+re^{i\theta})}{r^2 e^{2i\theta}} rie^{i\theta} d\theta = \frac{1}{2\pi r} \int_0^{2\pi} \zeta(s+re^{i\theta}) e^{-i\theta} d\theta$$

Veamos que dada la definición de r, para todo valor $z \in C$ se tiene que $\mathrm{Re}(z) \geq \sigma - r \geq \gamma = 0$ y $|\mathrm{Im}(z)| \geq |\tau| - r \geq 1$. Además la fórmula integral de Cauchy se puede ya que ζ era holomorfa en C y su interior, ya que su único polo en $s = 1$ está a una distancia de al menos $|\tau|$ de s y $r \leq |\tau| - 1 \leq |\tau|$. Esto permite aplicar la estimación de crecimiento de ζ anteriormente probada, obteniendo:

$$\begin{aligned}
|\zeta'(s)| &\leq \frac{1}{2\pi r}\int_0^{2\pi} |\zeta(s+re^{i\theta})|d\theta \leq \frac{c_\varepsilon^{(2)}}{2\pi r}\int_0^{2\pi} |\tau + rcos\theta|^{1+\varepsilon} d\theta \\
&\leq \frac{c_\varepsilon^{(2)}}{2\pi r}\int_0^{2\pi} (|\tau|+r)^{1+\varepsilon} d\theta = \frac{c_\varepsilon^{(2)}(|\tau|+r)^{1+\varepsilon}}{r} \\
&\leq \frac{2^{1+\varepsilon} c_\varepsilon^{(2)} |\tau|^{1+\varepsilon}}{r} \leq 2^{1+\varepsilon} c_\varepsilon^{(2)} |\tau|^{\varepsilon}
\end{aligned}$$

Por lo que $c_\varepsilon = 2^{1+\varepsilon} c_\varepsilon^{(2)}$ cumple la segunda desigualdad pedida en el teorema y como también $c_\varepsilon > c_\varepsilon^{(2)}$, también cumple con la primera, concluyendo la demostración.

El último teorema de esta sección crea una estimación para $1/\zeta$, que también será necesaria para la demostración del teorema de los números primos. Semejante al caso anterior, usando notación O grande de Landau, el teorema nos dice que $1/\zeta$ es $O(|\mathrm{Im}(s)|^\varepsilon)$ para todo s con $1 \leq \mathrm{Re}(s) \leq \sigma_0$ y $|\mathrm{Im}(s)|^\varepsilon \geq 1$, donde $\varepsilon > 0$ y $\sigma_0 > 1$ que influencian la constante.

La demostración también es un cálculo cuidadoso usando propiedades que ya conocemos. La idea será probar una cota inferior para la función dseta de la forma $k|\mathrm{Im(s)}|^{-\varepsilon}$. Para ello se parte del lema 3.2.4 en la forma:

$$|\zeta^3(\sigma)\zeta^4(s)\zeta(\sigma+2i\tau)| \geq 1$$

Recordar que $s = \sigma + i\tau$. Despejado $|\zeta(s)|$, como los otros ζ quedarán en el denominador, se pueden acotar superiormente (aumentar el denominador disminuye la fracción). Para $|\zeta(\sigma+2i\tau)|$ podemos usar el teorema 3.3.3 anterior, pero para $|\zeta(\sigma)|$ no porque su parte imaginaria es 0. Para este la cota viene del límite que da residuo al polo $s = 1$. Aplicando todo esto, nos queda una desigualdad de la forma:

$$|\zeta(s)| \geq C(\sigma-1)^{3/4}|\tau|^{-\mu/4}$$

Con $\mu > 0$ auxiliar y C alguna constante. Casi está lo pedido, el único problema es el término $(\sigma-1)^{3/4}$. Si fuera $\sigma - 1 \leq C'|\tau|^{-5\mu}$, no habría problema, se reemplaza el $\sigma - 1$ por su cota inferior y se reemplaza $\mu = \varepsilon/4$ para tenerlo deseado:

$$|\zeta(s)| \geq C''|\tau|^{-\varepsilon}$$

Pero si $\sigma - 1 < C'|\tau|^{-5\mu}$, habrá que usar otro camino. Consideremos la parte real menor que cumplía con el caso anterior $\sigma_1 = C'|\tau|^{-5\mu}$ y sea $s_1 = \sigma_1 + i\tau$. Entonces la desigualdad triangular nos marca el camino:

$$|\zeta(s)| \geq |\zeta(s_1)| - |\zeta(s) - \zeta(s_1)|$$

Al minuendo $|\zeta(s_1)|$ se le puede aplicar lo ya aprendido porque está en el caso anterior, mientras al sustraendo $|\zeta(s) - \zeta(s_1)|$ se le puede acotar superiormente usando el teorema del valor medio y la desigualdad del teorema 3.3.3 anterior para la derivada. Esto llevará con algunos cálculos a la misma desigualdad, aunque con una constante diferente:

$$|\zeta(s)| \geq C'''|\tau|^{-\varepsilon}$$

Tomando como constante final el inverso de la menor de las constantes conseguidas, se obtiene la demostración del teorema. Es necesario, eso sí, en cada paso tener un cuidado de siempre mantenerse dentro de las condiciones que exige el enunciado.

Antes de pasar a la demostración formal, es necesario hacer una acotación respecto al libro base [14]. Aunque con algunas ligeras desviaciones en las demostraciones y notaciones, los teoremas presentados mantienen los mismos enunciados hasta ahora que en el libro [14]. Pero en el caso del teorema 3.3.4 siguiente hay una ligera modificación, en cuanto se agrega un $\sigma_0 > 1$ arbitrario y se exige que la parte real σ sea menor o igual que σ_0, además la constante depende de este σ_0. Esto se debe a una ligera errata encontrada en [14]: Cuando se usa el límite $(\sigma-1)\zeta(\sigma) \to 1$ para acotar $\zeta(\sigma) \leq c/(\sigma-1)$, esta desigualdad solo se cumple localmente cerca de 1. Más específico, para cada intervalo $[1,\sigma_0]$ existe una constante c_{σ_0} que cumple esa desigualdad, pero no una sola constante para todo $(1,\infty)$ como usa [14]. De hecho, esa única constante no puede existir: De la definición vía serie se cumple siempre que $\zeta(\sigma) > 1$, mientras $c/(\sigma-1) \to 0$ para $\sigma \to \infty$, luego para un σ muy grande la desigualdad se debe quebrar. Esta errata es menor y con este pequeño cambio en el enunciado se arregla la demostración sin alterar las futuras aplicaciones.

Teorema 3.3.4 Sea $\sigma_0 > 1$. Sea $s = \sigma + i\tau$ con $\sigma, \tau \in \mathbb{R}$. Entonces para cada $\varepsilon > 0$, existe una constante k_{ε,σ_0} tal que: $\left|\frac{1}{\zeta(s)}\right| \leq k_{\varepsilon,\sigma_0}|\tau|^{\varepsilon}$ para $\sigma \in [1,\sigma_0]$ al igual que $|\tau| \geq 1$

Demostración. Sea $\varepsilon, \mu > 0$. Sea $\sigma_0 > 1$, $\sigma \in [1,\sigma_0]$ y $\tau \in \mathbb{R}$ tal que $|\tau| > 1$.

Del lema 3.2.4, se sabe que:

$$|\zeta^3(\sigma)\zeta^4(s)\zeta(\sigma+2i\tau)| \geq 1$$

Lo que implica:

$$|\zeta(s)| \geq \frac{1}{\sqrt[4]{|\zeta^3(\sigma)\zeta(\sigma+2i\tau)|}}$$

Aplicando la primera desigualdad del teorema 3.3.3 anterior para $\gamma = 1$, se tiene:

$$|\zeta(s)| \geq \frac{|t|^{-\mu/4}}{\sqrt[4]{c_\mu \zeta^3(\sigma)}}$$

Como $s = 1$ es un polo simple de residuo 1, existe c_{σ_0} tal que:

$$|\zeta(\sigma)| \leq \frac{c_{\sigma_0}}{\sigma-1} \qquad \forall \sigma \leq \sigma_0$$

Aplicando a lo que se tenía anteriormente, se deduce que:

$$|\zeta(s)| \geq \frac{|t|^{-\mu/4}(\sigma-1)^{3/4}}{\sqrt[4]{c_\mu c_{\sigma_0}^3}} = c_1(\sigma-1)^{3/4}|\tau|^{-\mu/4}$$

donde $c_1 \leq c_\mu^{-1/4} c_{\sigma_0}^{-3/4}$.

Escogemos c_1 lo suficientemente pequeño tal que $c_2 = (c_1/(2c_\mu))^4$ sea menor o igual que $\sigma_0 - 1$, siendo c_μ la constante usada en el teorema 3.3.3 anterior. Tenemos entonces, dos casos: **Si** $\sigma - 1 \geq c_2|\tau|^{-5\mu}$, entonces aplicando en la desigualdad anteriormente obtenida se concluye que:

$$|\zeta(s)| \geq c_1 c_2^{-3/4}|\tau|^{-4\mu}$$

Si $\sigma - 1 < c_2|\tau|^{-5\mu}$, entonces sea $\sigma_1 - 1 = c_2|\tau|^{-5\mu}$. Como $|\tau|^{-5\mu} \leq 1$, se tiene que $\sigma_1 \leq \sigma_0$. Sea $s_1 = \sigma_1 + i\tau$, Por desigualdad triangular, se sabe que:

$$|\zeta(s)| \geq |\zeta(s_1)| - |\zeta(s_1) - \zeta(s)|$$

El minuendo puede acotarse inferiormente aprovechando que $\sigma_1 \leq \sigma_0$, luego se puede utilizar la desigualdad final del caso anterior:

$$|\zeta(s)| \geq c_1(\sigma_1 - 1)^{3/4}|\tau|^{-\mu/4} - |\zeta(s_1) - \zeta(s)|$$

El sustraendo puede acotarse superiormente a través del teorema del valor medio, ya que existe algún $\sigma_2 \in [\sigma, \sigma_1]$ tal que:

$$|\zeta(s_1) - \zeta(s)| = |\zeta'(\sigma_2 + i\tau)||\sigma_1 - \sigma|$$

Como $\sigma_2 > 1$ y $|\tau| \geq 1$, aplicando la segunda desigualdad del teorema 3.3.3 anterior se tiene:

$$|\zeta(s_1) - \zeta(s)| \leq c_\mu|\tau|^\mu|\sigma_1 - \sigma| \leq c_\mu|\tau|^\mu(\sigma_1 - 1)$$

Aplicando la cota conseguida, se tiene:

$$|\zeta(s)| \geq c_1(\sigma_1 - 1)^{3/4}|\tau|^{-\mu/4} - c_\mu|\tau|^\mu(\sigma_1 - 1)$$

Al reemplazar $\sigma_1 - 1$ con $(c_1/(2c_\mu))^4|\tau|^{-5\mu}$ se tiene:

$$|\zeta(s)| \geq \left(\frac{c_1^4}{8c_\mu^3}\right)|\tau|^{-4\mu} - \left(\frac{c_1^4}{16c_\mu^3}\right)|\tau|^{-4\mu} = c_2|\tau|^{-4\mu}$$

con $c_2 = c_1^4/16c_\mu^3$. Tomando $\mu = \varepsilon/4$ y escogiendo k_{ε,σ_0} como el inverso multiplicativo del menor de c_2 y $c_1c_2^{3/4}$, se concluye que:

$$|\zeta(s)| \geq k_{\varepsilon,\sigma_0}^{-1}|\tau|^{-\varepsilon} \quad \Rightarrow \quad 1/|\zeta(s)| \leq k_{\varepsilon,\sigma_0}|\tau|^\varepsilon$$

3.4 Ceros de la función dseta

Para concluir este capítulo dedicado a la función dseta, es necesario discutir sobre sus ceros. Como se ha visto en el capítulo de historia y se profundizará más en el próximo capítulo, los ceros de la función dseta tienen una conexión especial con la manera en que se distribuyen los números primos, por lo que han sido tópico de interés de estudio para muchos matemáticos. La función dseta tiene una cantidad infinita de ceros, que son clasificados en dos tipos: Los **ceros triviales** que son los pares negativos y los **ceros no triviales** que están en la franja $\mathrm{Re}(s) = 0$ y $\mathrm{Re}(s) = 1$, que es llamada la franja crítica. La denominación trivial se refiere a que, mientras los ceros triviales son bien conocidos, los ceros no triviales les cubre un manto de misterio y desde los tiempos de Riemann hasta el día de hoy son objeto de estudio para revelar cómo se distribuyen en la franja crítica. Pero es en este estudio donde se encuentra la deseada recompensa: Los ceros no triviales esconden en parte el comportamiento caótico, en sentido coloquial, de la aparición de los números primos en la recta numérica.

¿Y cómo se sabe que la función dseta tiene ceros solo en la franja crítica y en los pares negativos? Aquello es materia del siguiente teorema 3.4.1. Lo primero es probar que la función dseta no tiene ceros en H_1. Esto se logra directamente al usar la Identidad de Euler y el teorema 5.1.5 del apéndice.

Lo segundo es que la simetría de la función ξ, vista en el teorema 3.1.3, implica una relación entre $\zeta(s)$ y $\zeta(1-s)$:

$$\zeta(s) = \frac{\pi^{s-0,5}\Gamma(\frac{1-s}{2})\zeta(1-s)}{\Gamma(s/2)}$$

Esto permite usar nuestro conocimiento sobre la ausencia de ceros de la función dseta en H_1 para estudiar la existencia de ceros en el semiplano $\mathrm{Re}(s) < 0$. Nunca $\pi^{s-0,5}$ o la función gamma se hacen cero, y para s en ese semiplano, $\zeta(1-s)$ no se hace cero. Luego los únicos ceros que la función dseta puede tener en $\mathrm{Re}(s) < 0$ son por $1/\Gamma$, lo que dan los pares negativos si se considera que está evaluado a $s/2$. Lamentablemente este método no nos otorga información sobre la franja crítica, por lo que todos los demás ceros deben encontrarse en esa región inexplorada por el teorema.

Teorema 3.4.1 Todos los ceros de la función ζ son de la forma:

- **Ceros triviales:** Son pares negativos, como $-2, -4, -6, \ldots$
- **Ceros no triviales:** Pertenecen a la franja crítica $FC := \{s \in \mathbb{C} \mid \mathrm{Re}(s) \in [0,1]\}$

Demostración. Para el caso $\sigma = \mathrm{Re}(s) > 1$, podemos expresar la función dseta según la identidad de Euler:

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

De un pequeño cálculo se tiene que:

$$\begin{aligned}\sum_{p \text{ primo}} \left|\frac{1}{1-p^{-s}} - 1\right| &= \sum_{p \text{ primo}} \left|\frac{p^{-s}}{1-p^{-s}}\right| = \sum_{p \text{ primo}} |p^s - 1|^{-1} \\ &\leq \sum_{p \text{ primo}} (|p^s| - 1)^{-1} = \sum_{p \text{ primo}} (p^\sigma - 1)^{-1} \\ &\leq \sum_{p \text{ primo}} 2p^{-\sigma} \leq 2\zeta(\sigma) < \infty\end{aligned}$$

Por el teorema 5.1.5 del apéndice, el producto infinito solo se puede anular si un término se anula, lo que claramente no puede suceder. Por ende ζ no tiene raíces con parte real mayor a 1.

Para el caso $\mathrm{Re}(s) < 0$, utilizamos la construcción de la función ζ en el corolario 3.1.4, y aplicando la simetría del teorema 3.1.3 y luego la definición 3.1.2 de la función ξ en H_1:

$$\zeta(s) = \frac{\pi^{s/2}\xi(s)}{\Gamma(s/2)} = \frac{\pi^{s/2}\xi(1-s)}{\Gamma(s/2)} = \frac{\pi^{s-0,5}\Gamma(\frac{1-s}{2})\zeta(1-s)}{\Gamma(s/2)}$$

Conocido es que $\pi^{s-0,5}$ no tiene raíces. Por el corolario 2.3.4, $\Gamma(\frac{1-s}{2})$ no tiene raíces. $1/\Gamma(s/2)$ se anula en $s \in \{-2, -4, ...\}$. Finalmente de que en este caso $\mathrm{Re}(1-s) > 1$, $\zeta(1-s)$ no tiene ceros. Luego ceros en esta región son los ceros triviales $\{-2, -4, ...\}$, mientras cualquier otro posible cero deberá estar en la franja crítica $\mathrm{Re}(s) \in [0,1]$.

El teorema 3.4.1 no brinda información sobre la existencia de ceros no triviales en la recta $\mathrm{Re}(s) = 1$. Como se ahonda en el primer capítulo, los geniales estudios de Riemann sobre la función dseta culminaron en que el Teorema de los Números Primos podía ser probado como una consecuencia de la no existencia de ceros en dicho eje. Y justamente serían Hadamard y De la Vallée Poussin quienes probarían independientemente este hecho en 1896. Desde entonces otras demostraciones han venido, pero la mayoría de las analíticas se basan directa o indirectamente en que la función dseta no tiene ceros en $\mathrm{Re}(s) = 1$. La demostración de esta obra no será diferente.

Aunque esta demostración se inspira en ideas que se remontan a Hadamard y De la Vallée Poussin, el tiempo ha simplificado la prueba al punto que todo se trata de un simple cálculo de límite. Es una reducción al absurdo: Supongamos que existe un cero de la función dseta de la forma $1+yi$, entonces se puede calcular que el siguiente límite se hace 0

$$\lim_{x\to 1}\left|\zeta^3(x)\zeta^4(x+iy)\zeta(x+2iy)\right| = 0$$

El cálculo de límite no tiene dificultad: Hay que reconstruir el límite del cálculo de residuo para $\zeta(x)$ y el límite que define la derivada para $\zeta(x+iy)$, todo esto multiplicando un "1 conveniente". La consecuencia es clara: Este resultado contradice el lema 3.2.4 que obliga a la expresión dentro del límite a ser mayor que 1. Y eso es suficiente. Si sorprende la sencillez del paso clave, cabe recordar lo señalado por Grothendieck sobre esta manera de hacer matemáticas: Una vez sumergida la nuez por suficiente tiempo, basta un ligero apretón para romper su cáscara. Este es el apretón.

Teorema 3.4.2 La función ζ no tiene ceros en $\mathrm{Re}(s) = 1$.

Demostración. Suponga lo contrario, es decir, que existe $y \in \mathbb{R}$ tal que $\zeta(1+iy) = 0$. Evidentemente $y \neq 0$ porque en $s = 1$ hay un polo. Para $x \in \mathbb{R}$, estudiaremos que sucede con $\left|\zeta^3(x)\zeta^4(x+iy)\zeta(x+2iy)\right|$ cuando $x \to 1^+$:

$$\begin{aligned}
&\lim_{x\to 1^+}\left|\zeta^3(x)\zeta^4(x+iy)\zeta(x+2iy)\right| \\
&= \lim_{x\to 1^+}\left|\zeta^3(x)\zeta^4(x+iy)\zeta(x+2iy)\right|\left|\frac{(x-1)^4}{(x-1)^4}\right| \\
&= \lim_{x\to 1^+}\left|(x-1)^3\zeta^3(x)\right| \lim_{x\to 1^+}\left|\frac{\zeta^4(x+iy)}{(x-1)^4}\right| \lim_{x\to 1^+}\left|\zeta(x+2iy)\right| \lim_{x\to 1^+}\left|(x-1)\right| = 0
\end{aligned}$$

El primer límite converge a una constante ya que es el cubo del residuo del polo simple de ζ en $s = 1$. El segundo límite también converge a una constante porque siendo $\zeta(1+iy) = 0$ por hipótesis, este límite resulta ser la derivada de ζ en $1+iy$ elevado a la cuarta potencia y como ζ es holomorfa este valor existe. El tercer límite también converge a un valor finito porque ζ es continua en $1+2iy$. Finalmente el último límite resulta ser 0, provocando que el límite resultante sea 0.

Esto se contradice con el teorema 3.2.4, ya que este implica que:

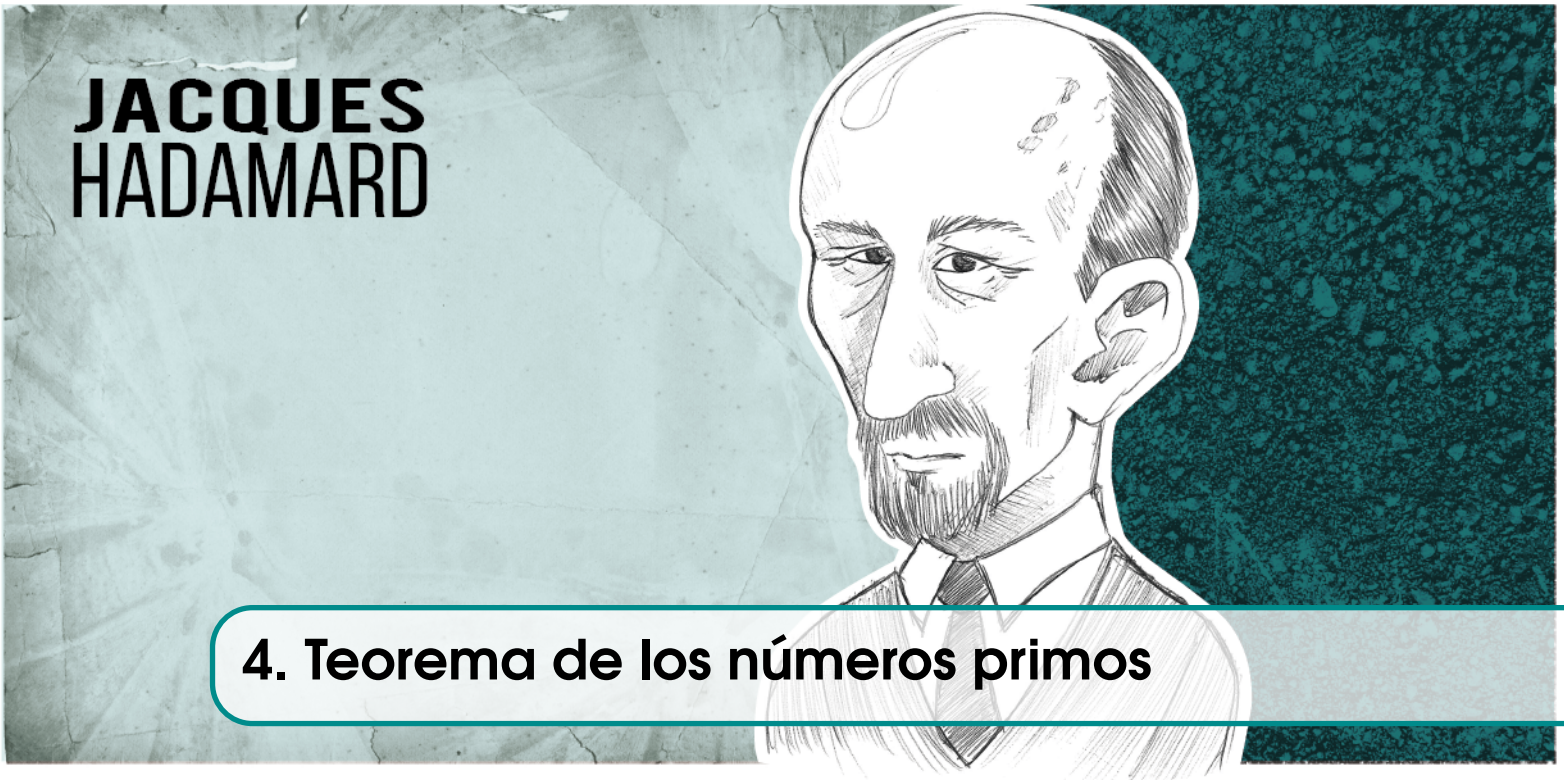
$$\left|\zeta^3(x)\zeta^4(x+iy)\zeta(x+2iy)\right| \geq 1 \quad \forall x > 1$$

Impidiendo que ese término pueda converger a 0 cuando $x \to 1^+$.

Como nota de cierre, si bien nuestro estudio de la función dseta se ha limitado a lo estrictamente necesario para demostrar el Teorema de los Números Primos, el universo matemático de la función dseta es mucho más extenso. Los ceros no triviales, solo sugeridos como una posibilidad en los dos teoremas de esta sección, en la realidad son infinitos y su distribución esconde formas de perfeccionar y estimar el error del teorema de los números primos. Varias conjeturas aún aguardan en este campo, una de ellas posiblemente la más famosa de las matemáticas actuales. Benhard Riemann, estudiando los ceros no triviales, observó que todos parecen cumplir la siguiente propiedad:

Problema 3.2 — Hipótesis de Riemann. Todos los ceros no triviales de la función ζ están en la línea $\mathrm{Re}(s) = \frac{1}{2}$.

Lamentablemente, este margen es demasiado pequeño para contener la demostración.



4. Teorema de los números primos

Hay frase memorable atribuida a Carl Gauss: “Las matemáticas son la reina de las ciencias y la teoría de números es la reina de las matemáticas”. Siendo así, seguramente el teorema de los números primos es una de las joyas más preciadas de la corona. Su enunciado otorgó un anhelado orden a la desconcertante distribución de los números primos y su demostración, fruto de casi 100 años de la labor cooperativa de la comunidad matemática, trajo una conexión también desconcertante entre dos ramas de las matemáticas tan alejadas en apariencia como son la aritmética, el campo que estudia lo discreto, y el análisis, el área que estudia lo continuo.

Y finalmente, tras haber pasado un largo camino entre diversas lecciones sobre extensión analítica y propiedades de la función dseta, estamos en condiciones de abordar el teorema de los números primos y demostrarlo. El teorema contesta la siguiente pregunta: De todos los números naturales hasta x, ¿qué proporción son números primos? La respuesta del teorema es: Si x es muy grande, aproximadamente $1/\log x$. Una respuesta engañosamente simple, pero que esconde bastantes secretos sobre la naturaleza de los números. Primero, impresiona en sí mismo que la aparición pseudoaleatoria de los números primos tenga una densidad tan sencilla. Segundo, nos otorga una primera forma simple para estudiar la secuencia de los números primos en grandes escalas, donde las antiguas cribas nos son inútiles incluso con los poderes computacionales actuales. Tercero, no todo es sobre el enunciado, sino también sobre la demostración. Ya se ha dedicado dos capítulos al análisis complejo, pero aún no se sospecha cómo se usarán tales técnicas para demostrar el teorema y son estas conexiones las que también nos dicen bastante sobre la naturaleza de los números y las matemáticas. Más aún, ¿cómo se demuestra una proposición de estas características donde lo que se plantea es un límite sobre una función que conocemos poco y nada?

Especialmente esta interrogante se responderá en este capítulo final. La primera sección formalizará el enunciado del teorema y su interpretación, la segunda veremos cómo reducir el problema siguiendo las ideas de Tchebyshev, en la tercera sección conectaremos estas reducciones con la función dseta y finalmente en la última sección concluiremos la demostración usando las propiedades aprendidas en los capítulos anteriores, dando fin al largo viaje emprendido.

4.1 Describiendo el problema

El teorema de los números primos versa sobre la función contadora de números primos, comúnmente denotada por la letra π minúscula, cuyo nombre encaja bastante bien su definición:

Definición 4.1.1 — Función contadora de números primos. La función $\pi : \mathbb{R}^+ \to \mathbb{R}$ está definida como

$$\pi(x) = \text{número de primos menores o iguales que } x$$

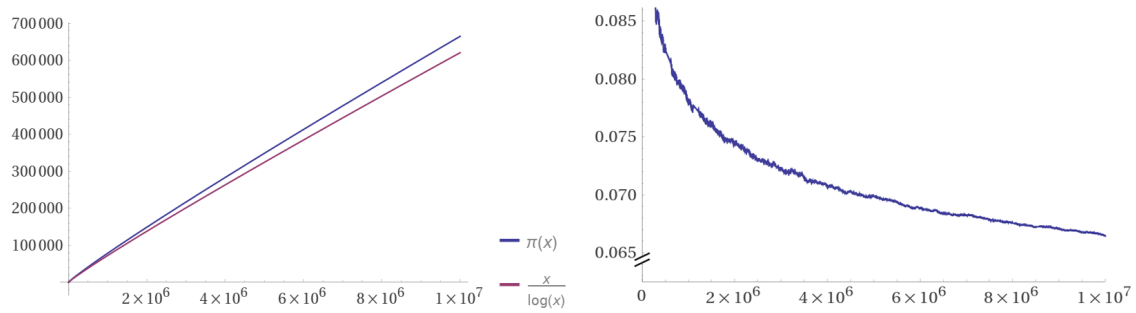
Con el protagonista claro, el enunciado formal del teorema de los números primos, que será el problema objetivo de todo este capítulo, es el siguiente:

Problema 4.1 — Teorema de los números primos. Demostrar que $\pi(x) \sim \frac{x}{\log x}$.

Donde el signo $\sim$ significa igualdad asintótica. Recordemos que dos funciones f y g se dicen asintóticamente iguales si $f(x)/g(x) \to 1$ cuando $x \to \infty$. Luego el enunciado del teorema de los números primos se puede expresar a probar el siguiente límite:

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

Es importante entender bien esta expresión. Este límite no significa que la diferencia entre $\pi(x)$ y $x/\log(x)$ se vaya a 0 a medida que x crece, sino que esta diferencia se va volviendo despreciable si se compara con el tamaño de x y de las funciones. En términos de errores, este límite no informa que el error absoluto desaparezca, sino sobre el error relativo se va haciendo cero. Esto se puede apreciar en los siguientes dos gráficos: Aunque a medida que x aumenta, $x/\log(x)$ y $\pi(x)$ se distancian en mayores cantidades absolutas, el error relativo entre ellos va disminuyendo cada vez más. Por lo que lo que el teorema de los números primos hace, no es ofrecernos una predicción sobre el valor de la función contadora de números primos que se vaya haciendo más certera a valores más grandes, sino brinda un entendimiento de cómo esta función $\pi(x)$ ser comporta en las grandes escalas.



(a) $\pi(x)$ vs $x/\log(x)$

(b) Error relativo de $x/\log(x)$ al aproximar $\pi(x)$

Figura 4.1: El primer gráfico muestra una comparativa entre la función contadora de números primos $\pi(x)$ y $x/\log(x)$. El segundo gráfico muestra el error relativo de $x/\log(x)$ al estimar $\pi(x)$. Se logra evidenciar que mientras el error absoluto puede aumentar, el error relativo disminuye cada vez más. Ambos gráficos son sobre el intervalo 2 a 10^7. **Fuente:** Wolfram Alpha.

Existen otras formas equivalentes de enunciar el teorema. Una es $\pi(x)/x \sim 1/\log(x)$. Matemáticamente, es lo mismo que lo enunciado anteriormente, pero su interpretación es ligeramente diferente: $\pi(x)/x$ es la proporción de cuantos números menores que x son primos y esta se comporta como $1/\log(x)$. Esta interpretación fue la original planteada por Gauss y Legendre.

Otra forma usual en la literatura es $\pi(x) \sim \text{Li}(x)$ donde $\text{Li}(x)$ es la integral logarítmica desplazada:

$$\text{Li}(x) = \int_2^x \frac{du}{\log(u)}$$

Esta fue una de las propuestas originales de Gauss. Aunque en los primeros valores de x tiene un peor desempeño que $x/\log(x)$, a grandes escalas termina siendo un mejor aproximador de $\pi(x)$. De todas formas, esta es una formulación equivalente del teorema de los números primos dado en el problema 4.1 porque $\text{Li}(x) \sim x/\log(x)$ y esto provoca que ambos límites sean iguales. En el teorema 4.1.1 se demuestra este hecho, que sale de integrar por partes a $\text{Li}(x)$ (usando como partes a $u = 1/\log x$ y $v = x$ según la notación usual). Lo resultante va a ser de la forma $\text{Li(x)} = x/(\log x) + o\,(x/\log x)$, siguiendo la notación o pequeña de Landau[1]. La prueba que los términos adicionales son $o\,(x/\log x)$ se hace por teorema del acotamiento.

Teorema 4.1.1 $\text{Li}(x) \sim \frac{x}{\log x}$. Luego $\pi(x) \sim \text{Li}(x)$ si y solo si $\pi(x) \sim \frac{x}{\log x}$.

Demostración. Vea que:

$$\frac{d}{dx}\left(\frac{1}{\log x}\right) = \frac{-1}{x(\log x)^2}$$

Luego si se aplica integración por partes:

$$\text{Li}(x) = \int_2^x \frac{du}{\log u} = \frac{u}{\log u}\bigg|_2^x + \int_2^x \frac{du}{(\log u)^2} = \frac{x}{\log x} - \frac{2}{\log 2} + \int_2^x \frac{du}{(\log u)^2}$$

Adicionalmente, como $(\log x)^-2$ es decreciente, se puede acotar esta última integral como sigue:

$$0 \leq \int_2^x \frac{du}{(\log u)^2} = \int_2^{\sqrt{x}} \frac{du}{(\log u)^2} + \int_{\sqrt{x}}^x \frac{du}{(\log u)^2} \leq \frac{\sqrt{x}}{(\log 2)^2} + \frac{x}{(\log \sqrt{x})^2}$$

Si se divide por $x/\log x$ y se aplica el límite al infinito se tiene:

$$\begin{aligned} 0 \leq \lim_{x\to\infty} \frac{\log x}{x} \int_2^x \frac{du}{(\log u)^2} &\leq \lim_{x\to\infty} \frac{\sqrt{x}\log x}{x(\log 2)^2} + \lim_{x\to\infty} \frac{x\log x}{x(\log\sqrt{x})^2} \\ &\leq \frac{1}{(\log 2)^2} \lim_{x\to\infty} \frac{\log x}{\sqrt{x}} + \lim_{x\to\infty} \frac{2}{\log\sqrt{x}} \to 0 \end{aligned}$$

Por lo que por teorema del acotamiento se tiene que dicho límite tiende a 0. Aplicando lo antes obtenido, se deduce:

$$\lim_{x\to\infty} \frac{\log(x)\text{Li}(x)}{x} = \lim_{x\to\infty} \frac{x\log x}{x\log x} - \lim_{x\to\infty} \frac{2\log x}{x\log 2} + \lim_{x\to\infty} \frac{\log x}{x} \int_2^x \frac{du}{(\log u)^2} = 1$$

Por lo que $\text{Li(x)} \sim \frac{x}{\log x}$. Por consecuencia se tiene que:

$$\lim_{x\to\infty} \frac{\pi(x)}{\text{Li}(x)} = \left(\lim_{x\to\infty} \frac{x}{\text{Li}(x)\log x}\right)\left(\lim_{x\to\infty} \frac{\pi(x)\log x}{x}\right) = \lim_{x\to\infty} \frac{\pi(x)\log x}{x}$$

Como los límites son iguales, se tiene que $\pi(x) \sim \text{Li}(x)$ si y solo si $\pi(x) \sim \frac{x}{\log x}$.

En lo que resta del capítulo, nos enfocaremos en demostrar el teorema de los números primos con el enunciado dado en el problema 4.1: $\pi(x) \sim x/\log x$.

[1] Una función f es $o(g)$ si y solo si $f(x)/g(x) \to 0$ para $x \to \infty$. Esta notación también es de Edmund Landau.

4.2 Reducciones de Tchebyshev

La principal complicación que enfrenta demostrar el teorema de los números primos es $\pi(x)$. La función contadora ni siquiera está expresada en términos de alguna operación que involucre a los números primos, sino que se limita a contarlos. Por decirlo de una manera informal, la función contadora no es ni siquiera sobre los primos en sí, sino exclusivamente sobre cómo se distribuyen. Y eso es un problema, porque poco se conoce sobre la distribución de los números primos, pues justamente eso es lo que queremos descubrir con este teorema.

Para abordar esta complejidad, el matemático ruso Pafnuti Tchebyshev dio con una idea clave: Hay que buscar otro enunciado equivalente, pero que se base en una función más trabajable. Tchebyshev consideró dos funciones que hoy llevan su nombre acompañado por una letra griega, pero nosotros nos limitaremos a la segunda de esas funciones: La función ψ de Tchebyshev.

Definición 4.2.1 La función ψ de Tchebyshev es $\psi : \mathbb{R}^+ \to \mathbb{R}$ dada por:

$$\psi(x) = \sum_{p^m \leq x} \log p$$

Esta función es bastante más trabajable, porque goza tanto de las propiedades del logaritmo natural, de las series como de las propiedades algebraicas de los números primos. De hecho, un uso meticuloso y exhaustivo de estas propiedades otorga una vía alternativa para la demostración descubierta por Atle Selberg a mediados del siglo XX. Pero en lo que respecta a nuestra demostración, requeriremos explorar dos formas equivalentes de expresar la función ψ en la proposición 4.2.1. La demostración es tan intuitiva como formal y no requiere de introducción:

Proposición 4.2.1 Otras formas equivalentes para la función ψ de Tchebyshev son:

- $\psi(x) = \sum_{p \leq x} \left[\frac{\log x}{\log p} \right] \log p$
- $\psi(x) = \sum_{n=1}^{[x]} \Lambda(n)$

donde $\Lambda : \mathbb{N} \to \mathbb{R}$ es la función de von Mangoldt que asume el valor $\log p$ si n es una potencia de un p primo y 0 en cualquier otro caso.

Demostración. Primero, note que no hay naturales entre x y $[x]$, luego el cambio en el término de la sumatoria no hace diferencia.

Para la primera equivalencia, sea M_p el mayor natural tal que $p^{M_p} \leq x$. En la suma que define a ψ, el término $\log p$ aparecerá M_p veces para cada p primo menor o igual que $[x]$. M_p puede expresarse como la parte entera de $\log_p x = \log x / \log p$, por lo que se deriva la identidad deseada.

Para la segunda equivalencia, es evidente desde que $\Lambda(n)$ coincide para cada término de la suma que define ψ y se anula en todos los demás.

Aunque más cómoda, de nada serviría la función ψ sin un resultado sobre ella que sea equivalente al teorema de los números primos. Tchebyshev probó la existencia de un elegante resultado equivalente: $\psi(x) \sim x$. Aunque se puede probar su recíproco, en el teorema 4.2.2 nos limitaremos a probar la implicación en el sentido que nos sirve a nuestro problema objetivo: $\psi(x) \sim x$ provoca que $\pi(x) \sim \frac{x}{\log x}$. ¿Y cómo se demuestra? Pues, asumiendo que $\psi(x) \sim x$ es cierto, buscaremos generar el siguiente acotamiento:

$$1 = \liminf_{x\to\infty} \frac{\psi(x)}{x} \leq \liminf_{x\to\infty} \frac{\pi(x)\log x}{x} \leq \lim_{x\to\infty} \frac{\pi(x)\log x}{x} \leq \limsup_{x\to\infty} \frac{\pi(x)\log x}{x} \leq \limsup_{x\to\infty} \frac{\psi(x)}{x} = 1$$

Con las igualdades dadas por la hipótesis y las desigualdades centrales dadas por el análisis, es suficiente probar las desigualdades externas. Ambas salen de acotar apropiadamente la primera expresión alternativa de $\psi(x)$ en la proposición 4.2.1, superiormente por $\log(x)\pi(x)$ e inferiormente por $\varepsilon\log(x)\pi(x)$, para cualquier $\varepsilon \in (0,1)$. De ahí se reconstruyen los límites y ya está. Sin mayor preámbulo, pasemos a la demostración formal:

Teorema 4.2.2 Si $\psi(x) \sim x$, entonces $\pi(x) \sim \frac{x}{\log x}$.

Demostración. Suponga que $\psi \sim x$. La demostración del límite procede por criterio de acotamiento:

Acotamiento inferior: De la proposición 4.2.1 anterior se tiene que:

$$\psi(x) = \sum_{p\leq x}\left[\frac{\log x}{\log p}\right]\log p \leq \sum_{p\leq x}\frac{\log x}{\log p}\log p = \pi(x)\log(x)$$

Dividiendo por x se tiene que:

$$\frac{\psi(x)}{x} \leq \frac{\pi(x)}{x/\log(x)}$$

Tomando el límite inferior se tiene:

$$1 = \liminf_{x\to\infty}\frac{\psi(x)}{x} \leq \liminf_{x\to\infty}\frac{\pi(x)}{x/\log(x)}$$

Acotamiento superior: Sea $\varepsilon \in (0,1)$. De la proposición 4.2.1, podemos acotar inferiormente como sigue

$$\psi(x) = \sum_{p\leq x}\left[\frac{\log x}{\log p}\right]\log p \geq \sum_{p\leq x}\log p \geq \sum_{x^\varepsilon\leq p\leq x}\log p \geq \sum_{x^\varepsilon\leq p\leq x}\log x^\varepsilon = \varepsilon\log x\pi(x)$$

Dividiendo por εx se tiene:

$$\frac{\psi(x)}{\varepsilon x} \geq \frac{\pi(x)}{x/\log x}$$

Tomando el límite superior se tiene que:

$$\varepsilon^{-1} = \limsup_{x\to\infty}\frac{\psi(x)}{\varepsilon x} \geq \limsup_{x\to\infty}\frac{\pi(x)}{x/\log x}$$

Luego tomando $\varepsilon \to 1$ se logra lo deseado.

Finalmente, como el límite superior y el límite inferior coinciden en 1, el límite al infinito de $x\pi(x)/\log(x)$ existe y es 1.

Uno de los problemas tanto de $\pi(x)$ como de $\psi(x)$ es que son funciones escalera, es decir, se mantienen constantes hasta que crecen a saltos discontinuos. Esto resulta molesto para utilizar herramientas de análisis como queremos hacer, por ende, un nuevo y último cambio a un enunciado equivalente se hace necesario. Para este efecto, la nueva función a considerar es la función suavizante de Tchebyshev, denotada por ψ_1, y que resulta ser la integral de la función ψ:

Definición 4.2.2 La función suavizante de Tchebyshev es $\psi_1 : \mathbb{R}^+ \to \mathbb{R}$ definida por la regla:

$$\psi_1(x) = \int_1^x \psi(t)dt$$

¿Y cuál sería el nuevo enunciado objetivo que implique el teorema de los números primos? Pues si el anterior era $\psi \sim x$ y ψ_1 es su integral, las sospechas recaen a la integral de x: $x^2/2$. Justamente este es el caso: $\psi_1 \sim \frac{x^2}{2}$. La demostración sigue la misma estrategia que la prueba anterior. Asumiendo que $\psi_1 \sim \frac{x^2}{2}$, queremos probar el siguiente acotamiento:

$$1 = \liminf_{x\to\infty} \frac{2\psi_1(x)}{x^2} \leq \liminf_{x\to\infty} \frac{\psi(x)}{x} \leq \lim_{x\to\infty} \frac{\psi(x)}{x} \leq \limsup_{x\to\infty} \frac{\psi(x)}{x} \leq \limsup_{x\to\infty} \frac{2\psi_1(x)}{x^2} = 1$$

Al igual que en el teorema previo, basta con probar las desigualdades externas y estas vienen de un acotamiento apropiado. Esta vez el acotamiento viene dado por el teorema del valor medio para $[\alpha x, x]$ y $[x, \beta x]$ para cualquier $0 \leq \alpha \leq 1 \leq \beta$. Como ψ es creciente, las medias en esos intervalos harán cotas inferiores y superiores basadas en la integral de ψ, es decir, ψ_1. Luego basta con reconstruir los límites y hacer $\alpha \to 1 \leftarrow \beta$ para lograr la demostración.

Teorema 4.2.3 Si $\psi_1 \sim \frac{x^2}{2}$, entonces $\pi(x) \sim \frac{x}{\log x}$.

Demostración. Suponga que $\psi_1 \sim \frac{x^2}{2}$. Observemos que ψ es una función creciente, ya que valores de x más elevados solo están incluyendo sumandos positivos a su valor. Luego tenemos que, para $0 < a < 1 < b$ cualesquiera, se cumple que:

$$\frac{1}{x-\alpha x}\int_{\alpha x}^{x} \psi(x)dx \leq \psi(x) \leq \frac{1}{\beta x - x}\int_{x}^{\beta x} \psi(x)dx$$

Por teorema del valor medio, el término del lado derecho es igual a ψ evaluado en algún valor entre $[\alpha x, x]$ y por ende menor a $\psi(x)$. El razonamiento recíproco explica la desigualdad del lado derecho. Escrita estas desigualdades en términos de psi_1 quedan:

$$\frac{1}{x-\alpha x}(\psi_1(x) - \psi_1(\alpha x)) \leq \psi(x) \leq \frac{1}{\beta x - x}(\psi_1(\beta x) - \psi_1(x))$$

Dividiendo por x y con una multiplicación de un 1 conveniente queda:

$$\frac{1}{1-\alpha}\left(\frac{\psi_1(x)}{x^2} - \frac{\psi_1(\alpha x)}{\alpha^2 x^2}\alpha^2\right) \leq \frac{\psi(x)}{x} \leq \frac{1}{\beta - 1}\left(\frac{\psi_1(\beta x)}{\beta^2 x^2}\beta^2 - \frac{\psi_1(x)}{x^2}\right)$$

Del supuesto, $\lim_{x\to\infty} \psi_1(cx)/c^2x^2 = 1/2$ para cualquier $c > 0$. Luego aplicando límites inferior en el lado izquierdo y límites superior en el lado derecho queda:

$$\liminf_{x\to\infty} \frac{\psi(x)}{x} \geq \liminf_{x\to\infty} \frac{1}{1-\alpha}\left(\frac{\psi_1(x)}{x^2} - \frac{\psi_1(\alpha x)}{\alpha^2 x^2}\alpha^2\right) = \frac{1}{1-\alpha}\left(\frac{1}{2} - \frac{\alpha^2}{2}\right) = \frac{1+\alpha}{2}$$

$$\limsup_{x\to\infty} \frac{\psi(x)}{x} \leq \limsup_{x\to\infty} \frac{1}{\beta - 1}\left(\frac{\psi_1(\beta x)}{\beta^2 x^2}\beta^2 - \frac{\psi_1(x)}{x^2}\right) = \frac{1}{\beta - 1}\left(\frac{\beta^2}{2} - \frac{1}{2}\right) = \frac{\beta+1}{2}$$

Tomando $\alpha \to 1^-$ y $\beta \to 1^+$, se tiene por coincidencia de los límites superiores e inferiores que el límite al infinito de $\psi(x)/x$ existe y es 1. Por el teorema 4.2.2, se concluye que $\pi(x) \sim x/\log x$

4.3 Relación con la función dseta

Ya habiendo definido el enunciado que vamos a querer demostrar, $\psi_1 \sim x^2/2$, lo siguiente es conectar esta función ψ_1 con la función dseta. De aquello tratará el teorema 4.3.3, sin embargo para poder probarlo necesitaremos de demostrar dos proposiciones previas. La primera es una igualdad que relaciona la función dseta, su derivada y la función de Van Mangoldt que apareció en la proposición 4.2.1. La demostración no tiene dificultad: Es la derivada de $\log\zeta(s)$ usando la primera igualdad del lema 3.2.3. Esta derivada se puede calcular derivando bajo la doble sumatoria porque estas convergen uniformemente, como se señaló en la demostración de dicho lema. Pasemos a la demostración:

Proposición 4.3.1 Para $s \in H_1$, se tiene que:

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

Demostración. Del lema 3.2.3, se cumple que:

$$\log\zeta(s) = \sum_{k\in\mathbb{N},\ p \text{ primo}} \frac{p^{-ks}}{k}$$

En la demostración del lema 3.2.3, dicha doble serie converge uniformemente, luego se puede diferenciar bajo la serie, donde con un pequeño tratamiento se obtiene la siguiente equivalencia:

$$\frac{\zeta'(s)}{\zeta(s)} = -\sum_{k\in\mathbb{N},\ p \text{ primo}} p^{-ks}\log p = -\sum_{k\in\mathbb{N},\ p \text{ primo}} p^{-ks}\Lambda(p^k) = -\sum_{n\in\mathbb{N}} n^{-s}\Lambda(n)$$

la última igualdad se justifica al observar que su lado izquierdo ya considera la suma de todas las potencias de primo y solo se están agregando sumandos nulos ya que Λ se anula en números distintos a una potencia de un primo.

El siguiente lema trata sobre resolver una integral que necesitaremos en el próximo teorema. Para $a, c > 0$, la integral es la siguiente:

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{a^s ds}{s(s+1)}$$

La solución de la integral es 0 para $a \leq 1$ y $1 - a^{-1}$ para $a > 1$. ¿Cómo se resuelve esta integral en todo $\mathrm{Re}(s) = c$? Lo primero es integrar para una versión acotada del dominio: Desde $c - iT$ a $c + iT$. Después que se calcule o se acoten los términos, se podrá tomar $T \to \infty$ para llegar al resultado final.

Lo segundo es notar que la función que integramos es meromorfa y recordar que las integrales de funciones meromorfas en curvas cerradas simples son sencillamente residuo de los polos del interior de la curva cerrada. Luego una buena estrategia para afrontar esta integral de $c - iT$ a $c + iT$ es incrustar ese dominio en una curva cerrada. Para escoger seguiremos la figura 4.2: Hay dos caminos posibles siguiendo medias circunferencias, una por la de la izquierda $C_1(T)$ y otra por la de la derecha $C_2(T)$. La integral conjunta con la de la izquierda tiene dos polos con residuo $1 - a^{-1}$ y con la de la derecha no tiene ningún polo por lo que da 0.

En tercer lugar, vamos a acotar la integral en la semicircunferencia. El término $|1/(s(s-1))|$ se puede acotar superiormente en ambos casos por un término de la forma C/T^2, con C constante. Para acotar $|a^s| = a^{\mathrm{Re}(s)}$ un buen candidato es a^c, pero para asegurar que $a^{\mathrm{Re}(s)} \leq a^c$ escogeremos que semicircunferencia usar. Si $a \leq 1$, esto se cumplirá para cuando $\mathrm{Re}(s) \geq c$, por ende, usaremos a la semicircunferencia de la derecha. Si $a > 1$, entonces esto se cumplirá cuando $\mathrm{Re}(s) \leq c$, por ende, usaremos la semicircunferencia de la izquierda. Haciendo esto la integral en la semicircunferencia es acotable por C'/T, por lo que cuando $T \to \infty$, esta se hará cero concluyendo lo que deseamos.

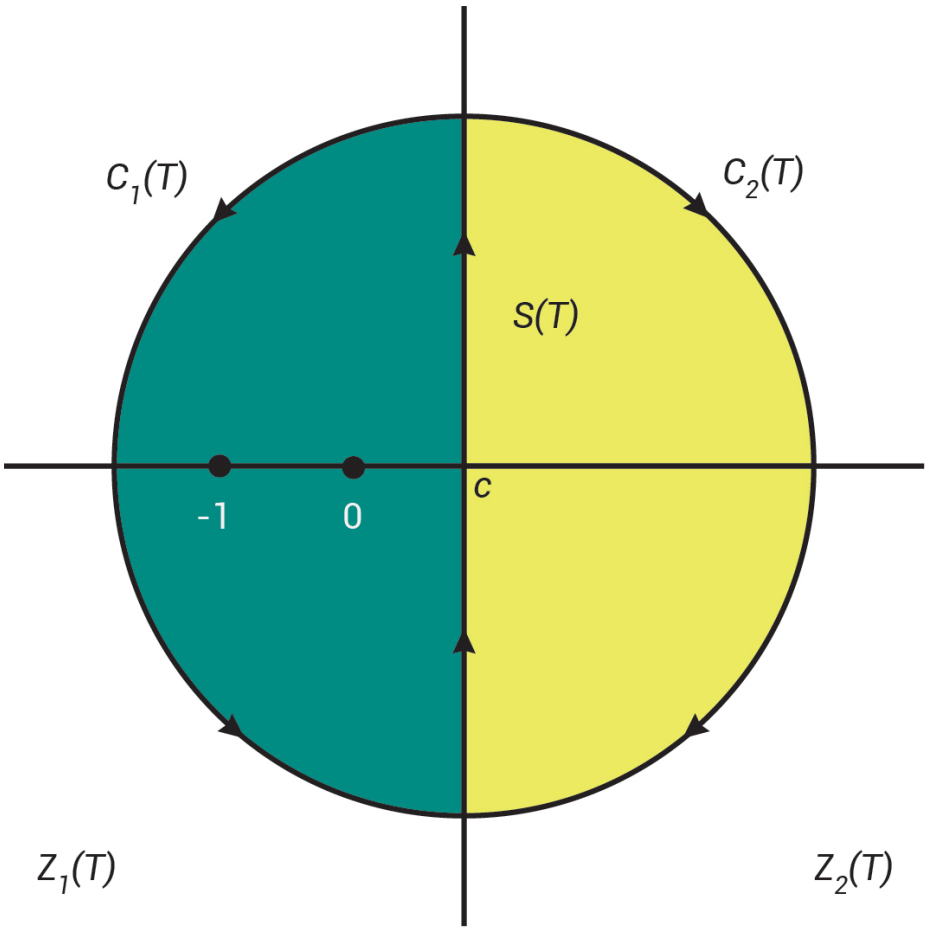


Figura 4.2: Gráfico de los caminos de integración de la demostración del lema 4.3.2, coloreados el interior de ambos: En ciánido el caso para a>1 y en amarillo para a entre 0 y 1.

Con la estrategia clara, procedamos a hacer las cuentas:

Lema 4.3.2 Si $c > 0$, entonces:

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{a^s ds}{s(s+1)}$$

vale 0 si $a \in]0,1]$ y $1-\frac{1}{a}$ si $a \geq 1$.

Demostración. Sea $T > 0$ cualquiera. Recordando que a^s es entera, el argumento de la integral solo tiene polos en $s = 0$ y $s = -1$. Con ello presente, vamos a considerar dos casos:

Si $a \in]0,1[$, considere la siguiente integral:

$$\frac{1}{2\pi i}\int_{Z_2(T)} \frac{a^s ds}{s(s+1)} = 0$$

Donde $Z_2(T)$ es el camino conformado por $S(T)$, que es la recta vertical orientada de forma ascendente desde $c-iT$ a $c+iT$, y $C_2(T)$ que es la semicircunferencia derecha orientada de forma descendente desde $c+iT$ a $c-iT$ (ver figura 4.2). Como el argumento de la integral es holomorfa tanto en el camino como en su interior, la integral se anula. Esto implica que:

$$\frac{1}{2\pi i}\int_{S(T)} \frac{a^s ds}{s(s+1)} + \frac{1}{2\pi i}\int_{C_2(T)} \frac{a^s ds}{s(s+1)} = 0$$

Lo siguiente es ver a que tiende el segundo sumando cuando $T \to \infty$. De que $a < 1$, se deduce que $|a^s| = a^{\mathrm{Re}(s)} \leq a^c$ para todo $s \in C_2(T)$. Como $C_2(T)$ es la mitad derecha de una circunferencia de radio T y centro $c > 0$, se tiene que $|s(s+1)| \geq T^2$ por desigualdad triangular. Luego se concluye que:

$$\left|\frac{1}{2\pi i}\int_{C_2(T)} \frac{a^s ds}{s(s+1)}\right| \leq \frac{1}{2\pi}\int_{C_2(T)} \left|\frac{a^s ds}{s(s+1)}\right| \leq \frac{a^c \pi T}{2\pi T^2} = \frac{a^c}{2T} \to 0$$

Aplicando a la expresión anterior, se logra cuando $T \to \infty$:

$$\frac{1}{2\pi i}\int_{S(T)} \frac{a^s ds}{s(s+1)} + \frac{1}{2\pi i}\int_{C_2(T)} \frac{a^s ds}{s(s+1)} \to \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{a^s ds}{s(s+1)} = 0$$

Si $a \geq 1$, considere la siguiente integral:

$$\frac{1}{2\pi i}\int_{Z_1(T)} \frac{a^s ds}{s(s+1)}$$

Donde $Z_1(T)$ es el camino conformado por $S(T)$ definido anteriormente y $C_1(T)$ que es la semicircunferencia izquierda orientada de forma descendente desde $c+iT$ a $c-iT$ (ver figura 4.2). Si T es suficientemente grande, se incluye ambos polos del argumento. Luego por fórmula del residuo se tiene:

$$\frac{1}{2\pi i}\int_{Z_1(T)} \frac{a^s ds}{s(s+1)} = \mathrm{res}_{s=0}\left(\frac{a^s}{s(s+1)}\right) + \mathrm{res}_{s=-1}\left(\frac{a^s}{s(s+1)}\right) = 1 - 1/a$$

La integral de la izquierda es separable en la suma de las siguientes dos integrales:

$$\frac{1}{2\pi i}\int_{S(T)}\frac{a^s ds}{s(s+1)}+\frac{1}{2\pi i}\int_{C_1(T)}\frac{a^s ds}{s(s+1)}=1-1/a$$

Para proceder similar a la parte anterior, nos basta probar que el segundo sumando tiende a 0 cuando $T\to\infty$. Como $a\geq 1$, $|a^s|=a^{\mathrm{Re}(s)}\leq a^c$ para todo $s\in C_1(T)$. $C_1(T)$ es la parte izquierda de una circunferencia centrada en $s=c$ de radio T. Se puede escoger un T suficientemente grande como para que esta circunferencia $D(c,T)$ contenga a $D(0,T/2)$ y $D(1,T/2)$, luego se tiene que $|s(s+1)|\geq T^2/4$ para todo $s\in C(T)$. Aplicando se obtiene para $T\to\infty$:

$$\left|\frac{1}{2\pi i}\int_{C_1(T)}\frac{a^s ds}{s(s+1)}\right|\leq\frac{4a^c\pi T}{2\pi T^2}=\frac{2a^c}{T}\to 0$$

Luego se concluye la demostración con que:

$$\frac{1}{2\pi i}\int_{S(T)}\frac{a^s ds}{s(s+1)}+\frac{1}{2\pi i}\int_{C_1(T)}\frac{a^s ds}{s(s+1)}\to\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{a^s ds}{s(s+1)}=1-1/a$$

Con los lemas aprendidos, se puede avanzar hacia la conexión entre la función dseta de Riemann y la función ψ de Tchebyshev, que como ya vimos conecta con el teorema de los números primos. Este vínculo viene dado por la siguiente integral:

$$\psi_1(x)=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{-x^{s+1}\zeta'(s)ds}{s(s+1)\zeta(s)}\qquad\forall c>1$$

Esta conexión proviene de las ideas de Benhard Riemann, formalizadas por Hans von Mangoldt. En esta formalización, esta expresión integral para ψ_1 vendría de una transformada inversa de Mellin[2] con reordenaciones de términos. Más precisamente, $x^{-2s-1}\psi_1(x)$ es la transformada inversa de Mellin de $\frac{-\zeta'(s)}{s(s+1)\zeta(s)}$. ¿Pero cómo se sabe aquello? Pues de la proposición 4.3.1 se puede expresar a $\frac{\zeta'(s)}{s(s+1)\zeta(s)}$ como la transformada de Mellin de $x^{-2s-1}\psi_1(x)$. Pues, como se mencionó en el capítulo de historia, la serie de la proposición 4.3.1 se puede entender como la integral de Riemann-Stieljes de x^{-s} con integrador $\psi(x)$. Si a esta integral se le hace integración por partes dos veces se tiene:

$$\frac{-\zeta'(s)}{\zeta(s)}=\sum_{n=1}^{\infty}\frac{\Lambda(n)}{n^s}=\int_1^{\infty}\frac{d\psi(x)}{x^{-s}}=-\int_1^{\infty}\frac{s\psi(x)}{x^{s+1}}dx=\int_1^{\infty}\frac{s(s+1)\psi_1(x)}{x^{s+2}}dx$$

Si se reordenan términos, se logra lo que se quiere. Sin embargo, nosotros probaremos el teorema 4.3.3 por otro camino, pero que no requiere conocimiento previo ni en transformadas de Mellín, ni en integrales de Riemann-Stieljes. La idea de la demostración es la siguiente: Demostrar que tanto el lado izquierdo como el lado derecho de la igualdad son iguales a la siguiente suma $\sum_{n=1}^{[x]}\Lambda(n)(x-n)$. Probar que la integral es igual a esta suma se requieren tres pasos: Primero, usar la proposición 4.3.1 para reemplazar $-\zeta'/\zeta$ por la serie. Segundo, usar el teorema de Fubini-Tonelli para invertir el orden entre la integral y la serie. Finalmente, tras factorizar por $\Lambda(n)x$, se reconstruye la integral del lema 4.3.2 con $a=x/n$, por lo que para $n>[x]$ la integral dará 0 y para los demás $(1-n/x)$. Demostrar que $\psi_1(x)$ se logra de aplicar la segunda forma para $\psi(x)$ de la proposición 4.2.1 en la definición de ψ_1 y luego hacer cálculos tradicionales. Esta última igualdad $\psi_1(x)=\sum_{n=1}^{[x]}\Lambda(n)(x-n)$ tiene una interpretación intuitiva, ya que siendo ψ_1 la integral de la función escalera ψ, es una función de rectas a trozos cuya pendiente va incrementándose. La suma $\sum_{n=1}^{[x]}\Lambda(n)(x-n)$ representa como se van acumulando estos incrementos lineales sucesivos.

[2]La transformada de Mellin, en honor al matemático finés Hjalmar Mellin, de f es $M[f](s)=\int_0^{\infty}x^{s-1}f(x)dx$. La transformada inversa de Mellin de g es $M^{-1}[g](x)=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}x^{-s}g(s)ds$.

Dichas las ideas, se procede a la demostración rigurosa:

Teorema 4.3.3 Para todo $c > 1$ se tiene:

$$\psi_1(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{-x^{s+1}\zeta'(s)ds}{s(s+1)\zeta(s)}$$

Demostración. Aplicando la proposición 4.3.1 a la integral del lado izquierdo, se obtiene que:

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{-x^{s+1}\zeta'(s)ds}{s(s+1)\zeta(s)} = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \sum_{n=1}^{\infty}\left(\frac{x^{s+1}}{s(s+1)}\right)\left(\frac{\Lambda(n)}{n^s}\right)ds$$

Considerando la serie como una integral de Stieljes, vemos que se corrobora el requisito para aplicar Fubini-Tonelli ya que:

$$\begin{aligned}
\int_{c-i\infty}^{c+i\infty} \sum_{n=1}^{\infty}\left(\frac{|x^{s+1}|}{|s||s+1|}\right)\left(\frac{\Lambda(n)}{|n^s|}\right)ds &= \int_{c-i\infty}^{c+i\infty}\left(\frac{x^{\mathrm{Re}(s)+1}}{|s||s+1|}\right)\sum_{n=1}^{\infty}\left(\frac{\Lambda(n)}{n^{\mathrm{Re}(s)}}\right)ds \\
&= \int_{c-i\infty}^{c+i\infty}\left(\frac{x^{\mathrm{Re}(s)+1}}{|s||s+1|}\right)\left(\frac{\zeta'(\mathrm{Re}(s))}{\zeta(\mathrm{Re}(s))}\right)ds \\
&\leq \int_{c-i\infty}^{c+i\infty}\left(\frac{x^{\mathrm{Re}(s)+1}}{|s|^2}\right)\left(\frac{\zeta'(\mathrm{Re}(s))}{\zeta(\mathrm{Re}(s))}\right)ds \\
&= x^{c+1}\left(\frac{\zeta'(c)}{\zeta(c)}\right)\int_{-\infty}^{\infty}\frac{dt}{c^2+t^2} \\
&= x^{c+1}\left(\frac{\zeta'(c)}{\zeta(c)}\right)\left(\frac{\tan^{-1}(t/c)}{c}\Bigg|_{-\infty}^{\infty}\right) \\
&= \frac{\pi x^{c+1}}{c}\left(\frac{\zeta'(c)}{\zeta(c)}\right) < \infty
\end{aligned}$$

Por lo que, aplicando Fubini-Tonelli y el lema 4.3.2 para $a = x/n$, se logra:

$$\begin{aligned}
\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{-x^{s+1}\zeta'(s)ds}{s(s+1)\zeta(s)} &= \sum_{n=1}^{\infty}\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\left(\frac{x^{s+1}}{s(s+1)}\right)\left(\frac{\Lambda(n)}{n^s}\right)ds \\
&= \sum_{n=1}^{\infty} x\Lambda(n)\left(\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{(x/n)^s ds}{s(s+1)}\right) \\
&= \sum_{n=1}^{[x]} x\Lambda(n)\left(\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{(x/n)^s ds}{s(s+1)}\right) \\
&\quad + \sum_{n=[x]+1}^{\infty} x\Lambda(n)\left(\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{(x/n)^s ds}{s(s+1)}\right) \\
&= \sum_{n=1}^{[x]} x\Lambda(n)(1-n/x) = \sum_{n=1}^{[x]}\Lambda(n)(x-n)
\end{aligned}$$

Y este último término es igual a ψ_1, como muestra el siguiente cálculo considerando $I_{n\leq x}(x)$ como la función característica del intervalo $[n,\infty)$:

$$\begin{aligned}\psi_1(x) &= \int_0^x \psi(x)dx = \int_0^x \left(\sum_{n=1}^{[x]} \Lambda(n)\right) dx \\ &= \int_0^x \left(\sum_{n=1}^{\infty} \Lambda(n) I_{n\leq x}(x)\right) dx = \sum_{n=1}^{\infty} \Lambda(n) \int_0^x I_{n\leq x}(x)dx \\ &= \sum_{n=1}^{[x]} \Lambda(n) \int_0^n I_{n\leq x}(x)dx + \sum_{n=1}^{[x]} \Lambda(n) \int_n^x I_{n\leq x}(x)dx + \sum_{n=[x]+1}^{\infty} \Lambda(n) \int_0^x I_{n\leq x}(x)dx \\ &= \sum_{n=1}^{[x]} \Lambda(n) \int_n^x dx = \sum_{n=1}^{[x]} \Lambda(n)(x-n)\end{aligned}$$

donde se puede aplicar Fubini-Tonelli ya que la expresión es de términos no negativos.

4.4 El último paso

Finalmente, solo queda el último paso. Nuestro objetivo: Demostrar la reducción del teorema de los números primos hecha por Tchebyshev.

$$\psi_1(x) \sim \frac{x^2}{2}$$

El punto de partida es la relación entre ψ_1 y la función dseta:

$$\psi_1(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{-x^{s+1}\zeta'(s)ds}{s(s+1)\zeta(s)}$$

Para entender la estrategia de demostración, debemos centrar la atención en la figura 4.3 que muestra tres caminos de integración. El primer camino es el eje $\mathrm{Re}(s) = c$. El segundo camino $\gamma(T)$ va por el eje $\mathrm{Re}(s) = 1$, en $1 - Ti$ se va a la derecha al eje $\mathrm{Re}(s) = c$, por este sube hasta $c + Ti$ y luego se devuelve a $\mathrm{Re}(s) = 1$. El tercer camino es similar, pero en vez de subir por $\mathrm{Re}(s) = c$ en ese breve segmento, lo hace por $\mathrm{Re}(s) = \delta$. Los parámetros $T > 1$ y $\delta \in (0,1)$ los fijaremos luego a nuestra conveniencia.

La estrategia de demostración consiste en dos partes. La primera parte es cambiar el dominio de integración desde $\mathrm{Re}(s) = c$ a $\gamma(T)$ y luego de $\gamma(T)$ a $\gamma(T,\delta)$. Más en específico, si se nombra como $F_x(s)$ al integrando, lo que se hará es ir probando las siguientes igualdades sucesivamente:

$$\psi_1(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} F_x(s)ds = \frac{1}{2\pi i} \int_{\gamma(T)} F_x(s)ds = \frac{x^2}{2} + \frac{1}{2\pi i} \int_{\gamma(\delta,T)} F_x(s)ds$$

La primera igualdad es el teorema 4.3.3, pero ¿cómo se prueban las siguientes igualdades? En ambos casos se partirá con la resta de las integrales involucradas. Como son integrales con el mismo argumento, pero distinto dominio, sucederá lo siguiente: En donde los caminos de integración coincidan las integrales se cancelarán, mientras lo demás de los dominios se conectarán, con la orientación invertida para la integral que esté restando. De forma intuitiva e informal, la resta de las integrales va a ser la integral de la "resta de los caminos". Observando la figura 4.3, la "resta" de $\gamma(\delta)$ menos $\mathrm{Re}(s) = c$ son dos Ues rectas, una que baja desde $c + i\infty$ hasta $c + iT$, de ahí se mueve horizontalmente hasta $1 + iT$ y luego sube a $1 + i\infty$, mientras la otra es su reflexión respecto al eje real. Por otro lado, la "resta" de $\gamma(\delta,T)$ menos $\gamma(T)$ es un rectángulo alrededor de $s = 1$.

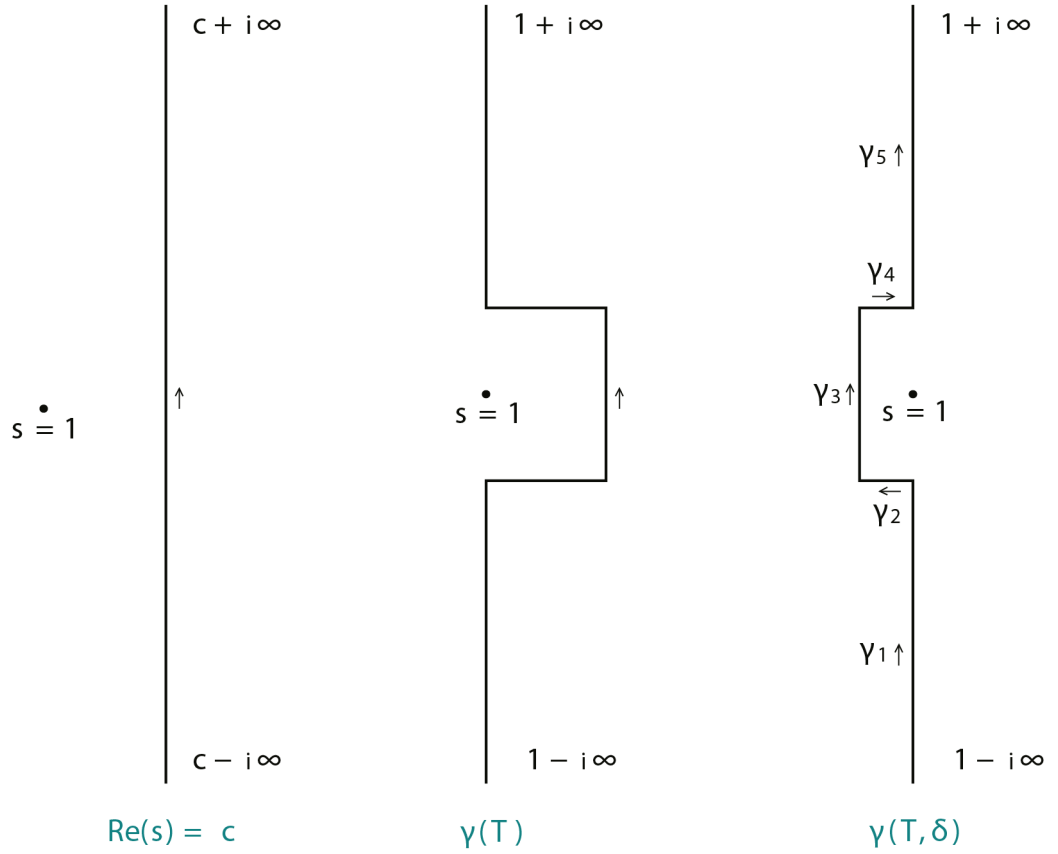


Figura 4.3: Gráfico de caminos de integración de la demostración del teorema 4.4.1

Para obtener las igualdades que deseamos, la integral sobre las Ues debe darnos 0, mientras la integral sobre el rectángulo alrededor de $s = 1$ debe dar $x^2/2$. ¿Cómo se pueden resolver estas integrales a pesar de la complicada expresión de $F_x(s)$? Pues se puede a través del teorema de los residuos sobre integrales cerradas simples. Entonces cabe preguntarse cuáles son los polos de $F_x(s)$ y, observando su expresión explícita, los únicos polos relevantes a la sección del plano complejo donde estamos trabajando es el polo $s = 1$ de ζ' y los ceros de la función ζ. De aquí que cobra importancia conocer sobre los ceros de la función dseta: Del teorema 3.4.1 sabemos que no hay ceros en $\mathrm{Re}(s) > 1$ y del teorema 3.4.2 sabemos que no hay ceros en $\mathrm{Re}(s) = 1$. Esto es suficiente para saber que en la integral sobre las Ues no habrá ningún polo, ni en el interior de la curva ni en la curva misma, por lo que ya sospechamos que su resultado es 0 (Aún queda un obstáculo adicional con esta integral a mencionar en el párrafo siguiente). Para el caso de la integral sobre el rectángulo, podemos asignar δ tan cercano a 1 que no haya ningún cero no trivial en el interior o el borde del rectángulo. Pero esto, ¿por qué se puede? La única forma de que no se pueda es que hubiera una sucesión de ceros no triviales que estuvieran tan cerca como se quiera de $\mathrm{Re}(s) = 1$. Sin embargo, como para un T fijo el rectángulo debe vivir en un conjunto compacto (entre $|\mathrm{Im}| \leq 1$ y $\mathrm{Re}(s) \in [0, c]$), esta sucesión debería tener una subsucesión convergente y ese límite sería un cero no trivial en $\mathrm{Re}(s) = 1$ contradiciendo el teorema 3.4.2. ¡De ahí la importancia del teorema 3.4.2 para el teorema de los números primos, al punto que es considerado su paso clave! Con esto el único polo al interior del rectángulo es $s = 1$ y calculando su residuo da justamente $x^2/2$.

En el caso de la integral sobre las Ues, hay otra complicación ya que las Ues no son curvas cerradas simples. Sin embargo, se puede corregir este problema usando un procedimiento similar al lema 4.3.2: Podemos ver las Ues como rectángulos donde uno de los lados se traslada hasta el infinito. Como no hay ningún polo ni en el borde ni en el interior, por Teorema de Cauchy las integrales sobre estos rectángulos acotados da siempre cero, luego su límite también es cero. Finalmente, basta mostrar que la integral sobre este lado que se va a infinito tiende a cero, a través de acotarla utilizando los teoremas de estimaciones 3.3.3 y 3.3.4.

Concluida esta primera parte, se habrá demostrado que:

$$\psi_1(x) = \frac{x^2}{2} + \frac{1}{2\pi i}\int_{\gamma(\delta,T)} F_x(s)ds$$

La integral funciona como un factor de error de aproximar a $\psi_1(x)$ con $x^2/2$, por lo que la demostración se logra si para x suficientemente grande se puede acotar la integral tan pequeña como se quiera. También recuerde que, más allá de exigir $T > 1$, aún tenemos ese parámetro libre a disposición. Por lo que esta segunda parte consiste en acotar la integral de $F_x(s)$ en $\gamma(\delta,T)$, para ello dividiremos el camino $\gamma(\delta,T)$ en sus segmentos lineales como muestra la figura 4.3. Se puede acotar de forma directa las integrales para cada subcaminos, siendo el único término complejo de acotar ζ'/ζ: En los subcaminos γ_1 y γ_5 este término se puede acotar usando los teoremas 3.3.3 y 3.3.4. En los otros subcaminos no se cumplen los requisitos de ambos teoremas, pero igualmente se puede acotar por su máximo garantizado por el Teorema de Weierstrass al ser estos caminos compactos. Al realizar todos estos cálculos, se llega a que:

$$\left|\frac{1}{2\pi i}\int_{\gamma(\delta,T)} F_x(s)ds\right| \leq (C_T^{(1)} + C_T^{(5)})x^2 + (C_{T,\delta}^{(2)} + C_{T,\delta}^{(4)})\frac{x^2}{\log x} + C_{T,\delta}^{(3)}x^{1+\delta}$$

Por lo que se puede concluir:

$$\left|\frac{2\psi_1(x)}{x^2} - 1\right| \leq 2(C_T^{(1)} + C_T^{(5)}) + (C_{T,\delta}^{(2)} + C_{T,\delta}^{(4)})\frac{2}{\log x} + 2C_{T,\delta}^{(3)}x^{\delta-1}$$

Cuando x es suficientemente grande, se puede reducir los últimos sumandos, mientras las dos primeras constantes $C_T^{(1)}$ y $C_T^{(5)}$ se pueden hacer tan pequeñas como se quieran incrementando el valor de T, propiedad que sale del acotamiento. Y con ello se logra la demostración del teorema de los números primos, a continuación su versión formal:

Teorema 4.4.1 $\psi_1 \sim \frac{x^2}{2}$. En consecuencia, $\pi(x) \sim \frac{x}{\log x}$.

Demostración. El punto de partida de la demostración está en el teorema 4.3.3: Sea $c > 1$, entonces

$$\psi_1(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{-x^{s+1}\zeta'(s)ds}{s(s+1)\zeta(s)} = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} F_x(s)ds$$

Para cualquier $x > 1$ y donde denominaremos al argumento de la integral como $F_x(s)$. Sea $T > 1$ por fijar luego. El primer paso consiste cambiar el camino de integración desde la recta $\mathrm{Re}(s) = c$ a $\gamma(T)$, que se puede observar en la figura 4.3, camino que asciende por la recta $\mathrm{Re}(s) = 1$ salvo por el intervalo $|\mathrm{Im}(s)| \leq T$ donde asciende por la recta $\mathrm{Re}(s) = c$, ambos subcaminos son conectados por caminos horizontales en una orientación coherente al sentido del camino total. La relación entre ambas integrales es de igualdad:

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} F_x(s)ds = \frac{1}{2\pi i}\int_{\gamma(T)} F_x(s)ds$$

Para demostrar este hecho, considere el rectángulo R_1 con orientación horaria cuyos vértices son $1+iT$, $1+iU$, $c+iU$ y $c+iT$, donde $U > T$ tan grande como queramos. De que la función $zeta'$ solo tiene un polo en $s = 1$ y que ζ no tiene ceros en $\mathrm{Re}(s) \geq 1$, F_x es una composición de funciones que son holomorfas tanto al interior como al borde de R_1, luego se deduce por teorema de Cauchy que:

$$\int_R F_x(s)ds = 0$$

$$\Rightarrow \int_T^U F_x(1+iy)dy + \int_1^c F_x(y+iU)dy + \int_U^T F_x(c+iy)dy + \int_c^1 F_x(y+iT)dy = 0$$

Concéntrese en el segundo sumando. Sea $\eta \in]0,2[$. Tomando $\sigma_0 = c > 1$, como $U > T > 1$ e $y \in [1,c]$, se puede utilizar los teoremas 3.3.3 y 3.3.4 para acotar dicho sumando:

$$\begin{aligned}
\left|\int_1^c F_x(y+iU)dy\right| &\leq \int_1^c \frac{|x^{y+iU+1}||\zeta'(y+iU)|}{|y+iU||y+1+iU||\zeta(y+iU)|}dy \\
&\leq x^{c+1}\int_1^c \frac{|\zeta'(y+iU)|}{|y+iU||y+1+iU||\zeta(y+iU)|}dy \\
&\leq x^{c+1}|U|^{-2}\int_1^c \frac{|\zeta'(y+iU)|}{|\zeta(y+iU)|}dy \\
&\leq x^{c+1}K^{(1)}_{\eta/2}|U|^{-2+\eta/2}\int_1^c \frac{dy}{|\zeta(y+iU)|} \\
&\leq (c-1)x^{c+1}K^{(1)}_{\eta/2}K_{c,\eta/2}|U|^{-2+\eta}
\end{aligned}$$

Por lo que si $U \to \infty$, el segundo sumando se va a cero dejando la siguiente igualdad:

$$\int_T^\infty F_x(1+iy)dy + \int_c^1 F_x(y+iT)dy = \int_T^\infty F_x(c+iy)dy$$

Por un argumento análogo al anterior para el rectángulo R_2 que se forma al reemplazar U con -U y T con -T, se tiene la siguiente igualdad:

$$\int_{-\infty}^{T} F_x(1+iy)dy + \int_1^c F_x(y+iT)dy = \int_{-\infty}^{T} F_x(c+iy)dy$$

Sumando ambas igualdades, luego sumando la integral de $F_x(s)$ en $\mathrm{Re}(s) = c$ desde $c-iT$ a $c+iT$ a ambos y finalmente dividiendo todo por $2\pi i$, se logra la igualdad deseada:

$$\psi_1(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} F_x(s)ds = \frac{1}{2\pi i}\int_{\gamma(T)} F_x(s)ds$$

Sea $\delta \in]0,1[$. El segundo paso es cambiar el camino de integración de $\gamma(T)$ a $\gamma(T,\delta)$, que se observa en la figura 4.3, que es un camino ascendente por la recta $\mathrm{Re}(s) = 1$, salvo en $|\mathrm{Im}(s)| \leq T$ donde el camino ascendente será en $\mathrm{Re}(s) = \delta$, conectándose ambos subcaminos por caminos horizontales con orientación coherente al sentido del camino total. Para ver la relación entre ambas integrales, veamos su resta:

$$\begin{aligned}
&\frac{1}{2\pi i}\left(\int_{\gamma(T)} F_x(s)ds - \int_{\gamma(T,\delta)} F_x(s)ds\right) \\
&= \frac{1}{2\pi i}\left(\int_1^c F_x(y-iT)dy + \int_{-T}^{T} F_x(c+iy)dy + \int_c^1 F_x(y+iT)dy - \int_{\gamma_2,\gamma_3,\gamma_4} F_x(s)ds\right) \\
&= \frac{1}{2\pi i}\int_{R(T,\delta)} F_x(s)ds
\end{aligned}$$

donde $R(T,\delta)$ es el rectángulo formado por los vértices $\delta - iT$, $c-iT$, $c+iT$ y $\delta+iT$ con orientación antihoraria. Se puede escoger un δ lo suficientemente cercano a 1 tal que ζ no tenga raíz alguna tanto en R como su interior. Esto se puede ya que si suponemos que no se puede, entonces para todo $\delta \in]0,1[$ debería existir una raíz de ζ en $R(T,\delta)$. Si tomamos la sucesión $\delta_n = 1 - 1/2n$, podríamos construir una sucesión de raíces de ζ que estén en $R(T,\delta_n)$. Todos estos rectángulos están contenidos por el primero $R(T,1/2)$ que es compacto, por lo que la sucesión de raíces debería tener una subsucesión convergente. Como $\delta_n \to 1$, el límite debe tener parte real mayor o igual a 1, y como ζ es continua, el límite también debe ser una raíz. Pero esto entra en contradicción con los teoremas 3.4.1 y 3.4.2 que prueba que no existe raíz de ζ con parte real mayor o igual 1, por contradicción entonces debe existir el δ deseado. Luego el único polo de F_x en $R(T,\delta)$ es $s = 1$ donde se indefine ζ', se tiene por el teorema de residuos que:

$$\frac{1}{2\pi i}\int_{R(T,\delta)} F_x(s)ds = \mathrm{res}_{s=1}\left(\frac{x^{s+1}}{s(s+1)}\right)\left(\frac{-\zeta'(s)}{\zeta(s)}\right) = \left(\frac{x^2}{2}\right)\mathrm{res}_{s=1}\frac{-\zeta'(s)}{\zeta(s)}$$

Del corolario 3.3.2, sabemos que $\zeta(s) = 1/(s-1) + f(s)$ con f holomorfa en H. Luego sabemos que $\zeta'(s) = -1/(s-1)^2 + f'(s)$ donde f' es holomorfa en H. Por lo que al calcular el residuo en el polo simple $s = 1$:

$$\begin{aligned}\text{Res}_{s=1}\frac{-\zeta'(s)}{\zeta(s)} &= \lim_{s\to 1}(s-1)\frac{-\zeta'(s)}{\zeta(s)} \\ &= \lim_{s\to 1}(s-1)\left(\frac{\frac{1}{(s-1)^2}+f'(s)}{\frac{1}{s-1}+f(s)}\right) \\ &= \lim_{s\to 1}\frac{1+(s-1)^2 f'(s)}{1+(s-1)f(s)} = 1\end{aligned}$$

Por lo que se logra en conclusión que:

$$\psi_1(x) = \frac{1}{2\pi i}\int_{\gamma(T)} F_x(s)ds = \frac{x^2}{2} + \frac{1}{2\pi i}\int_{\gamma(T,\delta)} F_x(s)ds$$

Finalmente, el tercer paso de la demostración es descomponer la última integral para cada subcamino presente en la figura 4.3 y acotar estas integrales por términos asintóticos a $x^2/2$. Cabe recordar que todavía podemos asignar T tan grande como se necesite mientras sea mayor a 1. Empezando tenemos:

$$\int_{\gamma(T,\delta)} F_x(s)ds = \int_{\gamma_1} F_x(s)ds + \int_{\gamma_2} F_x(s)ds + \int_{\gamma_3} F_x(s)ds + \int_{\gamma_4} F_x(s)ds + \int_{\gamma_5} F_x(s)ds$$

Para γ_1, como $|\text{Im(s)}| \geq T > 1$ y la parte real es 1, podemos usar tanto el teorema 3.3.3 y 3.3.4 para acotar usando $\eta \in (0,2)$ y $\sigma_0 = c$:

$$\begin{aligned}\left|\int_{\gamma_1} F_x(s)ds\right| &\leq \int_{-\infty}^{-T} \frac{|x^{2+iv}||\zeta'(1+iv)|dv}{|1+iv||2+iv||\zeta(1+iv)|} \\ &\leq x^2\int_{-\infty}^{-T}\frac{|\zeta'(1+iv)|dv}{|v|^2|\zeta(1+iv)|} \\ &\leq x^2 K^{(1)}_{\eta/2}K^{(2)}_{c,\eta/2}\int_{-\infty}^{-T}|v|^{\eta-2}dv \\ &= 2\pi C_T^{(1)}x^2\end{aligned}$$

Como la integral converge ya que $\eta < 2$, se puede escoger el T tan grande como se quiera para encoger el valor de esta integral y, por ende, reducir la constante positiva $C_T^{(1)}$ tanto como se necesite. Para γ_5, por un acotamiento análogo se obtiene que:

$$\left|\int_{\gamma_5} F_x(s)ds\right| \leq 2\pi C_T^{(5)}x^2$$

donde la relación de $C_T^{(5)}$ con T es análoga al caso anterior.

Para γ_2, se puede hacer la siguiente acotación usando el Teorema de Weierstrass por ser γ_2 un conjunto compacto:

$$\begin{aligned}
\left|\int_{\gamma_3} F_x(s)ds\right| &\leq \int_{\delta}^{1} \frac{|x^{u+1-iT}||\zeta'(u-iT)|du}{|u-iT||1+u+iT||\zeta(u-iT)|} \\
&\leq \frac{1}{T^2}\int_{\delta}^{1} \frac{x^{u+1}|\zeta'(u-iT)|du}{|\zeta(u-iT)|} \\
&\leq \left(\frac{1}{T^2}\right)\left(\max_{u\in[\delta,1]} \frac{|\zeta'(u-iT)|}{|\zeta(u-iT)|}\right)\int_{\delta}^{1} x^{u+1}du \\
&\leq \left(\frac{1}{T^2}\right)\left(\max_{u\in[\delta,1]} \frac{|\zeta'(u-iT)|}{|\zeta(u-iT)|}\right)\left(\frac{x^2-x^{\delta+1}}{\log x}\right) \\
&\leq \left(\frac{1}{T^2}\right)\left(\max_{u\in[\delta,1]} \frac{|\zeta'(u-iT)|}{|\zeta(u-iT)|}\right)\left(\frac{x^2}{\log x}\right) \\
&= 2\pi C^{(2)}_{T,\delta}\frac{x^2}{\log x}
\end{aligned}$$

Donde la constante $C^{(2)}_{T,\delta}$ depende de T y δ de forma desconocida a priori. Con un procedimiento análogo, para γ_4 se logra la desigualdad:

$$\left|\int_{\gamma_4} F_x(s)ds\right| \leq 2\pi C^{(4)}_{T,\delta}\frac{x^2}{\log x}$$

Cuya constante tiene la misma dependencia desconocida sobre los parámetros. Para γ_3 se acota por el siguiente procedimiento, también usando el teorema de Weierstrass al ser γ_3 compacto:

$$\begin{aligned}
\left|\int_{\gamma_3} F_x(s)ds\right| &\leq \int_{-T}^{T} \frac{|x^{\delta+1+iv}||\zeta'(\delta+iv)|dv}{|\delta+iv||1+\delta+iv||\zeta(\delta+iv)|} \\
&\leq \frac{x^{\delta+1}}{|\delta|^2}\int_{-T}^{T} \frac{|\zeta'(\delta+iv)|dv}{|\zeta(\delta+iv)|} \\
&\leq \left(\frac{2Tx^{\delta+1}}{|\delta|^2}\right)\left(\max_{v\in[-T,T]} \frac{|\zeta'(\delta+iv)|}{|\zeta(\delta+iv)|}\right) \\
&= 2\pi C^{(3)}_{T,\delta}x^{1+\delta}
\end{aligned}$$

Donde esta última constante también depende de los parámetros de formas a priori desconocidas. Considerando todas estas desigualdades, podemos concluir: Sea $\varepsilon > 0$. Entonces se tiene que:

$$\begin{aligned}
\left|\psi_1(x) - \frac{x^2}{2}\right| &= \frac{1}{2\pi}\left|\int_{\gamma(T,\delta)} F_x(s)ds\right| \\
&\leq \frac{1}{2\pi}\sum_{k=1}^{5}\left|\int_{\gamma_k} F_x(s)ds\right| \\
&\leq (C^{(1)}_T + C^{(5)}_T)x^2 + (C^{(2)}_{T,\delta} + C^{(4)}_{T,\delta})\frac{x^2}{\log x} + C^{(3)}_{T,\delta}x^{1+\delta}
\end{aligned}$$

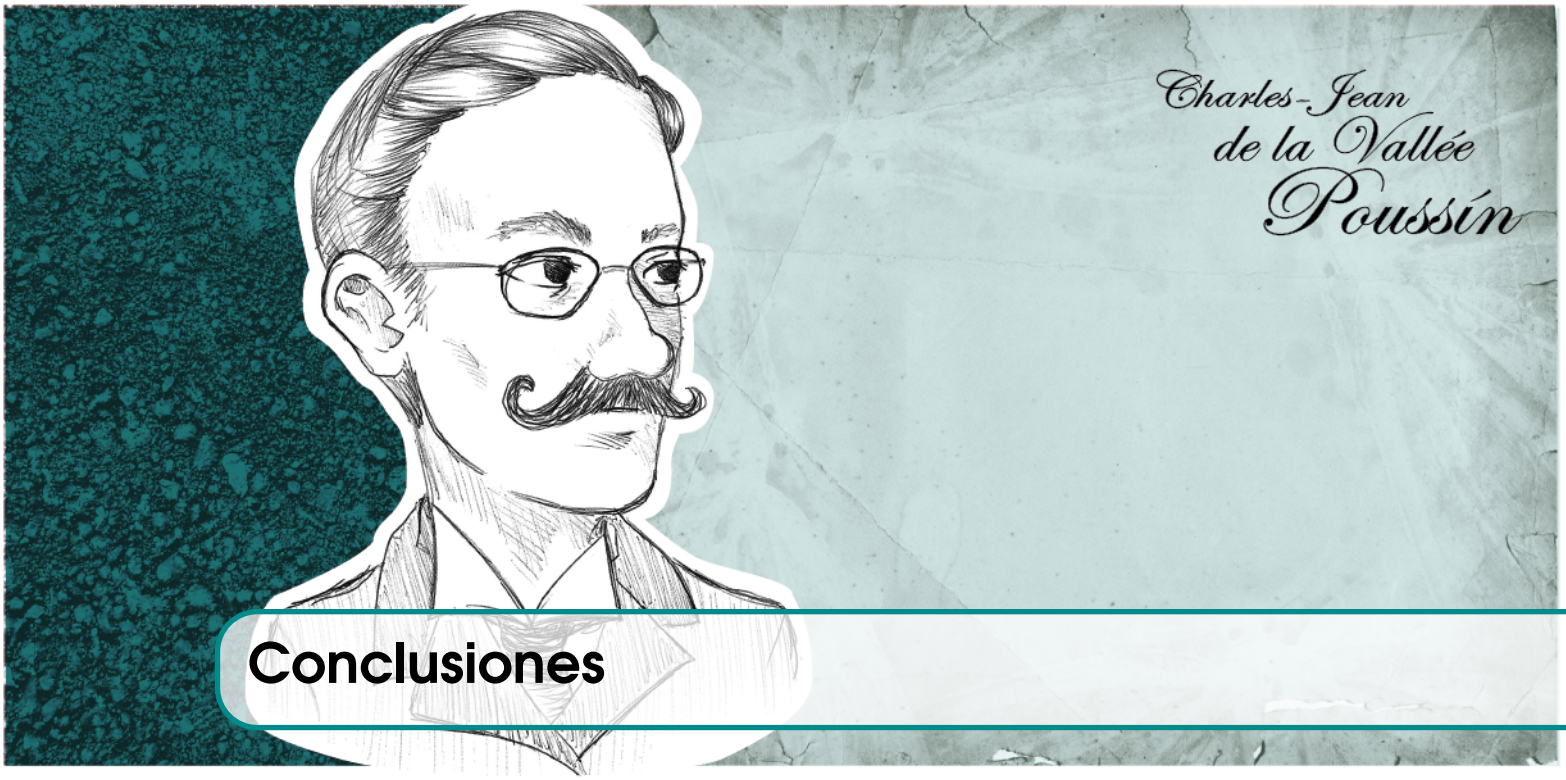
Dividiendo por $x^2/2$ se tiene:

$$\left|\frac{2\psi_1(x)}{x^2}-1\right| \leq 2(C_T^{(1)}+C_T^{(5)})+2(C_{T,\delta}^{(2)}+C_{T,\delta}^{(4)})\frac{1}{\log x}+2C_{T,\delta}^{(3)}x^{\delta-1}$$

Por la dependencia de las constantes 1 y 5, se escoge T lo suficientemente grande para que el primer sumando sea menor estricto a $\varepsilon/2$. Con T asignado, se escoge $\delta < 1$ lo suficientemente cercano a 1 tal que no hayan raíces en $R(T,\delta)$. Finalmente, se escoge x lo suficientemente grande para que los otros dos sumandos no superen en total $\varepsilon/2$ posible porque sumandos tienden a 0 cuando x tiende a infinito. Por lo que se tiene que para x suficientemente grande:

$$\left|\frac{2\psi_1(x)}{x^2}-1\right| < \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon \quad \Rightarrow \quad \psi_1 \sim \frac{x^2}{2}$$

Del teorema 4.2.3, se concluye entonces que $\pi(x) \sim \frac{x}{\log x}$, finalizando la demostración del teorema de los números primos.



Conclusiones

Se puede decir que el teorema de los números primos ocupa un lugar especial en las matemáticas y su historia. En parte esto proviene de su enunciado: El teorema fundamental de la aritmética enseña que los números primos son los bloques fundamentales que constituyen a los números enteros y, por ende, su estudio es valioso para comprender el comportamiento de lo discreto. Sin embargo, los números primos se distribuyen en la recta numérica en formas que escapan nuestro reconocimiento de patrones, aparentando una aleatoriedad que dificulta su estudio especialmente en escalas que superan nuestro poder computacional. De todas formas, esta aleatoriedad es simplemente una persistente ilusión, proveniente de la extraña recursividad que los forma: El siguiente primo está definido por los anteriores como el primer número natural que no es múltiplo de ninguno de ellos. Aunque todavía lejos de ser un patrón explícito, el teorema de los números primos revela un orden importante en esta distribución, especialmente a esas grandes escalas donde prevalecen sus enigmas.

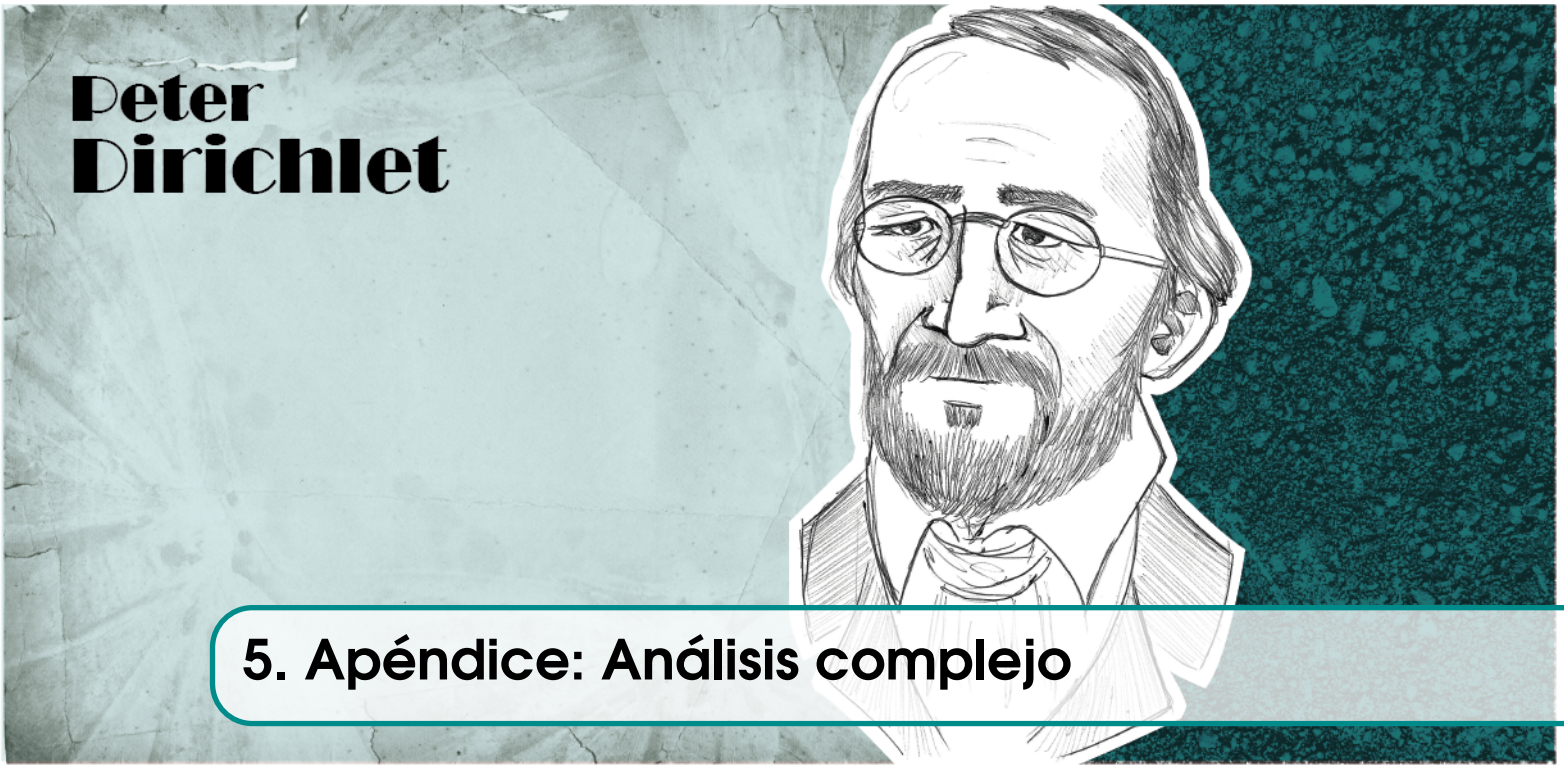
El teorema de los números primos ya sería remarcable solo por aquello, pero sería un error creer que toda su importancia se reduce a su enunciado. Su demostración revela una conexión inesperada entre lo continuo y lo discreto, entre el análisis complejo y la teoría numérica. El primer acercamiento a este puente fue realizado por Leonhard Euler para hacer una curiosa demostración alternativa sobre un hecho conocido desde los antiguos griegos: La existencia de una cantidad infinita de números primos. Pero sería primero Dirichlet y luego Riemann quienes revelarían todo el potencial de la identidad usada por Euler: Si la información sobre los números primos puede estar contenida en ciertas funciones, entonces se puede investigar utilizando las potentes herramientas del análisis. Fueron estos resultados los que fundamentaron el camino para lo que hoy conocemos como la Teoría Analítica de Números y que dan muestra de esta multidisciplinariedad interna de las matemáticas, donde a pesar de las diferencias entre sus distintas áreas están en una constante comunicación que ha sido una fuente fecunda de descubrimientos hasta nuestros días.

La historia del teorema de los números primos también es enriquecedora sobre cómo es la labor matemática. No todos los problemas son abordables de forma directa, a veces se descubren conjeturas que superan el entendimiento que tiene la humanidad en ese momento. Luego es necesario ir construyendo nuevas teorías, definiciones y teoremas alrededor del problema, hasta que las nuevas ideas vayan penetrando en la conjetura, como el agua en la cáscara de una nuez, hasta demostrarla con un esfuerzo final, muchas veces sencillo y elegante. La gran mayoría de las veces este avance será una contribución comunitaria entre mentes de distintas naciones y contextos, y que proseguirá más allá de la resolución del problema inicial, porque así como las nuevas ideas resolverán viejos problemas también levantarán nuevas interrogantes, así fue como para demostrar el teorema de los números primos llevó a la teoría sobre la función dseta que contenía en su haber la hipótesis de Riemann, cuya interrogante aún enfrenta -y al mismo tiempo une- a los matemáticos del mundo.

Grothendieck se refirió a esta generación de nuevas ideas como una marea subiente que iba erosionando la tierra de las dudas, que por más impenetrable que parezca terminará rindiéndose. Esta marea subiente responde otra cuestionante que suelen enfrentar las ciencias puras: El para qué de sus esfuerzos. Para quienes estén más avocados en las aplicaciones en campos ajenos a los que los números primos juegan un papel vital, podrían dudar la utilidad de estos esfuerzos centenarios más allá del valor del saber por el saber. Pero es esta marea subiente, esta nueva generación de ideas lo que va alimentando la ciencia hasta llegar a las aplicaciones. El teorema de números primos impulsó los estudios sobre el análisis complejo, que fueron centrales para entender el mundo cuántico que hoy define nuestra tecnología. Esta es la importancia de hacerse grandes preguntas, pues el desafío difícil lo que alimenta la generación de nuevos avances e ideas cuyas consecuencias son insospechadas. Gauss o Legendre seguramente no creían que el análisis complejo demostraría la relación que habían encontrado entre unas tablas de números primos y logaritmos, tampoco Euler podría haber sospechado que su curiosa demostración generaría una cascada de inspiraciones que alimenta hasta hoy la Teoría Analítica de Números. Son estos spillovers de la ciencia que alimentan la tecnología que define nuestra era.

En este sentido, cabe cerrar recordando las palabras del presidente estadounidense John Fitzgeral Kennedy sobre la importancia de grandes desafíos, en un discurso memorable para la historia de la ciencia:

> ¿Por qué, algunos dirán, la luna? ¿Por qué escoger este como nuestro objetivo? Y ellos tal vez pregunten ¿Por qué subir la montaña más alta? ¿Por qué hace 30 años volamos sobre el Atlántico? [...] ¡Nosotros escogimos ir a la luna en esta década y hacer las otras cosas no porque sean fáciles, sino porque son difíciles! Porque ese objetivo servirá para organizar y medir lo mejor de nuestras energías y habilidades, porque ese desafío es uno que estamos dispuestos a aceptar, uno que no estamos dispuestos a posponer y es uno que pretendemos ganar, y los otros también. [10]



5. Apéndice: Análisis complejo

Este apéndice contiene algunos teoremas que se usan a lo largo de la obra, pero que no conforman el núcleo del estudio. Se adjuntan ideas de demostración y referencias.

5.1 Teoremas sobre funciones holomorfas

Teorema 5.1.1 Sea $F : \Omega \times [0,1] \to \mathbb{C}$ con $\Omega \subset \mathbb{C}$ abierto. Si F cumple con:

- $F(s,x)$ es holomorfa en s para cada x fijo.
- F es continua.

Entonces es holomorfa la función $f : \Omega \to \mathbb{C}$ definida por:

$$f(s) = \int_0^1 F(s,x)dx$$

Idea de Demostración. Ver [14], capítulo 2 teorema 5.4. Aplicar Teorema de Fubini-Tonelli a

$$\int_T \int_0^1 F(s,x)dx$$

con T cualquier triángulo de un disco arbitrario D. Luego aplicar Teorema de Morera.

Teorema 5.1.2 Sea $f_n : \Omega \subset \mathbb{C} \to \mathbb{C}$ una sucesión de funciones holomorfas en Ω que converge uniformemente a $f : \Omega \to \mathbb{C}$, entonces f es holomorfa en Ω.

Idea de Demostración. Ver [14], capítulo 2 teorema 5.2. Para cualquier triángulo T contenido en un disco arbitrario D con clasura contenida en Ω, $\int_T f_n(s)ds = 0$ por Teorema de Goursat. Por convergencia uniforme, f es continua y $\int_T f_n(s)ds \to \int_T f ds = 0$. Por Teorema de Morera, f es holomorfa en D. Como D es arbitrario, f es holomorfa en Ω.

Proposición 5.1.3 Para $a \in (0,1)$:

$$\int_0^\infty \frac{x^{-a}dx}{1+x} = \frac{\pi}{\sin \pi a}$$

Idea de Demostración. Ver [14], capítulo 3 ejemplo 2.1. Hacer cambio de variable $x = e^u$. Resolver integral sobre el rectángulo de $|\mathrm{Re}(s)| \leq R$ y $\mathrm{Im}(s) \in [0, 2\pi]$ con el teorema de residuo, donde el único polo es $s = \pi i$. Cuando $R \to \infty$, la integral del lado inferior se vuelve la deseada, la del lado superior es un múltiplo de la deseada y las integrales de los otros dos lados se van a 0 por acotamiento. Se despeja y se obtiene lo deseado.

Teorema 5.1.4 — Criterio de Weierstrass o Prueba M. Sean $\Omega \subset \mathbb{C}$. Si para una sucesión de funciones $f_n : \Omega \to \mathbb{C}$ existe una sucesión de valores positivos ρ_n tal que $\sum_{n=1}^\infty \rho_n$ converge y $|f_n(s)| \leq \rho_n$ para todo $x \in \Omega$, entonces la sucesión de sumas parciales $\sum_{n=1}^N f_n(s)$ converge uniformemente.

Idea de Demostración. Sea $\varepsilon > 0$. Entonces $\sum_{n=1}^\infty |f(x)| \leq \sum_{n=1}^\infty \rho_n \leq \infty$, luego las sumas parciales convergen absolutamente. Además sea $\varepsilon > 0$:

$$\left| \sum_{n=1}^{\infty} f_n(s) - \sum_{n=1}^{N} f_n(s) \right| \leq \sum_{n=N+1}^{\infty} |f_n(s)| \leq \sum_{n=N+1}^{\infty} \rho_n$$

Y se puede escoger N para que esta última serie sea menor que ε para todo $s \in \Omega$, teniendo la convergencia uniforme.

Teorema 5.1.5 Si $\sum_{n=1}^\infty |a_n| < \infty$, entonces $\prod_{n=1}^\infty 1 + a_n$ converge. Además, $\prod_{n=1}^\infty 1 + a_n \neq 0$ si y solo si $1 + a_n \neq 0$ para todo $n \in \mathbb{N}$.

Idea de Demostración. Ver [14], capítulo 5 proposición 3.1. Como $\sum |a_n| < \infty$, considere desde el N que $|a_n| < 1/2$. Usando que el $\log(1+z)$ dado por la clásica serie de potencias cumple $1 + z = e^{\log(1+z)}$, se tiene:

$$\prod_{n=N}^{M} 1 + a_n = e^{\sum_{n=N}^{M} \log(1+a_n)}$$

Y esta serie converge absolutamente cuando $M \to \infty$ por comparación $\log|1 + a_n| \leq 2|a_n|$. Como $e^x \neq 0$, la única forma que sea cero es que un factor anterior a N sea cero.

Teorema 5.1.6 — Criterio de Weierstrass para producto. Sean $\Omega \subset \mathbb{C}$. Si para una sucesión de funciones $f_n : \Omega \to \mathbb{C}$ existe una sucesión de valores positivos ρ_n tal que $\sum_{n=1}^\infty \rho_n$ converge y $|f_n(s) - 1| \leq \rho_n$ para todo $x \in \Omega$, entonces la sucesión de productos parciales $\prod_{n=1}^N f_n(s)$ converge uniformemente.

Idea de Demostración. Ver [14], capítulo 5 proposición 3.2. Aplicar el teorema 5.1.5 para $a_n = f_n(s) - 1$ para cada $s \in \Omega$, cuyo $|a_n| \leq \rho_n$. Como la comparación es la misma independiente de s, la convergencia será uniforme.

5.2 Función Theta

Definición 5.2.1 — Función Theta. La función theta es $\theta : \mathbb{R}^+ \to \mathbb{R}$ dada por:

$$\theta(t) = \sum_{n=-\infty}^{\infty} e^{-\pi n t}$$

Proposición 5.2.1 La función theta es continua.

Idea de Demostración. Expresar $\theta = 1 + 2\sum_{n=1}^{\infty} e^{-\pi n t}$. Usar Criterio de Weierstrass con $|e^{-\pi n t}| \leq e^{-\pi n}$ para demostrar convergencia uniforme. Por convergencia uniforme, se tiene la continuidad.

Proposición 5.2.2 La función theta es exponencialmente dominada, es decir, existe $C > 0$ tal que:

$$|\theta(t) - 1| \leq C_1 e^{-\pi t}$$

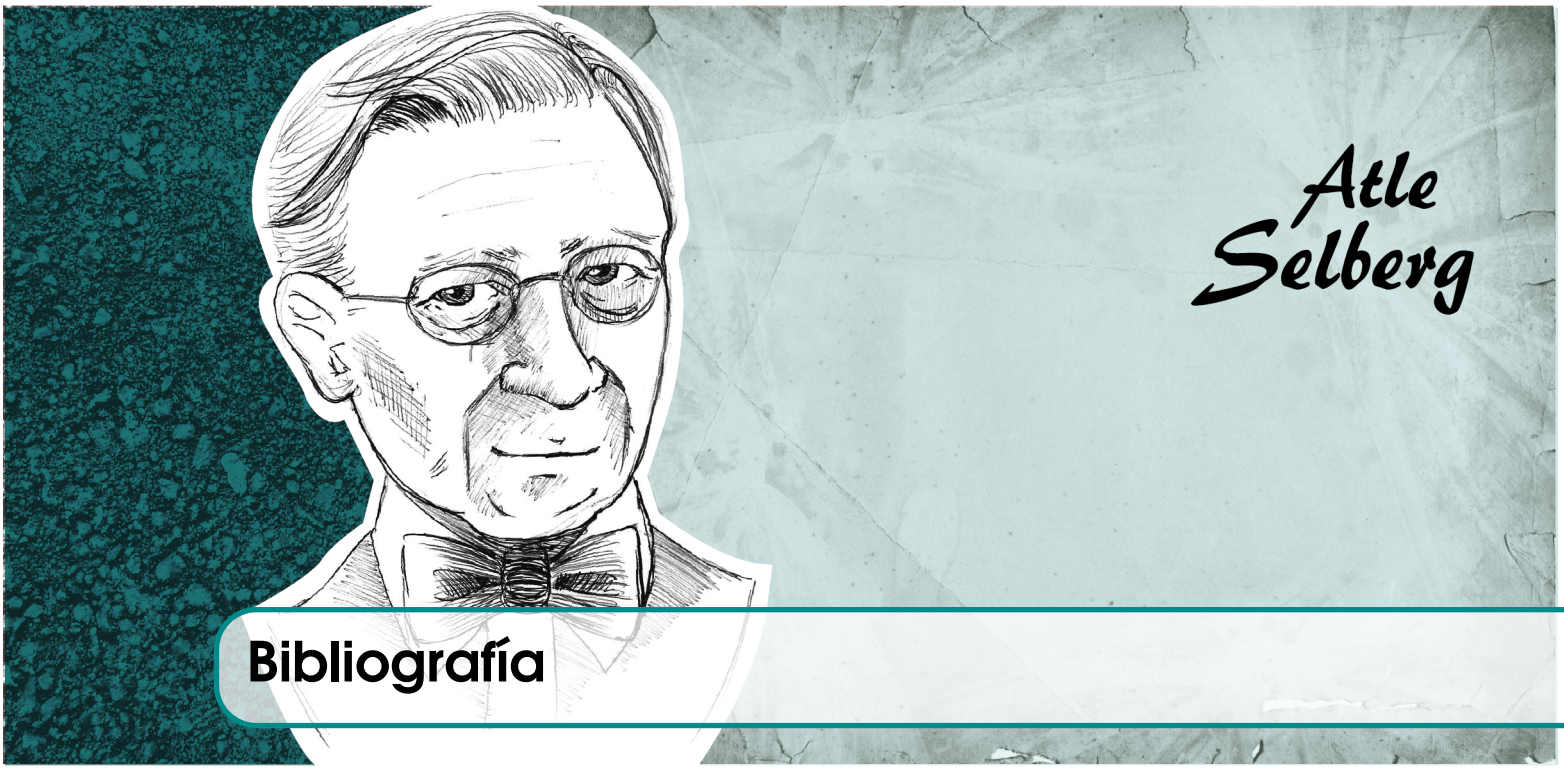
Idea de Demostración. Ver [14], capítulo 6 página 169. De la siguiente desigualdad:

$$|\theta(t) - 1| \leq 2\sum_{n=1}^{\infty} e^{-\pi n^2 t} \leq 2\sum_{n=1}^{\infty} e^{-\pi n t} = \frac{2}{e^{\pi t} - 1} \leq C e^{-\pi t}$$

Proposición 5.2.3 La función theta cumple la siguiente igualdad:

$$\theta(t) = t^{-1/2}\theta(1/t)$$

Idea de Demostración. Ver [14], en capítulo 4 teorema 2.4 y ecuación 4 (pág. 120). Directo de la Fórmula de Poisson.



Bibliografía

[1] APOSTOL, Tom M. Introduction to Analytic Number Theory. Estados Unidos, Springer Science+Business Media, 1976. p. 329. (Serie Undergraduate Texts in Mathematics).

[2] APOSTOL, Tom M. A Centennial History of the Prime Number Theorem. En: BAMBAH R.P., DUMIR V.C., HANS-GILL R.J. (eds.). Number Theory. India, Birkhäuser Basel, 2000, pp. 1-14.

[3] AUDIN, Michèle, ROYO-LETELIER, Jimena (trad.), DE VILLEGAS, Julio E. (trad.). Jacques Hadamard y el Teorema de los Números Primos [en línea]. Paisajes Matemáticos. <https://images.math.cnrs.fr/Nueva-traduccion-Jacques-Hadamard-y-el-teorema-de-los-numeros-primos.html?lang=es> [Consulta: 29 de noviembre 2019]

[4] CALDWELL, Chris K. y XIONG, Yeng. What is the smallest prime? [en línea]. Journal of Integer Sequences. 2012. Volumen 15. Artículo 12.9.7. <https://cs.uwaterloo.ca/journals/JIS/vol15.html> [Consulta: 29 de diciembre 2019].

[5] DWYMARK. Meaning of Rays in Polar Plot of Prime Numbers. Math Stack Exchange. <https://math.stackexchange.com/questions/885879/meaning-of-rays-in-polar-plot-of-prime-numbers> [Consulta: 5 de marzo de 2020]

[6] GOLDFELD, Dorian. The Elementary Proof of the Prime Number Theorem: An Historical Perspective. En: CHUDNOVSKY D., CHUDNOVSKY G., NATHANSON M. (eds.). Number Theory. Estados Unidos, Springer, 2004, pp. 179-192.

[7] GOLDSTEIN, Larry J. A History of the Prime Number Theorem. The American Mathematical Monthly. 80(6):599-615, jun-jul. 1973.

[8] HAWKING, Stephen. Dios creo los números 5 ed. España, Crítica, 2008. p. 1031.

[9] HEATH, Thomas L. The thirteen books of Euclid's Elements, Volume II. 2a. ed. Estados Unidos, Dover Publications Inc, 1956.

[10] KENNEDY, John F. Moon Speech. [en línea]. Estados Unidos, 1962. En: Johnson Space Center. <https://er.jsc.nasa.gov/seh/ricetalk.htm> [Consulta: 12 de octubre 2020]

[11] KLYVE, Dominic. The Origin of the Prime Number Theorem [en línea]. Ursinus College, Number Theory 9. 2018. <https://digitalcommons.ursinus.edu/triumphs_number/9/> [Consulta: 12 de marzo de 2020].

[12] SANDERSON, Grant. Why do prime numbers make these spirals?. 3Brown1Blue (Youtube). <https://www.youtube.com/watch?v=EK32jo7i5LQ> [Consulta: 5 de marzo de 2020]

[13] SELBERG, Atle. An Elementary Proof of the Prime-Number Theorem. Annals of Mathematics. Second Series, 50(2):305-313, apr. 1949.

[14] STEIN, Elías M. y SHAKARCHI, Rami. Complex Analysis. Estados Unidos, Pricenton University Press, 2003. p. 379. (Serie Pricenton Lectures in Analysis).

[15] WEISSTEIN, Eric W. Prime counting function [en línea]. MathWorld - A Wolfram Web Resource. <https://mathworld.wolfram.com/PrimeCountingFunction.html> [Consulta: 24 de marzo 2020]