UNIVERSIDAD TECNOLÓGICA DE LA MIXTECA

"**La distribución de los números primos**"

Tesis

que para obtener el título de

Licenciado en Matemáticas Aplicadas

presenta

José Hernández Santiago

Director de tesis:

M. C. Vulfrano Tochihuitl

Co-director:

Dr. Eugenio Balanzario

Huajuapan de León, Oaxaca. Agosto de 2010.

A todos aquellos que han estado a punto de perder la esperanza.

Prefacio

En este trabajo se lleva a cabo un estudio del resultado central sobre la distribución de los números primos entre los números naturales: el Teorema del Número Primo.

Las primeras versiones (conjeturales) de dicho teorema fueron presentadas por C. F. Gauss y A. M. Legendre hacia finales del siglo XVIII, pero no fue sino hasta aproximadamente cien años después que J. Hadamard y C. J. de la Vallée Poussin publicaran las primeras demostraciones del teorema.

Una característica de nuestro tratamiento del Teorema del Número Primo es que, aún cuando se optó por enfocar la atención hacia las pruebas más recientes que se conocen del resultado, jamás se perdió de vista el desarrollo histórico que conllevó a las demostraciones de Hadamard y de de la Vallée Poussin. Una muestra de esto es el análisis que hemos hecho de la memoria en [3], en la cual, P. L. Chebyshev derivó las primeras estimaciones concretas sobre el orden de magnitud de la función contadora de primos.

El trabajo consta de tres capítulos. En el primero de ellos se motiva el tema central de la obra desde la interesante temática de la generación de números primos mediante fórmulas elementales. En ese capítulo se presentan también las formulaciones *efectivas* del Teorema del Número Primo que probamos posteriormente.

En el capítulo 2 se ponen de manifiesto, una y otra vez, algunas de las sorprendentes conexiones entre la función zeta de Riemann y los números

primos. La primera de ellas es la representación en producto de Euler para la función zeta y la más impactante de todas es la equivalencia entre el Teorema del Número Primo y la no anulación de la función zeta de Riemann en la línea $\Re(z) = 1$. Es en este capítulo donde presentamos las pruebas del Teorema del Número Primo que optamos abordar. La primera de la pruebas expuestas se debe a Donald J. Newman y la segunda se obtiene como una aplicación del teorema tauberiano de N. Wiener y S. Ikehara. Al final de ese capítulo se presentan algunas aplicaciones notables del Teorema del Número Primo.

El resultado básico del capítulo 3 es la prueba de un célebre teorema debido a J. P. G. L. Dirichlet, a saber, la infinitud de primos en una progresión aritmética de números naturales donde el término inicial es coprimo con la diferencia común. Las investigaciones de Dirichlet son un referente obligado en las recolecciones sobre el Teorema del Número Primo porque, de acuerdo con L. J. Goldstein, *con ellas se introdujo a la Teoría de Números, la fértil idea de que los métodos analíticos podían ser exitosamente aplicados a problemas aritméticos*.

Dado que el trabajo de Dirichlet sobre primos en progresiones aritméticas (c. 1837) antecedió a la memoria de Chebyshev mencionada arriba (c. 1850), la disposición hecha de los capítulos puede resultar un tanto paradójica a primera vista. No obstante, la explicación de dicho seguimiento es simple: en el capítulo 3 se contempla también, entre otros puntos, la cuestión del comportamiento asintótico de las funciones contadoras de primos dentro de progresiones aritméticas y las respuestas que brindamos dependen del teorema de Wiener-Ikehara que se introduce al final del capítulo 2.

Están en orden ahora algunos comentarios en cuanto a la notación empleada. Como es usual en la literatura de Teoría de Números, la expresión $\log n$ denota siempre al logaritmo natural de n. A lo largo del trabajo, el símbolo $\mathbf{P}$ denota al conjunto de números primos positivos, p es siempre un primo fijo o una variable que toma valores en $\mathbf{P}$ y la expresión $p^{\alpha} \| n$

indica que p^α divide a n, pero $p^{\alpha+1}$ no. La parte entera del número x la denotamos con $\lfloor x \rfloor$ y $\phi(n)$ es la función de Euler evaluada en n.

Ahora bien, si $s \in \mathbb{C}$ entonces $s = \sigma + it$ para algunos σ y t en $\mathbb{R}$. Luego, para $A \in \mathbb{R}$, la relación $s \in \sigma > A$ indica que s está en el semiplano de los números complejos con parte real mayor que A. La expresión $\Re(z)$ es la parte real del número complejo z y $\Im(z)$ la parte imaginaria.

El símbolo O de Bachmann-Landau se emplea con la denotación usual. Esto es, si $f(n) = O(g(n))$ entonces existe una constante $A > 0$ tal que $|f(n)| \leq A|g(n)|$ siempre que n es suficientemente grande. La notación $f(n) = o(g(n))$ indica, por su parte, que $f(n)/g(n) \to 0$ cuando $n \to \infty$. El símbolo $\asymp$ ubicado entre dos funciones indica que las funciones tienen el mismo orden de magnitud. Esto es, si $f(n) \asymp g(n)$ entonces $A_1|g(n)| \leq |f(n)| \leq A_2|g(n)|$ para todo n suficientemente grande. Por último, la relación $f(n) \sim g(n)$ indica que f y g son asintóticamente equivalentes, esto es, que $f(n)/g(n) \to 1$ cuando $n \to \infty$.

Por último, pero no por ello menos importante, deseo expresar en este espacio mis agradecimientos a los sinodales Lic. Juan Luis Hernández, M.C. Adolfo Maceda y M. C. Alma Lidia Piceno por la minuciosa revisión que hicieron de esta tesis y por sus atinadas correcciones. En particular, agradezco al M. C. Vulfrano Tochihuitl por su apoyo durante la realización de este proyecto y al Dr. Eugenio Balanzario por ayudarme a definir la ruta que este trabajo tenía que seguir. Agradezco también a la M. C. Luz del Carmen Álvarez Marín por toda la ayuda que recibí de ella durante mi tiempo en la UTM.

José Hdz. Stgo.
Agosto de 2010.

Índice general

Capítulo 1

Preludio

§**1.** La infinitud del conjunto de los números primos era un resultado bien conocido para los antiguos griegos. La prueba que de este hecho Euclides dejara para nosotros es una hermosa ilustración de la efectividad de las pruebas constructivas. Aún así, nuestro trabajo no empezará con la exposición del argumento aquel sino con el estudio de una aseveración un tanto más fina que el sólo apunte sobre el cardinal del conjunto de primos, a saber,

Proposición 1.1. La serie de los recíprocos de los números primos diverge.

Prueba. Por contradicción. Si denotamos con **P** al conjunto de primos y suponemos que $\sum_{p\in\mathbf{P}} \frac{1}{p}$ converge, el criterio de Cauchy para series nos permite asegurar la existencia de $k \in \mathbb{N}$ tal que

$$\sum_{i\geq k+1} \frac{1}{p_i} < \frac{1}{2}. \tag{1.1}$$

Una vez que se ha fijado dicha k hacemos las siguientes convenciones: los primos $p_1, p_2, \ldots, p_k$ serán primos tipo **A** y los primos $p_{k+1}, p_{k+2}, \ldots$ serán primos tipo **B**. Además, si N es un natural mayor que 1 denotamos con $N_{\mathbf{B}}$ al número de naturales menores o iguales a N que son divisibles por al menos

un primo tipo **B**. $N_{\mathbf{A}}$ será, por su parte, el número de naturales menores o iguales a N que son divisibles sólo por primos tipo **A**. Claramente, para cada natural N distinto de 1, las convenciones hechas implican que

$$N_{\mathbf{A}} + N_{\mathbf{B}} = N. \tag{1.2}$$

Procedemos ahora a estimar $N_{\mathbf{B}}$. La expresión $\left\lfloor \frac{N}{p_i} \right\rfloor$ es igual al número de naturales menores o iguales a N que son divisibles por p_i. La desigualdad en (1.1) implica entonces que

$$N_{\mathbf{B}} \leq \sum_{i \geq k+1} \left\lfloor \frac{N}{p_i} \right\rfloor \leq \sum_{i \geq k+1} \frac{N}{p_i} < \frac{N}{2}. \tag{1.3}$$

Ahora bien, para obtener información con respecto a $N_{\mathbf{A}}$ notamos en primer lugar que todo natural n menor o igual a N que sólo tiene divisores del tipo **A** se puede escribir en la forma $n = a_n b_n^2$, donde a_n es libre de cuadrados y por tanto producto de distintos primos tipo **A**. Puesto que sólo hay k primos tipo **A**, se sigue que a lo sumo hay 2^k maneras diferentes de formar la parte libre de cuadrados de n. Más aún, las desigualdades $b_n \leq \sqrt{n} \leq \sqrt{N}$ implican que $N_{\mathbf{A}} \leq 2^k \sqrt{N}$. En particular, si $N = 2^{2(k+1)}$ la desigualdad previa indica que

$$N_{\mathbf{A}} \leq 2^k \sqrt{2^{2(k+1)}} = \frac{N}{2}. \tag{1.4}$$

Luego, para esta elección particular de N se tiene, en la luz de (1.3), que

$$N_{\mathbf{A}} + N_{\mathbf{B}} < \left(\frac{N}{2}\right) + \left(\frac{N}{2}\right) = N,$$

lo cual contradice lo afirmado en (1.2) y la prueba termina.

□

La primera demostración en la historia de la proposición anterior se atribuye a L. Euler. No obstante, la prueba que se ha presentado arriba es básicamente el argumento que P. Erdős presentara en su trabajo *Über die Reihe* $\sum \frac{1}{p}$ de 1938. El resultado proporciona tanto una demostración incidental de la infinitud de **P** como evidencia temprana sobre la *peculiaridad*

de la aparición de los primos entre los números naturales. Para aclarar un poco lo que se pretende dar a entender aquí con el adjetivo *peculiar* vamos a considerar la siguiente pregunta: ¿será cierto que si A es un subconjunto propio e infinito de $\mathbb{N}$ entonces la serie $\sum_{a\in A}\frac{1}{a}$ diverge?

Si la respuesta a nuestra interrogante fuera siempre afirmativa entonces estaríamos claros en que la proposición **1.1** sería típica de todo subconjunto propio e infinito de $\mathbb{N}$ y que por tanto no habría nada de particular en el hecho de que **P** la cumpliera, menos aún cuando se consideran los $2^{\aleph_0}$ subconjuntos infinitos propios[1] de $\mathbb{N}$. El hecho de que la respuesta sea en muchas ocasiones negativa indica que la pregunta es digna de consideración. Aún cuando no resulte aparente cómo proponer una caracterización sencilla de los subconjuntos de $\mathbb{N}$ para los cuales la respuesta es afirmativa, es interesante notar que el hecho de que la respuesta sea positiva para un subconjunto A de los naturales indica, de cierto modo, que A es un subconjunto *sustancioso* de los números naturales. Una manera de precisar la intuición detrás del uso del término *sustancioso* se tiene al dotar a los números naturales de una topología τ de tal manera que los subconjuntos densos de $\mathbb{N}$ bajo dicha topología se correspondan biunívocamente con los subconjuntos *sustanciosos* de $\mathbb{N}$, esto es, aquellos subconjuntos A de $\mathbb{N}$ para los cuales la serie $\sum_{a\in A}\frac{1}{a}$ diverge.

Otro punto importante a favor de la peculiaridad de **P** como consecuencia de lo establecido en **1.1** está dado por la siguiente conjetura de Erdős: si $\{a_n\}_{n\in\mathbb{N}}$ es una sucesión estrictamente creciente de enteros positivos y

[1]La demostración de este hecho nos proporciona una interesante ilustración de la ubiquidad de la noción de número primo en Matemáticas. Claramente, es suficiente con demostrar que el conjunto de sucesiones finitas de números naturales es numerable. Denotemos con p_i al $i-$ésimo número primo y a la sucesión finita $a_1,\ldots,a_k$ asociemos el número natural determinado por $p_1^{a_1}\cdots p_k^{a_k}$. El teorema fundamental de la Aritmética nos permite garantizar que la correspondencia anterior determina una función biyectiva entre el conjunto de sucesiones finitas de números naturales y $\mathbb{N}$ y de aquí nuestro resultado.

además $\sum_{n\in\mathbb{N}} \frac{1}{a_n}$ diverge entonces la sucesión $\{a_n\}_{n\in\mathbb{N}}$ contiene progresiones aritméticas arbitrariamente largas. Dicha conjetura no ha sido establecida en el caso general todavía, no obstante, Ben J. Green y Terence Tao demostraron en 2004 que[2]:

Teorema 1.2. *Hay progresiones aritméticas arbitrariamente largas en* **P**.

El trabajo de Green y Tao es notable y representa un aporte grandioso a nuestro conocimiento de los números primos. Como la gran mayoría de los eventos importantes en Matemáticas, el teorema de Green-Tao no fue fruto de esfuerzos aislados. Uno de los resultados precursores de esa línea de investigación fue un célebre teorema de van der Waerden de 1927 que asegura que *dado $k \in \mathbb{N}$ siempre es posible encontrar un segmento inicial de los números naturales de tal manera que cada vez que se realice una partición de dicho segmento inicial en k clases al menos una de ellas retenga una progresión aritmética de longitud prescrita ℓ*. Fascinados por el problema de la determinación de cotas para las longitudes de los segmentos iniciales mencionados en el anterior teorema de van der Waerden, Erdős y Turán conjeturarían en 1936 que todo subconjunto $\mathcal{A}$ de $\mathbb{N}$ con densidad superior positiva[3] contiene progresiones aritméticas arbitrariamente grandes. K. F. Roth dió en 1952 el primer paso hacia el establecimiento de dicha conjetura al demostrar que todo $\mathcal{A} \subseteq \mathbb{N}$ con densidad superior positiva contiene progresiones aritméticas de longitud al menos 3. En 1969, E. Szemerédi demostró que, en

[2] B. Green; T. Tao. *The primes contain arbitrarily long arithmetic progressions.* Ann. of Math. **167** 2 (2008), págs. 481-547.

[3] Como es de esperarse, si S y T son subconjuntos de $\mathbb{N}$, la densidad superior de S relativa a T se define como

$$\limsup_{N\to\infty} \frac{|\{n \in S : n \leq N\}|}{|\{n \in T : n \leq N\}|}.$$

Cuando $T = \mathbb{N}$ únicamente se habla de densidad superior de S. La noción de densidad inferior de S relativa a T se define de modo análogo y cuando los límites para una densidad y otra coinciden se dice que S tiene densidad asintótica relativa a T igual al valor común de ambos límites.

realidad, siempre es posible encontrar progresiones aritméticas de longitud 4 en un $\mathcal{A}$ tal y seis años más tarde presentaría la respuesta afirmativa al caso general. El hecho de que el teorema de Green-Tao no fuera establecido sino hasta casi treinta años después nos proporciona de paso un importante metadato sobre la caprichosa naturaleza de los números primos: ¡su densidad asintótica es **cero**! Intuitivamente, esto indica que los primos se hacen cada vez más escasos en cuanto más se avanza en la sucesión de naturales. La prueba de este hecho la posponemos para el §**3** de este capítulo.

Ahora bien, aunque la conjetura de Erdős mencionada en un principio no ha sido probada en toda su generalidad aún, tampoco ha sido refutada. Atendiendo a la idea intuitiva que suele tenerse sobre el papel fundamental de la sucesión de primos entre los enteros lo que se esperaría es que la solución general a dicha propuesta de Erdős haga algún uso de las ideas desarrolladas en el establecimiento del caso primo. En todo caso, cualquiera que sea el desarrollo futuro de dicha problemática, es importante mencionar ahora una conexión del resultado de Green y Tao con una de las primeras interrogantes que suelen originarse en torno al tema de los números primos: ¿hay una fórmula *simple* que permita determinar al n-ésimo número primo?

La idea que se tiene al introducir el requerimiento de simplicidad es que la evaluación de la fórmula en cuestión sea menos laboriosa que la aplicación del método estándar (criba de Erastóstenes). En la actualidad no se conoce una fórmula óptima en dicho rubro y la mayoría de las que se tienen o son meras curiosidades que definen al n-ésimo primo en términos de sí mismo o lo determinan bajo el costo de más manipulaciones de las que se requieren con la criba de Eratóstenes, o bien, no son más que alguna reformulación de dicha criba. La necesidad de modificar nuestro enfoque se acaba de hacer más que patente. Consideremos entonces la pregunta: ¿se puede dar una función de n cuya imagen consista exclusivamente de números primos? Algunas fórmulas proporcionan *muchos* primos. Por

ejemplo, la famosa propuesta de Euler

$$p_1(n) = n^2 + n + 41$$

determina un primo para cada entero n en el intervalo $[0, 39]$. Más aún, si $n > 0$ se tiene que $p_1(-n) = (n-1)^2 + (n-1) + 41$ y de ahí que $p_1(n)$ sea un número primo para cada $n \in \mathbb{Z} \cap [-40, 39]$. De hecho, un importante resultado de la teoría establece que $q = 41$ es el número más grande tal que el trinomio $n^2 + n + q$ determina un número primo para cada entero n en $[-(q-1), q-2]$. Por otro lado, una notable observación de C. Goldbach (c. 1752) indica que ningún polinomio $f(x) \in \mathbb{Z}[x] \setminus \mathbb{Z}$ es tal que $f(n)$ es un número primo para cada entero n mayor o igual a algún n_0. La prueba es tan atractiva como la observación misma: supongamos que $f(n_0) = p \in \mathbf{P}$. Claramente, para $k \in \mathbb{Z}$ se tiene que

$$f(n_0 + pk) \equiv f(n_0) \text{ mód } p.$$

Luego, si $k \geq 0$ se sigue que $p | f(n_0 + pk)$ y por consiguiente $f(n_0 + pk) = \pm p$. Esto último implica que alguno de los polinomios en el conjunto

$$\{f(x) + p, f(x) - p\}$$

admite una infinidad de ceros. Esto último sólo puede ser posible si f es constante (contradicción).

Al parecer el nuevo enfoque que se le ha dado al cuestionamiento sobre la posibilidad de encontrar un patrón que nos permita representar primos de forma *simple* se encuentra lejos de retribuir algo. Es preciso notar, sin embargo, que el teorema de Green-Tao asegura que siempre es posible dar con polinomios de grado 1 de tal manera que la imagen de segmentos iniciales de $\mathbb{N}$ (de longitud predefinida ℓ) consista exclusivamente de números primos. ¡*Voilà*! Prácticamente el tipo de resultado al que aspirabamos, salvo por un detalle crucial: el teorema de Green-Tao es un resultado puramente de existencia, esto es, en su demostración no se proporciona un algoritmo explícito para determinar las dichosas progresiones aritméticas. Otro

punto en detrimento del teorema de Green-Tao es que lo deseable es tener una expresión que genere infinitos primos y lo más que el resultado de Green y Tao ofrece es la posibilidad de construir polinomios lineales cuyo rango intercepte bloques arbitrariamente grandes de **P**. La sospecha de que una condición tal debe implicar, *a fortiori*, la existencia de una progresión aritmética infinita de primos no tiene lugar y el contraejemplo no puede hacerse esperar: la sucesión de naturales de la forma $\{10^i + j\}$ donde $i \in \mathbb{N}$ y $j \in \{1, \ldots, i\}$ contiene progresiones aritméticas arbitrariamente grandes, pero falla en albergar una progresión aritmética de longitud infinita.

Una vez que se ha recabado evidencia a favor de la imposibilidad de tener una fórmula polinomial para generar primos, surge como opción natural el análisis de la representación de primos por medio de funciones racionales. El resultado básico en este respecto está dado por el siguiente

Teorema 1.3. Sea $R(x)$ una función racional. Si $R(n)$ es primo para cada $n \in \mathbb{Z}^+$ entonces $R(x)$ es constante.

Prueba. Supongamos que $R(x) = f(x)/g(x)$ donde f y g son polinomios coprimos de coeficientes enteros. Aplicando el algoritmo de la división a $f(x)$ y $g(x)$ resulta que podemos encontrar un $q \in \mathbb{Z} \setminus \{0\}$ tal que $qR(x) = G(x) + r(x)$ donde $G(x) \in \mathbb{Z}[x]$ y $r(x)$ es una función racional cuyo numerador tiene grado menor al grado de su denominador. Por hipótesis se tiene que $r(n)$ es siempre un número entero. Luego, al tenerse que $\lim_{x\to\infty} r(x) = 0$ se sigue que la función $r(x)$ es idénticamente igual a cero pues $r(n) = 0$ para todo n suficientemente grande y una función racional que no se anula idénticamente sólo puede ser 0 un número finito de veces. La igualdad $qR(x) = G(x)$ implica ahora que los números

$$R(0) = \frac{G(0)}{q}, R(1) = \frac{G(1)}{q}, R(2) = \frac{G(2)}{q}, R(3) = \frac{G(3)}{q}, \ldots$$

son todos primos. Dado que ninguno de estos primos se puede repetir más de $m = \deg[G(x)]$ veces se sigue que en el conjunto $\left\{ \dfrac{G(n)}{q} : n \in \mathbb{Z}^+ \right\}$ hay un número infinito de primos.

La deducción anterior nos permite asegurar la existencia de $a, b \in \mathbb{N}$ con $R(a) = p_1$, $R(b) = p_2$ y $p_2 > p_1 > m$. Así[4], si c es un número natural tal que $c \equiv a$ mód p_1 y $c \equiv b$ mód p_2 se sigue que

$$m!R(c) \equiv 0 \text{ mód } p_1p_2,$$

lo cual contradice la primalidad de $R(c)$.

□

Antes de seguir con nuestros intentos de encontrar algún dejo de estructura en **P**, vale la pena hacer unos comentarios sobre una fórmula para primos que en su momento originó sensación. Pierre de Fermat creía que la expresión $F_n = 2^{2^n} + 1$ devolvía un número primo siempre que n se tomaba en los enteros no negativos. Las cinco primeras evaluaciones de la fórmula de Fermat dan

$$F_0 = 3, \qquad F_1 = 5, \qquad F_2 = 17, \qquad F_3 = 257 \qquad \text{y} \qquad F_4 = 65537$$

[4] **I**. La posibilidad de elegir tal número c es garantizada por el teorema chino del resto. **II**. El polinomio $H(x) = m!R(x)$ tiene coeficientes enteros. En efecto, si con $\binom{x}{n}$ denotamos al polinomio $\dfrac{x(x-1)\cdots(x-n+1)}{n!}$ se sigue que

$$R(x) = b_m + b_{m-1}\binom{x}{1} + b_{m-2}\binom{x}{2} + \ldots + b_0\binom{x}{m}$$

donde

$$\begin{aligned} R(0) &= b_m, \\ R(1) &= b_m + \binom{1}{1}b_{m-1}, \\ R(2) &= b_m + \binom{2}{1}b_{m-1} + \binom{2}{2}b_{m-2}, \\ &\cdots \\ R(m) &= b_m + \binom{m}{1}b_{m-1} + \ldots + \binom{m}{m}b_0. \end{aligned}$$

Dado que los números $R(0), \ldots, R(m)$ son enteros se tiene que los coeficientes $b_0, \ldots, b_m$ también lo son. La conclusión deseada es ahora una consecuencia directa de este hecho.

y todas ellas corresponden a números primos. Sin embargo, Euler mostró en 1732 que $F_5 = 2^{2^5} + 1$ es compuesto pues $641 = 5^4 + 2^4 = 5 \cdot 2^7 + 1$ divide tanto a $5^4 \cdot 2^{28} + 2^{32}$ como a $5^4 \cdot 2^{28} - 1$ y por tanto divide a la diferencia de ambos números, la cual es exactamente F_5. A la fecha, no se ha encontrado otro primo en la lista de evaluaciones de la fórmula de Fermat; luego, su conjetura ha resultado más bien desafortunada. Paradójicamente, los números en la sucesión $\{F_n\}_{n\in\mathbb{N}}$, bautizados a la postre como números de Fermat, no reducen su existencia a lo relacionado con esta conjetura de Fermat. Un notable resultado de Gauss y P. Wantzel asegura que el polígono regular de n lados puede construirse con regla y compás si y sólo si n es de la forma $2^h p_1 \cdots p_k$ donde h es un entero no negativo y los números p_i son números de Fermat primos. Otra interesante aparición de los números de Fermat se produce en una demostración de la infinitud de primos debida a Goldbach. La idea de dicha prueba se puede poner en medio párrafo y los detalles son fácilmente verificables: si $m > n$ entonces $F_n | (F_m - 2)$ y por tanto los números de Fermat son primos relativos dos a dos. Esto indica que cada número de Fermat aporta a **P** un factor primo distinto al que aportan los demás números de Fermat y de aquí nuestro QED. La prueba anterior suele atribuirse, erróneamente, a G. Pólya[5].

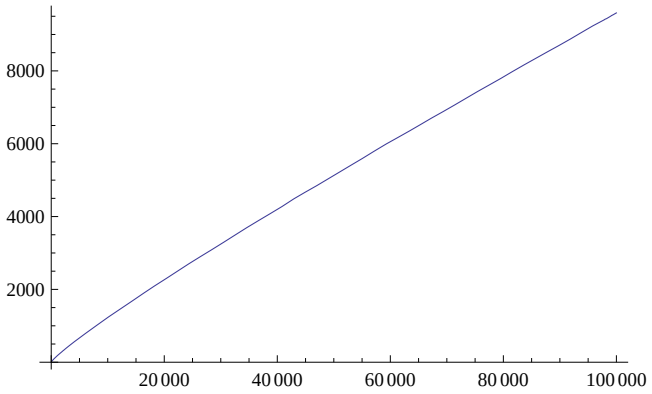
Dado que los resultados obtenidos al haber contemplado las dos preguntas anteriores han sido prácticamente nulos, daremos ahora un giro dramático al sentido de nuestra investigación. La nueva pregunta que vamos a considerar es: dado un número $x > 1$, ¿cuántos números primos hay en el intervalo $[1, x]$?

Denotemos, como es usual, con $\pi(x)$ al número de primos que son menores o iguales al número x. La función π así determinada suele ir bajo la denominación de **función contadora de primos**. Notemos que si p_n es el n-ésimo número primo entonces $\pi(p_n) = n$ y por tanto, vistas como funciones de una variable en $\mathbb{N}$, la función contadora de primos es inversa

[5]La prueba de la coprimalidad de números de Fermat distintos aparece ya en una carta de Golbach a Euler con fecha 20 de julio de 1730.

izquierda de p_n. Luego, el preguntar por una fórmula *simple* para $\pi(x)$ es prácticamente regresar a la primer interrogante de párrafos arriba. Así, la pregunta que recién se ha formulado la interpretaremos mejor como una solicitud de *una fórmula para el número aproximado de números primos en el intervalo* $[1, x]$.

En la ilustración siguiente aparece la gráfica de la función contadora de primos en el intervalo $[1, 100000]$:



La regularidad del comportamiento de la función $\pi(x)$, que en la gráfica de arriba ha quedado de manifiesto, es un indicio de que finalmente nuestro análisis ha sido encauzado por una ruta promisoria. Siguiendo a Zagier mostraremos ahora que es relativamente fácil encontrar una fórmula empírica que de una buena descripción del crecimiento de la función contadora de primos.

Hay 25 números primos menores o iguales a 100, esto es, una cuarta parte del total de números en el intervalo; debajo de 1000 hay 168 números primos o aproximadamente un sexto del total en el rango; debajo de 10000 hay 1229 números primos o aproximadamente un octavo de los números en el intervalo respectivo. En la tabla de la página siguiente se concentran los resultados que se obtienen al continuar con estos cálculos hasta alcanzar la undécima potencia de 10.

Los números listados en la tercer columna de la tabla nos indican que los cocientes $x/\pi(x)$ aumentan aproximadamente en 2.3 cuando pasamos

x	$\pi(x)$	$x/\pi(x)$
10	4	2.5
100	25	4.0
1000	168	6.0
10,000	1,229	8.1
100,000	9,592	10.4
1,000,000	78,498	12.7
10,000,000	664,579	15.0
100,000,000	5,761,455	17.4
1000,000,000	50,847,534	19.7
10,000,000,000	455,052,511	22.0
100,000,000,000	4,118,054,813	24.3

de una potencia de 10 a la siguiente. Dado que $2.3 \approx \log 10$, una conjetura que en este punto resulta natural proponer es que

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

La relación asintótica anterior, la cual no fue demostrada sino hasta 1896, es conocida en la actualidad como Teorema del Número Primo.

El objetivo primordial de este escrito será presentar un par de demostraciones para este importante resultado así como dar un recuento de algunos de los trabajos que influyeron en el establecimiento de tan notable teorema.

§**2.** Hemos visto ya que el conjunto de números primos es infinito. Una manera de ver que la sucesión de distancias entre primos consecutivos no está acotada es como sigue: si $k \in \mathbb{N}$ definamos como N al producto de todos los números primos que son menores que $k+2$. Dado que cada $i \in \{2, 3, \ldots, k+1\}$ tiene un divisor primo en común con N se sigue que ninguno de los números en $\{N+i : 2 \leq i \leq k+1\}$ es primo. Como una ilustración particular del resultado anterior tenemos, por ejemplo, que

cada uno de los cincuenta números siguientes es compuesto:

614889782588491412	614889782588491413	614889782588491414
614889782588491415	614889782588491416	614889782588491417
614889782588491418	614889782588491419	614889782588491420
614889782588491421	614889782588491422	614889782588491423
614889782588491424	614889782588491425	614889782588491426
614889782588491427	614889782588491428	614889782588491429
614889782588491430	614889782588491431	614889782588491432
614889782588491433	614889782588491434	614889782588491435
614889782588491436	614889782588491437	614889782588491438
614889782588491439	614889782588491440	614889782588491441
614889782588491442	614889782588491443	614889782588491444
614889782588491445	614889782588491446	614889782588491447
614889782588491448	614889782588491449	614889782588491450
614889782588491451	614889782588491452	614889782588491453
614889782588491454	614889782588491455	614889782588491456
614889782588491457	614889782588491458	614889782588491459
	614889782588491460	614889782588491461

Evidentemente, la receta permite crear bloques arbitrariamente grandes de números compuestos consecutivos. Aún con eso, es posible determinar cotas *locales* para el tamaño de los huecos en el conjunto de los números primos. Una de las cotas más famosas al respecto indica básicamente que *la distancia hacia el próximo número primo no puede ser más grande que el número a partir del cual se empieza a buscar*. El resultado es conocido como postulado de Bertrand pues fue el matemático francés J. Bertrand quien en 1845 lo conjeturó y verificó empíricamente para todos los naturales menores que tres millones.

La primera demostración del postulado de Bertrand fue publicada por

P. Chebyshev en 1850. Srinivasa Ramanujan proporcionaría en 1919 una prueba más simple. La demostración que a continuación se presenta se retoma del primer artículo en la prolífica carrera de Erdős.

Proposición 1.4. (Postulado de Bertrand) Sea $n \in \mathbb{N}$. Siempre hay un número primo en el intervalo $(n, 2n]$.

Prueba. La demostración la llevaremos a cabo en cuatro pasos.

1. En este punto vamos a establecer una notable desigualdad sobre el producto de primos en un rango, a saber,

$$\prod_{\substack{p \leq n, \\ p \in \mathbf{P}}} p < 4^{n-1}. \tag{1.5}$$

La desigualdad es claramente cierta para $n = 2$ y $n = 3$. Probaremos entonces que también se tiene para $n+1$ siempre que sea cierta para todos los naturales menores o iguales a n. Si $n+1$ es par entonces no hay nada que probar pues

$$\prod_{\substack{p \leq n+1, \\ p \in \mathbf{P}}} p = \prod_{\substack{p \leq n, \\ p \in \mathbf{P}}} p < 4^{n-1} < 4^n.$$

Supongamos entonces que $n+1$ es igual $2m+1$ para algún natural m mayor que 1. En tal caso se tiene que

$$\prod_{\substack{p \leq 2m+1, \\ p \in \mathbf{P}}} p = \left(\prod_{\substack{p \leq m+1, \\ p \in \mathbf{P}}} p \right) \left(\prod_{\substack{m+1 < p \leq 2m+1, \\ p \in \mathbf{P}}} p \right) < 4^m \binom{2m+1}{m} \leq 4^m 2^{2m} = 4^{2m}.$$

Todas las desigualdades en la línea anterior son fáciles de justificar. De hecho,

$$\prod_{\substack{p \leq m+1, \\ p \in \mathbf{P}}} p < 4^m$$

se cumple en virtud del paso de inducción. La desigualdad

$$\prod_{\substack{m+1<p\leq 2m+1,\\ p\in\mathbf{P}}} p \leq \binom{2m+1}{m}$$

es consecuencia de que el coeficiente binomial $\binom{2m+1}{m}$ es un múltiplo de cada uno de los primos en el intervalo $(m+1, 2m+1]$. Finalmente, la desigualdad

$$\binom{2m+1}{m} \leq 2^{2m} \tag{1.6}$$

se sigue de la observación de que tanto $\binom{2m+1}{m}$ como $\binom{2m+1}{m+1}$ son sumandos que aparecen en la expresión del lado izquierdo de la identidad

$$\sum_{k=0}^{2m+1} \binom{2m+1}{k} = 2^{2m+1}.$$

2. De la fórmula de Polignac[6] se tiene que el exponente del primo p en

[6]El exponente del primo p en la factorización canónica de $n!$ es

$$\sum_{k\geq 1} \left\lfloor \frac{n}{p^k} \right\rfloor.$$

Vamos a probar la validez de la fórmula haciendo inducción sobre n. La relación es cierta cuando $n = 1$. Supongamos que también vale para n y hagamos $n+1 = p^u m$, donde p no divide a m. De la hipótesis de inducción se sigue que el exponente de p en la factorización de $(n+1)!$ es

$$\sum_{k=1}^{u} \left(\left\lfloor \frac{n}{p^k} \right\rfloor + 1 \right) + \sum_{k>u} \left\lfloor \frac{n}{p^k} \right\rfloor.$$

La conclusión se tiene ahora al notar que $\left\lfloor \frac{n+1}{p^k} \right\rfloor = \left\lfloor \frac{n}{p^k} \right\rfloor + 1$ cuando $1 \leq k \leq u$ y $\left\lfloor \frac{n+1}{p^k} \right\rfloor = \left\lfloor \frac{n}{p^k} \right\rfloor$ cuando $k > u$.

la factorización de $\binom{2n}{n} = \frac{(2n)!}{n!n!}$ es

$$\sum_{k\geq 1}\left(\left\lfloor \frac{2n}{p^k}\right\rfloor - 2\left\lfloor \frac{n}{p^k}\right\rfloor\right). \tag{1.7}$$

Ahora bien, dado que

$$\left\lfloor \frac{2n}{p^k}\right\rfloor - 2\left\lfloor \frac{n}{p^k}\right\rfloor < \frac{2n}{p^k} - 2\left(\frac{n}{p^k} - 1\right) = 2$$

se concluye que cada uno de los sumandos de la serie en (1.7) es a lo más 1. Luego, la descomposición en primos de $\binom{2n}{n}$ contiene al factor p a lo sumo α_p veces donde $\alpha_p = \max\{k : p^k \leq 2n\}$. De esto último se obtiene, en particular, que los primos mayores que $\sqrt{2n}$ aparecen a los más una vez como factores de $\binom{2n}{n}$.

Mostremos ahora que para $n \geq 3$ los primos que satisfacen $\frac{2}{3}n < p \leq n$ no dividen a $\binom{2n}{n}$. En efecto, la desigualdad $3p > 2n$ implica que $p > 2$ y que p y $2p$ son los únicos múltiplos de p que aparecen en el numerador de $\frac{(2n)!}{n!n!} = \binom{2n}{n}$. Como $2p > \frac{4}{3}n$, el número de apariciones de p en el denominador de la fracción anterior es también 2. Luego, al hacer las cancelaciones correspondientes se concluye que el primo p ya no permanece en la descomposición de $\binom{2n}{n}$, tal como deseabamos establecer.

3. A continuación procedemos a efectuar algunas estimaciones relacionadas con el coeficiente binomial $\binom{2n}{n}$. Para $n \geq 3$ se tiene[7], según lo

[7]La desigualdad $\frac{4^n}{2n} \leq \binom{2n}{n}$ es bastante fácil de obtener. En la lista,

$$\binom{2n}{0} + \binom{2n}{2n}, \binom{2n}{1}, \ldots, \binom{2n}{2n-1}$$

visto en el paso **2**, que

$$\frac{4^n}{2n} \leq \binom{2n}{n} \leq \left(\prod_{p\leq \sqrt{2n}} 2n\right)\left(\prod_{\sqrt{2n}<p\leq \frac{2}{3}n} p\right)\left(\prod_{n<p\leq 2n} p\right).$$

Dado que el número de primos en el intervalo $[1, \sqrt{2n}]$ no excede a $\sqrt{2n}$, las desigualdades anteriores nos permiten asegurar que

$$4^n \leq (2n)^{1+\sqrt{2n}}\left(\prod_{\sqrt{2n}<p\leq \frac{2}{3}n} p\right)\left(\prod_{n<p\leq 2n} p\right) \tag{1.8}$$

cuando $n \geq 3$.

4. Supongamos ahora que n es un número natural tal que el intervalo $(n, 2n]$ no contiene primo alguno. Esto implica que el tercer factor en el miembro derecho de (1.8) es igual a 1. La desigualdad en (1.5) indica entonces que

$$4^n < (2n)^{1+\sqrt{2n}} \cdot 4^{\frac{2}{3}n}$$

o bien

$$4^{\frac{1}{3}n} < (2n)^{1+\sqrt{2n}}. \tag{1.9}$$

Vamos a tratar de obtener información adicional sobre n en base a la desigualdad en (1.9). La aplicación de una instancia de la conocida desigualdad de Bernoulli[8] nos permite asegurar que

$$2n = (\sqrt[6]{2n})^6 < (\lfloor \sqrt[6]{2n} \rfloor + 1)^6 \leq 2^{6\lfloor \sqrt[6]{2n} \rfloor} \leq 2^{6\sqrt[6]{2n}}. \tag{1.10}$$

---

el término más grande es $\binom{2n}{n}$. Luego,

$$(2n)\binom{2n}{n} \geq \left[\binom{2n}{0} + \binom{2n}{2n}\right] + \binom{2n}{1} + \ldots + \binom{2n}{2n-1} = 2^{2n} = 4^n.$$

[8]Dicha desigualdad asegura que $(1+x)^k \geq 1 + kx$ siempre que $x > -1$ y $k \in \mathbb{N}$.

Si suponemos que el natural n es tal que $18 < 2\sqrt{2n}$, las desigualdades en (1.9) y (1.10) nos proporcionan que

$$2^{2n} < (2n)^{3(1+\sqrt{2n})} < 2^{\sqrt[6]{2n}(18+18\sqrt{2n})} < 2^{20\sqrt[6]{2n}\sqrt{2n}} = 2^{20(2n)^{2/3}}.$$

De esto último se sigue que $(2n)^{1/3} < 20$ y por tanto n tiene que ser menor que 4000.

El establecimiento de la proposición se ha reducido entonces a analizar los intervalos $(n, 2n]$ donde n es un natural menor que 4000. Claramente, ello no significa que la proposición deba irse probando caso a caso entre todos los naturales n menores que 4000. De hecho, es suficiente con notar que

$$\mathcal{L} = \{2, 3, 5, 7, 13, 23, 43, 83, 163, 317, 631, 1259, 2503, 4001\}$$

es una lista de números primos donde cada uno es menor que dos veces el anterior. En efecto, dado $n \in [1, 4000]$, sea N igual al mayor de los primos en $\mathcal{L}$ que es menor o igual a n. Se cumple entonces que el primo siguiente a N en $\mathcal{L}$, digamos N', es mayor que n y además $N' \leq 2n$. Así , $N' \in (n, 2n]$ y el resultado se sigue.

□

Es natural preguntarse en este momento el porqué la afirmación recién probada ha pasado a la historia bajo la denominación de *postulado* si en realidad antes de que se le demostrara no se trataba más que de una simple conjetura aritmética. W. J. LeVeque expresa en su célebre *Fundamentals of Number Theory* que el origen de dicha denominación se encuentra en el hecho que J. Bertrand diera por cierto el resultado e incluso lo utilizara en sus investigaciones en torno a un problema de Teoría de Grupos. LeVeque no da más detalles sobre el problema en que Bertrand trabajara y en una primera lectura la conexión podría resultar inesperada. Con el fin de mostrar que las interacciones del postulado de Bertrand con el Álgebra son una realidad presentamos ahora mismo un interesante

Ejemplo 1.5. Sea p un número primo impar. ¿Será cierto que el polinomio $p(x) = x^p + (p-1)!$ es irreducible en $\mathbb{Q}[x]$?

Al variar p entre los primeros números primos, intuimos la posible respuesta (afirmativa) al caso general. La demostración correspondiente es como sigue: sea q el mayor primo menor que p. Dado que q divide a los coeficientes de los términos cuyo grado varía entre $p-1$ y 0, el criterio de Eisenstein transfiere el estudio de la irreducibilidad de $p(x)$ a una cuestión bastante concreta: ¿es $(p-1)!$ un múltiplo de q^2? La negativa a dicha interrogativa implicaría de inmediato la irreducibilidad de $p(x)$ en $\mathbb{Q}[x]$.

Consideremos los factores de $(p-1)! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (p-1)$. Claramente, $1 \cdot 2 \cdot \ldots \cdot (q-1)$ no es divisible por q. Resta entonces buscar múltiplos de q en el producto $(q+1)(q+2)\ldots(p-1)$. Si hubiera un múltiplo de q en el producto anterior entonces debería ser el caso que $2q \leq p-1$. Por otro lado, del postulado de Bertrand se tendría la existencia de un primo p' con $p' \in (q, 2q]$. Luego, p' sería tal que $q < p' \leq 2q \leq p-1 < p$, lo cual entra en contradicción con la maximalidad de q. De todo lo anterior se sigue que $q \| (p-1)!$ y de ahí la irreducibilidad de $p(x)$ en $\mathbb{Q}[x]$.

□

Una aplicación del postulado de Bertrand, un tanto más cercana al objeto de nuestro estudio, tiene que ver con la temática de las fórmulas generadoras de primos que ya discutieramos al final de §**1**. Bertrand implica que si n es un número compuesto mayor que 2 entonces siempre hay un número primo en el intervalo $(n, 2n-2)$. Esto permite asegurar la existencia de una sucesión $\{q_1, q_2, \ldots\} \subseteq \mathbf{P}$ que satisface, para cada $n \in \mathbb{N}$, las siguientes desigualdades

$$2^{q_n} < q_{n+1} < 2^{q_n+1} - 1. \tag{1.11}$$

De (1.11) se sigue ahora que la sucesión $\{q_1, q_2, \ldots\}$ es tal que

$$\log_2^1 q_1 < \log_2^2 q_2 < \ldots < \log_2^n q_n < \ldots \tag{1.12}$$

y

$$\log_2^1(q_1+1) > \log_2^2(q_2+1) > \ldots > \log_2^{n+1}(q_{n+1}+1) > \ldots$$

Luego, las desigualdades en la línea anterior nos permiten asegurar que la sucesión

$$\{\log_2^1 q_1, \log_2^2 q_2, \ldots, \log_2^n q_n, \ldots\}$$

es estrictamente creciente. Por otra parte, si n es un natural cualquiera se tiene que $\log_2^n q_n < \log_2^n(q_n+1) < \ldots < \log_2^1(q_1+1)$. Esto último implica que

$$\{\log_2^1 q_1, \log_2^2 q_2, \ldots, \log_2^n q_n, \ldots\}$$

es una sucesión acotada superiormente. Al ser además creciente se concluye que la sucesión converge. Supongamos que su límite es μ. Denotemos a partir de este momento a $\log_2^n q_n$ con a_n y a $\log_2^n(q_n+1)$ con b_n. Si existiera $N \in \mathbb{N}$ con $a_N \geq \mu$ entonces para cada $k \in \mathbb{N}$ se tendría que

$$|a_{N+k} - \mu| = |(a_{N+k} - a_N) + (a_N - \mu)| \geq a_{N+k} - a_N \geq a_{N+1} - a_N$$

y de aquí que

$$\lim_{k\to\infty} |a_{N+k} - \mu| \geq a_{N+1} - a_N > 0$$

lo cual entra en contradicción con el hecho de que $a_n \to \mu$ cuando $n \to \infty$. Así, debe ser el caso que $a_n < \mu$ para todo $n \in \mathbb{N}$. Ahora bien, la existencia de b_N tal que $b_N \leq \mu$ implica que para todo $k \in \mathbb{N}$

$$|\mu - a_{N+k}| = |(\mu - b_N) + (b_N - a_{N+k})| \geq b_N - a_{N+k} > b_N - b_{N+1}.$$

Esto último contradice la convergencia de la sucesión $\{a_n\}_{n\in\mathbb{N}}$ y de ahí que la desigualdad $b_n > \mu$ resulte justa para cada $n \in \mathbb{N}$. De todo lo hecho en los párrafos anteriores se desprende ahora el notable

Teorema 1.6. Todos los elementos de la sucesión $\left\{\lfloor 2^\mu \rfloor, \left\lfloor 2^{2^\mu} \right\rfloor, \left\lfloor 2^{2^{2^\mu}} \right\rfloor, \ldots\right\}$ son primos.

Prueba. Se ha probado en la discusión preliminar que para cada $n \in \mathbb{N}$ las desigualdades

$$\log_2^n q_n < \mu < \log_2^n (q_n + 1)$$

son ciertas. Ahora bien, si denotamos con $f(x)$ a la función inversa de $\log_2(x)$ se cumple que

$$q_n < f^n(\mu) < q_n + 1$$

y de aquí que $\lfloor f^n(\mu) \rfloor = q_n$ para cada $n \in \mathbb{N}$. El resultado se sigue ahora al notar que

$$f^1(\mu) = 2^\mu, \quad f^2(\mu) = 2^{2^\mu}, \quad f^3(\mu) = 2^{2^{2^\mu}}, \ldots$$

□

El teorema anterior proporciona un ejemplo concreto de una fórmula para generar infinitos primos. La fórmula fue propuesta por E. M. Wright alrededor de 1951 (cf. E. M. Wright. *A prime representing function.* Amer. Math. Monthly **58** (1951), págs. 617-618.). El mismo Wright probaría algunos años después que los valores admisibles para la variable μ en el teorema recién probado forman un subconjunto nunca denso de los reales, de cardinal $2^{\aleph_0}$ y medida cero. La opinión generalizada con respecto a la fórmula de Wright es que no es un buen ejemplo de lo que se espera al solicitar una fórmula para representar primos. Los dos incovenientes principales que se le encuentran a una fórmula así son: primero, la expresión depende de un número μ que se define de modo indirecto y segundo, la fórmula no depende de propiedades intrínsecas a **P**. De hecho, no es díficil presentar fórmulas análogas para otros subconjuntos de $\mathbb{N}$ con una propiedad técnica similar al postulado de Bertrand.

§**3.** En este apartado vamos a derivar los primeros resultados puntuales relativos a la función $\pi(x)$. Nuestro primer teorema contempla, entre otras cosas, la prueba de un hecho central comentado ya en el §**1**:

Teorema 1.7.

a) La función $\pi(x)$ es $o(x)$.

b) La sucesión $\left\{\dfrac{n}{\pi(n)}\right\}_{n\in\mathbb{N}_{\geq 2}}$ contiene al conjunto de números naturales mayores que 1. En particular, existen infinitos números naturales n tales que $\pi(n)|n$.

Prueba. La demostración del inciso a depende fuertemente de la desigualdad en (1.5). Sea k un número natural mayor que 1. Para x suficientemente grande se tiene, en la luz de (1.5), que

$$x\log 4 > \log \prod_{\substack{p\leq x,\\ p\in\mathbf{P}}} p \geq \sum_{\substack{k<p\leq x,\\ p\in\mathbf{P}}} \log p \geq (\pi(x)-\pi(k))\log k$$

y por consiguiente

$$\frac{\pi(x)}{x} < \frac{\pi(k)}{x} + \frac{\log 4}{\log k}.$$

Puesto que ambos términos en el lado derecho de esta desigualdad van a 0 cuando $x\to\infty$ se sigue que $\lim\limits_{x\to\infty}\dfrac{\pi(x)}{x} = 0$, tal como queríamos probar.

Resta entonces dar la prueba de lo aseverado en b. Sea $m\in\mathbb{N}\setminus\{1\}$ y

$$X = \left\{n\in\mathbb{N} : \frac{\pi(mn)}{mn} \geq \frac{1}{m}\right\}.$$

Este conjunto X es claramente no vacío. Además, como $\pi(mn)/mn\to 0$ cuando $n\to\infty$, se tiene que X está acotado. Por lo tanto X posee un elemento maximal, digamos k.

Si $\dfrac{\pi(mk)}{mk} = \dfrac{1}{m}$ se sigue que m es un término de la sucesión $\left\{\dfrac{n}{\pi(n)}\right\}_{n\in\mathbb{N}}$. En caso contrario, se tendría que

$$\pi(m(k+1)) \geq \pi(mk) \geq (k+1),$$

lo cual implica que $(k+1)\in X$. Esta relación contradice la maximalidad del elemento k y de aquí el resultado. □

La prueba de la parte a es, sin lugar a dudas, la demostración más breve que se puede tener de dicho resultado. La afimación en b es notable y más aún la técnica empleada en su prueba. Es evidente que en dicha demostración la función contadora de primos puede reemplazarse por cualquier sucesión creciente $\{a_n\}_{n\in\mathbb{N}} \subseteq \mathbb{N}$ tal que $a_n = o(n)$.

Es importante mencionar en este momento una notable consecuencia de la proposición **1.1** y del teorema **1.7**: la función $\pi(x)$ no puede ser racional:

Si $\pi(x) = \dfrac{p(x)}{q(x)}$ entonces

$$+\infty = \lim_{x\to\infty} \pi(x) = \lim_{x\to\infty} \frac{p(x)}{q(x)}$$

y de aquí que $\deg[p(x)] \geq \deg[q(x)] + 1$. Por otro lado, al ser

$$0 = \lim_{x\to\infty} \frac{\pi(x)}{x} = \lim_{x\to\infty} \frac{p(x)}{xq(x)}$$

se sigue que $\deg[p(x)] \leq \deg[q(x)]$ y por tanto

$$\deg[p(x)] \leq \deg[q(x)] < \deg[q(x)] + 1 \leq \deg[p(x)],$$

lo cual es decididamente absurdo.

Inspirados por los términos logarítmicos que han aparecido en la verificación de la primera parte del teorema **1.7**, procedemos a considerar la función auxiliar T definida del modo siguiente

$$T(x) = \sum_{n\leq x} \log n.$$

Parecería, en un principio, que al introducir la función T, la discusión abandona el terreno de los números primos. No obstante, vamos a ver ahora que esto no es así. Sabemos que $\pi(x)$ es la función suma de la aplicación característica de $\mathbf{P}$ evaluada en x. Intentar obtener información sobre $\pi(x)$ directamente de la definición anterior es difícil y por eso, a la hora de los cálculos, suelen considerarse las funciones

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

y

$$\psi(x) = \sum_{p \leq x} a_p \log p \tag{1.13}$$

donde $\vartheta(x)$ no es más que una versión ponderada de $\pi(x)$ y los a_p en (1.13) son tales que p^{a_p} es la mayor potencia del primo p que es menor o igual a x. Las relaciones centrales entre las funciones T, ϑ y ψ se condensan en la

Proposición 1.8. Para todo $x \in \mathbb{R}$ se cumple que

$$T(x) = \psi(x) + \psi(x/2) + \psi(x/3) + \ldots \quad \text{y} \quad \psi(x) = \vartheta(x) + \vartheta(x^{1/2}) + \vartheta(x^{1/3}) + \ldots$$

Prueba. La idea de ambas pruebas es la misma: hay que determinar, para cada primo p menor o igual a n, el coeficiente del término $\log p$ en cada miembro de la potencial identidad.

Vamos a comenzar con la *identidad* de más a la derecha. La definición de la función ψ indica que el coeficiente del término $\log p$ es igual a a_p donde p^{a_p} es la mayor potencia del primo p que es menor o igual a x, esto es, a_p es un número entero que satisface

$$p^{a_p} \leq x < p^{a_p+1}.$$

Al tomar logaritmo en base en p en las desigualdades previas se llega a que $a_p \leq \log_p x < a_p + 1$ y de ahí que $a_p = \lfloor \log_p x \rfloor$. Por otra parte, el coeficiente del término $\log p$ en la serie $\vartheta(x) + \vartheta(x^{1/2}) + \vartheta(x^{1/3}) + \ldots$ es

$$\sum_{i \in \mathbb{N}} [x^{1/i} \geq p]$$

donde [] son los corchetes de Iverson[9]. Se cumple entonces que

$$\left(\sum_{i \in \mathbb{N}} [x^{1/i} \geq p] \right) = k$$

[9]La expresión $[P]$ es igual a 1 cuando P es cierta y 0 en otro caso. Veáse, por ejemplo, el capítulo 2 del *Concrete Mathematics* de R. L. Graham, D. E. Knuth y O. Patashnik.

donde $k = \text{máx}\{i \in \mathbb{N} : x^{1/i} \geq p\}$. Lo anterior indica que k es un entero que cumple con el par de restricciones siguientes: $x^{1/k} \geq p$ y $x^{1/(k+1)} < p$. Así, k es tal que $k \leq \log_p x < (k+1)$ y por tanto $k = \lfloor \log_p x \rfloor = a_p$, como deseabamos obtener.

Para la obtención de la identidad restante empezamos por notar que $T(x) = \sum_{n \leq x} \log n = \log \lfloor x \rfloor !$ La fórmula de Polignac (ver pie de página 6) nos permite asegurar ahora que $T(x)$ puede ponerse en la forma

$$\sum_{p \leq x} c_p \log p$$

donde $c_p = \sum_{i \in \mathbb{N}} \left\lfloor \frac{\lfloor x \rfloor}{p^i} \right\rfloor$. Determinemos ahora el coeficiente de $\log p$ en la expresión $S(x) = \psi(x) + \psi(x/2) + \psi(x/3) + \ldots$ El coeficiente de $\log p$ en la suma correspondiente a $\psi(x/i)$ es $\left\lfloor \log_p \frac{x}{i} \right\rfloor$. Luego, el coeficiente con el que $\log p$ aparece en la suma que define a $S(x)$ es igual a

$$\sum_{i \in \mathbb{N}} \left\lfloor \log_p \frac{x}{i} \right\rfloor .$$

Claramente, el número de sumandos distintos de cero en la serie anterior es igual a $\left\lfloor \frac{x}{p} \right\rfloor$. Más aún, la identidad

$$\left\lfloor \log_p \frac{x}{i} \right\rfloor - 1 = \left\lfloor \left(\log_p \frac{x}{i} \right) - 1 \right\rfloor = \left\lfloor \log_p \frac{x}{p \cdot i} \right\rfloor$$

nos permite asegurar que

$$\sum_{i \in \mathbb{N}} \left\lfloor \log_p \frac{x}{i} \right\rfloor = \sum_{i \in \mathbb{N}} \left(1 + \left\lfloor \log_p \frac{x}{p \cdot i} \right\rfloor \right) = \left\lfloor \frac{x}{p} \right\rfloor + \sum_{i \in \mathbb{N}} \left\lfloor \log_p \frac{x}{p \cdot i} \right\rfloor$$

y de aquí que el coeficiente de $\log p$ en $S(x)$ es $\sum_{i \in \mathbb{N}} \left\lfloor \frac{x}{p^i} \right\rfloor$. El resultado se sigue ahora al notar que para cada natural P se cumple que[10] $\left\lfloor \frac{\lfloor x \rfloor}{P} \right\rfloor = \left\lfloor \frac{x}{P} \right\rfloor$. □

[10]La identidad en cuestión es bien conocida y se demuestra como sigue: x es igual a

Hagamos $V(x) = T(x)+T(x/30)-T(x/2)-T(x/3)-T(x/5)$. De la proposición **1.8** se obtiene que

$$\begin{aligned} V(x) &= \psi(x)+\psi\left(\frac{x}{2}\right)+\ldots+\psi\left(\frac{x}{30}\right)+\psi\left(\frac{x}{2\cdot 30}\right)+\ldots \\ &- \psi\left(\frac{x}{2}\right)-\psi\left(\frac{x}{2\cdot 2}\right)-\ldots-\psi\left(\frac{x}{3}\right)-\psi\left(\frac{x}{2\cdot 3}\right)-\ldots-\psi\left(\frac{x}{5}\right)-\psi\left(\frac{x}{2\cdot 5}\right)-\ldots \end{aligned}$$

Vamos a determinar ahora el coeficiente exacto, A_k, del término $\psi(x/k)$ en la serie del lado derecho de la igualdad anterior.

Sea t_k el coeficiente que corresponde a $\psi(x/k)$ en el desarrollo en términos de ψ de $T(x/i)$. Se cumple entonces que $A_k = 1-t_2-t_3-t_5+t_{30}$. Además, t_k es 1 ó 0 dependiendo de si i divide a k o no. Por tanto, $t_k = \lfloor 2^{\lfloor k/i\rfloor - k/i}\rfloor$ y de aquí que

$$A_k = 1-\left\lfloor 2^{\lfloor k/2\rfloor-(k/2)}\right\rfloor-\left\lfloor 2^{\lfloor k/3\rfloor-(k/3)}\right\rfloor-\left\lfloor 2^{\lfloor k/5\rfloor-(k/5)}\right\rfloor+\left\lfloor 2^{\lfloor k/30\rfloor-(k/30)}\right\rfloor.$$

De la expresión recién obtenida para A_k resulta claro que la sucesión $\{A_k\}_{k\in\mathbb{N}}$ tiene período 30 y que su valor en un natural dado sólo depende de la clase módulo 30 a la que este pertenece. Los primeros 30 términos de la sucesión son como sigue:

$$\begin{aligned} &1,0,0,0,0,-1,1,0,0,-1,1,-1,1,0,-1, \\ &0,1,-1,1,-1,0,0,1,-1,0,0,0,0,1,-1. \end{aligned}$$

Luego, la serie que define a $V(x)$ es alternante y sus primeros términos son:

$$V(x) = \psi(x)-\psi\left(\frac{x}{6}\right)+\psi\left(\frac{x}{7}\right)-\psi\left(\frac{x}{10}\right)+\psi\left(\frac{x}{11}\right)-\psi\left(\frac{x}{12}\right)+\ldots$$

La monotonía de ψ implica entonces que la función V es tal que

$$\psi(x)-\psi\left(\frac{x}{6}\right)\leq V(x)\leq\psi(x). \tag{1.14}$$

$\lfloor x\rfloor + Q$ para algún $Q\in[0,1)$ y el algoritmo de la división en $\mathbb{Z}$ garantiza la existencia de $(q,r)\in\mathbb{Z}^2$ tal que $\lfloor x\rfloor = qP+r$ donde $0\leq r<P$. Así, $\left\lfloor\frac{\lfloor x\rfloor}{P}\right\rfloor = q$. Por otro lado, de $x = qP+Q+r$ y del hecho que $Q+r\in[0,P)$ se sigue que $\left\lfloor\frac{x}{P}\right\rfloor = q$, tal como se deseaba obtener.

Ahora bien, si hacemos $a = \lfloor x \rfloor$ se tiene que $T(x) = \log a!$ ó bien $T(x) = \log a!(a+1) - \log(a+1)$. Aplicando la fórmula de Stirling[11] a cada una de estas expresiones resulta que

$$T(x) < \log \sqrt{2\pi} + a\log a - a + \frac{1}{2}\log a + \frac{1}{12a} \tag{1.15}$$

y

$$T(x) > \log \sqrt{2\pi} + (a+1)\log(a+1) - (a+1) - \frac{1}{2}\log(a+1). \tag{1.16}$$

Luego, si $x \geq 1$ se tiene que

$$T(x) < \log \sqrt{2\pi} + x\log x - x + \frac{1}{2}\log x + \frac{1}{12}$$

y

$$T(x) > \log \sqrt{2\pi} + x\log x - x - \frac{1}{2}\log x.$$

Las dos desigualdades anteriores son consecuencia de los estimados en (1.15) y (1.16), respectivamente, y de las inecuaciones

$$\begin{aligned} x\log x - x + \frac{1}{2}\log x + \frac{1}{12} &\geq a\log a - a + \frac{1}{2}\log a + \frac{1}{12a} \\ x\log x - x - \frac{1}{2}\log x &\leq (a+1)\log(a+1) - (a+1) - \frac{1}{2}\log(a+1), \end{aligned}$$

las cuales son válidas[12] siempre que x es mayor o igual a 2.

En la luz de los desarrollos anteriores tenemos entonces que

$$T(x) + T(x/30) < 2\log \sqrt{2\pi} + \frac{1}{6} + \frac{31}{30}x\log x - x\log 30^{1/30} - \frac{31}{30}x + \log x - \frac{1}{2}\log 30$$

y

$$T(x) + T(x/30) > 2\log \sqrt{2\pi} + \frac{31}{30}x\log x - x\log 30^{1/30} - \frac{31}{30}x - \log x + \frac{1}{2}\log 30.$$

[11] Para cada natural a se tiene que $\log a! = \log \sqrt{2\pi} + a\log a - a + \frac{1}{2}\log a + \frac{\theta_a}{12a}$, donde $\theta_a \in (0,1)$.

[12] Ambas desigualdades se pueden establecer mediante el conocido criterio que relaciona el signo de la primera derivada de una función y sus intervalos de monotonía.

Análogamente, se tiene que $T(x/2)+T(x/3)+T(x/5)$ es una expresión dominada por

$$3\log\sqrt{2\pi}+\frac{1}{4}+\frac{31}{30}x\log x-x\log 2^{1/2}3^{1/3}5^{1/5}-\frac{31}{30}x+\frac{3}{2}\log x-\frac{1}{2}\log 30$$

y siempre mayor a

$$3\log\sqrt{2\pi}+\frac{31}{30}x\log x-x\log 2^{1/2}3^{1/3}5^{1/5}-\frac{31}{30}x-\frac{3}{2}\log x+\frac{1}{2}\log 30.$$

Luego, si $A=\log\dfrac{2^{1/2}3^{1/3}5^{1/5}}{30^{1/30}}$, los cuatro estimados previos implican que

$$Ax-\frac{5}{2}\log x+\frac{1}{2}\log\frac{450}{\pi}-\frac{1}{4}<V(x)<Ax+\frac{5}{2}\log x-\frac{1}{2}\log 1800\pi+\frac{1}{6} \tag{1.17}$$

siempre que $x\geq 30$. A fin de obtener cotas válidas para cada $x\geq 1$, procedemos a reemplazar ambos extremos de la desigualdad en (1.17) por expresiones un tanto más burdas. Por ejemplo, como la suma de los últimos dos términos en el extremo derecho de (1.17) es negativa, podemos eliminarlos de la expresión y la función restante conserva el sentido de la desigualdad. En el extremo izquierdo, por su parte, lo que hacemos es reemplazar los dos términos constantes por un -1 sin que esto conlleve alguna modificación en el sentido de la desigualdad correspondiente. Se tiene entonces para cada $x\geq 1$ que

$$Ax-\frac{5}{2}\log x-1<V(x)<Ax+\frac{5}{2}\log x. \tag{1.18}$$

La veracidad de las desigualdades anteriores para los naturales en el intervalo $[1,30)$ fue verificada con ayuda de un sistema de álgebra computacional.

De (1.14) y (1.18) se tiene entonces que

$$\psi(x)>Ax-\frac{5}{2}\log x-1 \quad \text{y} \quad \psi(x)-\psi\left(\frac{x}{6}\right)<Ax+\frac{5}{2}\log x \tag{1.19}$$

siempre que $x \geq 1$. La primera desigualdad nos proporciona una cota inferior para $\psi(x)$. Se espera entonces que con ayuda de la segunda desigualdad se pueda derivar una expresión que domine a ψ.

Consideremos la función f definida como

$$f(x) = \frac{6}{5}Ax + \frac{5}{4\log 6}\log^2 x + \frac{5}{4}\log x.$$

Dicha función f es tal que

$$f(x) - f(x/6) = Ax + \frac{5}{2}\log x$$

para cada $x > 0$. De esta notable identidad se sigue ahora que

$$\psi(x) - f(x) < \psi(x/6) - f(x/6)$$

siempre que $x > 0$ y por tanto

$$\psi(x) - f(x) < \psi(x/6) - f(x/6) < \ldots < \psi(x/6^{k+1}) - f(x/6^{k+1}) \tag{1.20}$$

para cada $x > 0$ y $k \in \mathbb{N}$. En particular, si 6^k es la mayor potencia de 6 menor o igual a x se tiene que $\dfrac{x}{6^{k+1}} \in [1/6, 1)$. Claramente, la función ψ es nula a lo largo de dicho intervalo y $-f$ no es más grande que 1. Las desigualdades en (1.20) nos permiten concluir ahora que

$$\psi(x) < \frac{6}{5}Ax + \frac{5}{4\log 6}\log^2 x + \frac{5}{4}\log x + 1 \tag{1.21}$$

para cada $x \geq 1$. El esfuerzo llevado a cabo en la determinación de los estimados *explícitos* en (1.19) y (1.21) comenzará por rendir frutos en el siguiente apartado.

§**4.** La primera conclusión que derivamos de las inecuaciones en (1.19) y (1.21) es que la función ψ es $O(x)$, donde O es el símbolo de Bachmann-Landau. Aunque a principios de cuenta dicha formulación sobre el carácter asintótico de ψ es suficiente para la mayoría de nuestros fines a largo plazo, vamos a darnos a la búsqueda de estimaciones un tanto más concretas para la función ψ.

Empezamos con la desigualdad de más a la izquierda en (1.19). Sea α un real en el intervalo $(0, A)$. Dado que $\lim_{x\to\infty} \dfrac{\frac{5}{2}\log x + 1}{x} = 0$, aseguramos la existencia de $x_0 \in \mathbb{R}$ tal que $\left|\frac{\frac{5}{2}\log x+1}{x}\right| < \alpha$ cuando $x > x_0$. En particular, si $x > M_\alpha = \text{máx}(1, x_0)$ tenemos que

$$\frac{5}{2}\log x + 1 < \alpha \cdot x \tag{1.22}$$

y el estimado correspondiente en (1.19) sería

$$\psi(x) > (A - \alpha)x \tag{1.23}$$

cuando $x > M_\alpha$. Para lograr obtener una expresión equivalente para la cota en (1.21) procedemos de manera análoga. En este caso caso, para $\beta > 0$ tenemos que existe un $x_1 \in \mathbb{R}$ tal que

$$\left|\frac{5}{4\log 6}\log^2 x + \frac{5}{4}\log x + 1\right| < \beta|x|$$

cuando $x > x_1$. Por tanto, si $x > M_\beta = \text{máx}(1, x_1)$ se cumple que

$$\frac{5}{4\log 6}\log^2 x + \frac{5}{4}\log x + 1 < \beta \cdot x$$

y de aquí que en dicho rango se tenga

$$\psi(x) < \left(\frac{6}{5}A + \beta\right)x. \tag{1.24}$$

De (1.23) y (1.24) se desprende entonces que

$$(A - \alpha)x < \psi(x) < \left(\frac{6}{5}A + \beta\right)x \tag{1.25}$$

siempre que $x > M = \text{máx}(M_\alpha, M_\beta)$. Las desigualdades en la línea previa ya nos permiten concluir cuál es el orden de magnitud de la función ψ y además nos proporcionan una demostración bastante transparente del

Teorema 1.10. Existen constantes positivas a y b, con $a < b$, tales que

$$a \cdot x < \vartheta(x) < b \cdot x$$

siempre que x es suficientemente grande.

Prueba. De la proposición **1.8** tenemos que

$$\psi(x) = \vartheta(x) + \vartheta(x^{1/2}) + \vartheta(x^{1/3}) + \dots$$

Luego, si sustituimos x por $x^{1/2}$ resulta que

$$\psi(x^{1/2}) = \vartheta(x^{1/2}) + \vartheta(x^{1/4}) + \vartheta(x^{1/6}) + \dots$$

y por tanto

$$\begin{aligned} \psi(x) - \psi(x^{1/2}) &= \vartheta(x) + \vartheta(x^{1/3}) + \vartheta(x^{1/5}) + \dots \\ \psi(x) - 2\psi(x^{1/2}) &= \vartheta(x) - \vartheta(x^{1/2}) + \vartheta(x^{1/3}) + \dots \end{aligned}$$

Al ser $\vartheta(x)$ una función no decreciente se sigue que

$$\psi(x) - 2\psi(x^{1/2}) \leq \vartheta(x) \leq \psi(x) - \psi(x^{1/2}).$$

Por otra parte, al aplicar las acotaciones hechas en (1.25) se obtiene que

$$\mathbf{a}x - 2\mathbf{b}\sqrt{x} < \vartheta(x) < \mathbf{b}x - \mathbf{a}\sqrt{x} \tag{1.26}$$

siempre que $x > M^2$, donde $\mathbf{a} = (A - \alpha)$ y $\mathbf{b} = \frac{6}{5}A + \beta$. Ahora bien, al tenerse que $\dfrac{2\mathbf{b}\sqrt{x}}{x} \to 0$ cuando $x \to \infty$, para $\gamma \in (0, \mathbf{a})$ se garantiza la existencia de $M_\gamma \in \mathbb{R}$ tal que $2\mathbf{b}\sqrt{x} < \gamma \cdot x$ cuando $x > M_\gamma$. Así, si $x > \text{máx}(M^2, M_\gamma)$ se tiene que

$$(\mathbf{a} - \gamma)x < \vartheta(x) < \mathbf{b} \cdot x$$

y la prueba termina.

□

Vamos a ilustrar ahora mismo una ventaja importante de haber insistido tanto en la obtención de estimados *concretos* para las funciones ψ y ϑ.

Sea ϵ un número real mayor que $\dfrac{\mathbf{b} - \mathbf{a}}{\mathbf{a}}$. Consideremos la diferencia

$$\vartheta[(1 + \epsilon)x] - \vartheta(x). \tag{1.27}$$

Claramente, si x es tal que $\vartheta[(1+\epsilon)x] - \vartheta(x) > 0$ entonces el intervalo $(x,(1+\epsilon)x]$ **contiene** al menos un número primo. Determinemos entonces una condición suficiente para x que nos permita asegurar la validez de esa desigualdad. De las desigualdades en (1.26) se sigue que

$$\begin{aligned}\vartheta[(1+\epsilon)x] - \vartheta(x) &> [\mathbf{a}(1+\epsilon)x - 2\mathbf{b}\sqrt{(1+\epsilon)x}] - (\mathbf{b}x - \mathbf{a}\sqrt{x}) \\ &= [\mathbf{a}\epsilon - (\mathbf{b}-\mathbf{a})]x - (2\mathbf{b}\sqrt{1+\epsilon} - \mathbf{a})\sqrt{x}.\end{aligned}$$

Luego, una condición suficiente para que la expresión $\vartheta[(1+\epsilon)x] - \vartheta(x)$ sea positiva es que

$$[\mathbf{a}\epsilon - (\mathbf{b}-\mathbf{a})]x - (2\mathbf{b}\sqrt{1+\epsilon} - \mathbf{a})\sqrt{x} > 0. \tag{1.28}$$

Dado que $\epsilon > \dfrac{\mathbf{b}-\mathbf{a}}{\mathbf{a}}$, la desigualdad en (1.28) se cumple siempre y cuando $x > \text{máx}(X_0, M^2)$ donde X_0 es la raíz positiva de la ecuación

$$[\mathbf{a}\epsilon - (\mathbf{b}-\mathbf{a})]x - (2\mathbf{b}\sqrt{1+\epsilon} - \mathbf{a})\sqrt{x} = 0.$$

Al resolver dicha ecuación llegamos a que

$$X_0 = \left(\frac{2\mathbf{b}\sqrt{1+\epsilon} - \mathbf{a}}{\mathbf{a}\epsilon - (\mathbf{b}-\mathbf{a})}\right)^2. \tag{1.29}$$

Finalmente, al tenerse que

$$\epsilon > \frac{\mathbf{b}-\mathbf{a}}{\mathbf{a}} = \frac{\left(\frac{6}{5}A + \beta\right) - (A-\alpha)}{A-\alpha} = \frac{A + 5(\alpha+\beta)}{5(A-\alpha)} > \frac{A}{5A} = \frac{1}{5}$$

podemos dar por cierta la tesis del siguiente

Teorema 1.11. Si $\epsilon > \dfrac{1}{5}$ y X_0 es como en (1.29), entonces el intervalo

$$(x,(1+\epsilon)x]$$

contiene al menos un número primo siempre que $x > \text{máx}(X_0, M^2)$.

□

Notemos que al fijar $\epsilon > \frac{1}{5}$ podemos ir hacia atrás en el argumento y determinar de modo explícito las constantes involucradas en el teorema **1.11**. De hecho, al final todo se reduce a la elección de un par (α, β) —las constantes presentes en (1.25)—, a partir de ese punto todo puede irse construyendo paso a paso según las especificaciones hechas en la demostración. Como ya se habrá notado, el teorema **1.11** mejora sustancialmente lo demostrado en la proposición **1.4**. Chebyshev fue el primero en establecer estos resultados en su afán por demostrar el postulado de Bertrand. La introducción de las funciones ψ y ϑ se da también de manera natural en el marco de dicha problemática y nunca bajo el amparo del mezquino interés de generar teoría sólo por el deseo de hacerlo. Comenta H. M. Edwards en [7] que estas aportaciones de Chebyshev fueron relegadas al olvido hasta que las investigaciones de A. Selberg y Erdős hicieron patente el hecho de que las ideas de Chebyshev eran lo suficientemente maleables como para deducir de ellas pruebas elementales de resultados señeros de la Aritmética que se pensaba no podían ser probados sin apelar al Análisis de Fourier o a la Teoría de las Funciones Analíticas.

§**5.** Una vez que hemos determinado el orden de magnitud de la función ψ vamos a emplear la proposición **1.8** con el fin de extraer infomación con respecto al comportamiento asintótico de la función ϑ. Dado que el valor $\vartheta(x)$ se puede considerar, heurísticamente, cercano a $\pi(x)\log x$, la esperanza es que una buena comprensión de la función ϑ nos pueda ayudar a entender mejor el carácter asintótico de la función contadora de primos. Nuestro punto de partida en esta labor se encuentra dado por el

Teorema 1.12. Se verifica la siguiente relación asintótica

$$\psi(x) = \vartheta(x) + O(\sqrt{x}).$$

Prueba. De las definiciones de ψ y ϑ se sigue de modo inmediato que la diferencia $\psi(x) - \vartheta(x)$ es no negativa para cada x real. Para conseguir una

cota superior para esta diferencia partimos de la identidad

$$\psi(x) = \sum_{k\in\mathbb{N}} \vartheta(x^{1/k})$$

la cual se ha probado ya en la proposición **1.8**. Notamos ahora que en la expresión $\sum_{k\in\mathbb{N}} \vartheta(x^{1/k})$ siempre se cumple que los términos a partir de cierto punto se vuelven nulos y por tanto terminan aportando nada a la suma. Vamos a introducir entonces una literal que nos permita estimar sólo la porción significativa de esa suma. Sea $\mathbf{K} = \lfloor \log x/\log 2 \rfloor$. Si $k > \mathbf{K}$ se sigue que $k > \log x/\log 2$. En consecuencia, $1/k < \log 2/\log x = \log_x 2$ y de ahí que $x^{1/k} < x^{\log_x 2} = 2$. Lo anterior indica que los términos de la suma $\vartheta(x) + \vartheta(x^{1/2}) + \ldots$ que aparecen después del $\mathbf{K}$-ésimo son nulos y por tanto

$$\psi(x) - \vartheta(x) = \sum_{2\leq k\leq \mathbf{K}} \vartheta(x^{1/k}) \leq \sum_{2\leq k\leq \mathbf{K}} \psi(x^{1/k}) = \psi(x^{1/2}) + \sum_{2<k\leq \mathbf{K}} \psi(x^{1/k}).$$

Ahora bien, al tenerse que $\psi(x) \asymp x$ se sigue que

$$\sum_{2<k\leq \mathbf{K}} \psi(x^{1/k}) = O(x^{1/3}\log x).$$

Luego, $\psi(x) - \vartheta(x) = O(\sqrt{x}) + O(\sqrt[3]{x}\log x) = O(\sqrt{x})$, que era lo que se deseaba demostrar.

□

La identidad asintótica que viene jugará un papel central en el resto de la discusión sobre la naturaleza asintótica de la función π.

Teorema 1.13. $\pi(x) = \dfrac{\vartheta(x)}{\log x} + O(x/\log^2 x)$.

Prueba. Mostraremos primero que si $x \geq 2$ entonces

$$\pi(x) = \frac{\vartheta(x)}{\log x} + \int_2^x \vartheta(u)u^{-1}\log^{-2} u\, du. \tag{1.30}$$

Sean $p_1, p_2, \ldots, p_r$ los primos comprendidos en el intervalo $[2, x]$ y denotemos con I a la integral $\displaystyle\int_2^x \vartheta(u)u^{-1}\log^{-2} u\, du$. Se cumple entonces que

$$I = \sum_{k=1}^{r} \lim_{n\to\infty} \int_{p_k}^{p_{k+1}-\frac{1}{n+1}} \vartheta(u)u^{-1}\log^{-2} u\, du$$

donde $p_{r+1} = x$. Dado que la función ϑ es igual a $\log p_1 + \ldots + \log p_i$ en el intervalo $[p_i, p_{i+1} - \frac{1}{n+1}]$ se sigue que

$$I = \sum_{k=1}^{r} \sum_{i=1}^{k} \log p_i \lim_{n\to\infty} \int_{p_k}^{p_{k+1}-\frac{1}{n+1}} u^{-1} \log^{-2} u \, du.$$

Ahora bien, si intercambiamos las sumas en el lado derecho de la igualdad anterior llegamos a que

$$\begin{aligned} I &= \sum_{i=1}^{r} \sum_{k=i}^{r} \log p_i \lim_{n\to\infty} \int_{p_k}^{p_{k+1}-\frac{1}{n+1}} u^{-1} \log^{-2} u \, du \\ &= \sum_{i=1}^{r} \log p_i \lim_{n\to\infty} \int_{p_i}^{p_{r+1}-\frac{1}{n+1}} u^{-1} \log^{-2} u \, du \\ &= \sum_{i=1}^{r} \log p_i \left(\frac{1}{\log p_i} - \frac{1}{\log p_{r+1}} \right) \\ &= \pi(x) - \frac{\vartheta(x)}{\log x} \end{aligned}$$

como ya se anunciara en (1.30). La veracidad de la fórmula asintótica prometida depende ahora sólo de la estimación de $\displaystyle\int_2^x \vartheta(u) u^{-1} \log^{-2} u \, du$. Del teorema **1.10** se tiene que $\vartheta(x) \asymp x$. Por tanto,

$$\begin{aligned} \int_2^x \vartheta(u) u^{-1} \log^{-2} u \, du &= O\left(\int_2^x \log^{-2} u \, du \right) \\ &= O\left(\int_2^{\sqrt{x}} \log^{-2} u \, du + \int_{\sqrt{x}}^x \log^{-2} u \, du \right). \end{aligned}$$

La integral del primer término es $O(\sqrt{x})$. Por otra parte, el integrando del segundo término está comprendido entre $\log^{-2} x$ y $\log^{-2} \sqrt{x}$ y de ahí que

$$\int_{\sqrt{x}}^x \log^{-2} u \, du \leq \int_{\sqrt{x}}^x \log^{-2} \sqrt{x} \, du = \frac{4(x - \sqrt{x})}{\log^2 x}.$$

Se tiene así que

$$\int_2^x \vartheta(u) u^{-1} \log^{-2} u \, du = O(\sqrt{x}) + O(x/\log^2 x) = O(x/\log^2 x)$$

y la prueba concluye.

□

De los teoremas **1.10** y **1.13** se obtiene ahora que si $\epsilon > 0$ y x es suficientemente grande entonces

$$(\mathbf{a} - \gamma - \epsilon)\frac{x}{\log x} < \pi(x) < (\mathbf{b} + \epsilon)\frac{x}{\log x}$$

y por tanto, $\pi(x) \asymp x/\log x$. Más aún, dado que tenemos libertad de elección sobre ϵ, γ y las constantes α y β en las definiciones de **a** y **b** (respectivamente) podemos afirmar que $\pi(x)\log x$ se encuentra aproximadamente entre $0.92129x$ y $1.10555x$. Esto demuestra que si el cociente

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

tiene límite, entonces dicho límite debe estar entre 0.92129 y 1.1055.

Es posible que en este punto se salte a la conclusión de que el trabajo preliminar no ha rendido fruto alguno, pues lo único que se ha obtenido a partir de él es una estimación no muy fina para la posible ubicación del límite que se pretende calcular. Si bien se tiene todo el derecho de pensar de tal modo, la verdad es que seguir el camino histórico de cerca nos permite dimensionar adecuadamente el calibre de las diversas contribuciones hechas a la teoría en el afán de determinar el valor del límite en cuestión.

§**6.** Iniciamos el apartado con un resultado que jugará un papel importante en los cálculos posteriores. Es necesario notar que aún cuando los estimados que están por enunciarse no son los mejores que pueden presentarse, sí que son adecuados para las aplicaciones inmediatas que se tiene contempladas.

Lema 1.14. Para cada $n \in \mathbb{N}$ se cumple que

a) $\log n! = \sum_{k\geq 1} \psi\left(\frac{n}{k}\right)$,

b) $\log n! = n\log n + O(n)$, y

c) $\displaystyle\sum_{k=1}^{n} \frac{1}{k} = \log n + O(1)$.

Prueba. La primer identidad es una reformulación de una de las identidades establecidas en la proposición **1.8**. Para demostrar la segunda empezamos por observar que la monotonía de la función $\log x$ implica de inmediato que $\displaystyle\int_{k-1}^{k} \log u \, du$ es siempre menor o igual a $\log k$. Luego,

$$\log n! = \sum_{k=2}^{n} \log k \geq \sum_{k=2}^{n} \int_{k-1}^{k} \log u \, du = \int_{1}^{n} \log u \, du = n \log n - n + 1. \tag{1.31}$$

Para obtener una cota superior para $\log n!$ notamos ahora que $\displaystyle\int_{k}^{k+1} \log u \, du$ es mayor o igual a $\log k$ para cada $k \in \mathbb{N}$ y de aquí que

$$\begin{aligned} \log n! &= \log n + \sum_{k=1}^{n-1} \log k \\ &\leq \log n + \sum_{k=1}^{n-1} \int_{k}^{k+1} \log u \, du \\ &= \log n + \int_{1}^{n} \log u \, du \\ &= \log n + n \log n - n + 1. \end{aligned} \tag{1.32}$$

De (1.31) y (1.32) se desprende ahora que

$$|\log n! - n \log n| \leq (n - 1)$$

y la prueba del inciso b termina. La estimación en c depende sólo del hecho de que $\log n = \displaystyle\int_{1}^{n} \frac{dt}{t}$. En efecto, de

$$\int_{1}^{n} \frac{dt}{t} = \sum_{k=1}^{n-1} \int_{k}^{k+1} \frac{dt}{t} \leq \sum_{k=1}^{n-1} \frac{1}{k} = \sum_{k=1}^{n} \frac{1}{k} - \frac{1}{n}$$

se sigue que la expresión $S(n) = \sum_{k=1}^{n} \frac{1}{k} - \int_1^n \frac{dt}{t}$ es no negativa para cada natural n. Por otro lado, de

$$\int_1^n \frac{dt}{t} = \sum_{k=2}^{n} \int_{k-1}^{k} \frac{dt}{t} \geq \sum_{k=2}^{n} \frac{1}{k}$$

obtenemos que $1 \geq \sum_{k=1}^{n} \frac{1}{k} - \int_1^n \frac{dt}{t}$ y por consiguiente $|S(n)| \leq 1$. Luego,

$$\left| \sum_{k=1}^{n} \frac{1}{k} - \log n \right| \leq 1,$$

lo cual se deseaba establecer.

□

Consideremos ahora un natural m mayor que 1. Si n es un natural arbitrario mayor que m se tiene, en virtud de la parte a del lema recién establecido, que

$$\sum_{k=1}^{\lfloor n/m \rfloor} \frac{1}{k} \frac{\psi\left(\frac{n}{k}\right)}{\frac{n}{k}} \leq \frac{\log n!}{n} = \sum_{k=1}^{\lfloor n/m \rfloor} \frac{1}{k} \frac{\psi\left(\frac{n}{k}\right)}{\frac{n}{k}} + \sum_{k=\lfloor n/m \rfloor + 1}^{n} \frac{1}{k} \frac{\psi\left(\frac{n}{k}\right)}{\frac{n}{k}}. \tag{1.33}$$

Si definimos ahora a L_m^-, L_m^+, λ_m como

$$L_m^- = \operatorname{ínf}_{x \geq m} \frac{\psi(x)}{x}, \quad L_m^+ = \sup_{x \geq m} \frac{\psi(x)}{x}, \quad \lambda_m = \sup_{1 \leq x < m} \frac{\psi(x)}{x},$$

las relaciones en (1.33) nos permiten asegurar que

$$L_m^- \sum_{k=1}^{\lfloor n/m \rfloor} \frac{1}{k} \leq \frac{\log n!}{n} \leq L_m^+ \sum_{k=1}^{\lfloor n/m \rfloor} \frac{1}{k} + \lambda_m \sum_{k=\lfloor n/m \rfloor + 1}^{n} \frac{1}{k}.$$

De los apartados b y c del lema **1.14** y de las desigualdades en el renglón anterior se sigue que

$$\begin{aligned} L_m^-(\log n - \log m + O(1)) &\leq \log n + O(1) \\ &\leq L_m^+(\log n - \log m + O(1)) + \lambda_m(\log m + O(1)). \end{aligned}$$

Luego,

$$L_m^- \leq \frac{\log n + O(1)}{\log n - \log m + O(1)}$$

y

$$L_m^+ \geq \frac{\log n + O(1) - \lambda_m(\log m + O(1))}{\log n - \log m + O(1)}.$$

Puesto que ambas desigualdades son válidas para cada natural n mayor que m y λ_m es finito, el análisis previo permite concluir que

$$L_m^- = \operatorname*{ínf}_{x\geq m} \frac{\psi(x)}{x} \leq 1 \quad \text{y} \quad L_m^+ = \sup_{x\geq m} \frac{\psi(x)}{x} \geq 1;$$

de esto último se desprende, al hacer $m \to \infty$, que

$$\liminf_{x\to\infty} \frac{\psi(x)}{x} \leq 1 \quad \text{y} \quad \limsup_{x\to\infty} \frac{\psi(x)}{x} \geq 1. \tag{1.34}$$

Resulta más que justo condensar ahora todos estos razonamientos en la siguiente

Proposición 1.15. Se cumplen las siguientes desigualdades

$$0 < \liminf_{x\to\infty} \frac{\psi(x)}{x} \leq 1 \leq \limsup_{x\to\infty} \frac{\psi(x)}{x} < \infty.$$

Prueba. El par de desigualdades internas es consecuencia de lo concluido en (1.34). Las dos desigualdades externas son consecuencia de lo expresado en (1.25).

□

¿En qué se ha mejorado nuestra comprensión de la naturaleza asintótica de la función contadora de primos después del estudio que se acaba de efectuar? En el §**5** habíamos establecido ya que $\pi(x)$ es del mismo orden de magnitud que la función $x/\log x$ y de hecho, al final de dicho apartado, se mencionaron algunas cotas para la ubicación del límite de $\pi(x)\log x/x$ cuando $x \to \infty$. Lo que se acaba de hacer nos permite ir un poco más

allá. Como consecuencia inmediata de la proposición **1.15** tenemos que si acaso el límite de $\frac{\psi(x)}{x}$ existe entonces dicho límite tiene que ser exactamente igual a 1. Este hecho tendría repercusiones de peso sobre nuestro conocimiento de la función contadora de primos pues si el mencionado límite es 1 entonces el teorema **1.12** nos permitiría asegurar que $\frac{\vartheta(x)}{x} \to 1$ cuando $x \to \infty$. Este último resultado, en conjunción con el teorema **1.13**, darían lugar a su vez a la identidad

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log x} = 1. \tag{1.35}$$

La discusión precedente pone a la fórmula empírica derivada al final de §**1** en un terreno mucho más firme. En el capítulo siguiente nos daremos a la tarea de presentar demostraciones concretas de la validez de dicha fórmula.

Por ahora, lo único que agregaremos es que la aparición de la primer conjetura cercana a la tesis del Teorema del Número Primo tomó lugar en el *Essai sur la Théorie de Nombres* de A. M. Legendre en 1798. Legendre afirmaba que la función $\pi(x)$ era de la forma $x/(A\log x + B)$ para algunas constantes A y B. En 1808, Legendre refinaría su anterior conjetura al afirmar que

$$\pi(x) = \frac{x}{\log x - A(x)}$$

donde $A(x)$ es "aproximadamente 1.08366 . . .".

Aún cuando Legendre fue la primer persona en publicar una versión conjetural del Teorema del Número Primo, se sabe que C. F. Gauss había hecho extensas investigaciones en el problema de la distribución de los números primos entre 1792 y 1793. El estudio minucioso de las tablas matemáticas de Johann Carl Schulze permitió a Gauss conjeturar que el número de primos en el intervalo $[1, x]$ es aproximadamente igual a

$$\int_2^x \frac{dt}{\log t}.$$

Es importante mencionar que esta versión de Gauss del Teorema del Número Primo es totalmente equivalente a la que aparece en (1.35). Empero, la conjetura hecha por Gauss tiene la ventaja de que en ella aparece la integral

$$\int_2^x \frac{dt}{\log t}$$

la cual proporciona, por sí misma, una mejor aproximación númerica a $\pi(x)$ que $x/\log x$.

Capítulo 2

Dos pruebas

§**1.** Los trabajos de Chebyshev en Teoría de Números representaron el primer avance concreto hacia el establecimiento de la conjetura de Gauss-Legendre sobre la equivalencia asintótica de las funciones $\pi(x)$ y $x/\log x$. Un cambio de paradigma introducido en la teoría por la mente visionaria de Bernhard Riemann permitiría establecer algunos años después la validez de dicha conjetura. La motivación inicial de las aportaciones de Riemann provino de la siguiente identidad introducida en el siglo XVIII por L. Euler:

Proposición 2.1. Si $s \in \mathbb{C}$ es tal que $\Re(s) > 1$ entonces

$$\prod_{p\in\mathbf{P}}\left(1-\frac{1}{p^s}\right)^{-1} = \sum_{n=1}^{\infty}\frac{1}{n^s}. \tag{2.1}$$

Antes de pasar a la demostración de la proposición mencionaremos, haciendo honor a la verdad, que la identidad conocida por Euler era en realidad la versión *real* de la presentada en (2.1); esto es, Euler restringió su atención, como era la usanza en su tiempo, a la consideración del caso s real con $s > 1$. Lo anterior no es indicio de una falta de previsión de Euler, más bien, es una muestra de que el uso de parámetros complejos no estaba lo suficientemente afianzado en el consciente coletivo de su época. Otro punto que es importante mencionar es la convergencia absoluta y

uniforme[1] de la serie $\sum_{n=1}^{\infty} \frac{1}{n^s}$ en subconjuntos compactos del semiplano $\sigma > 1$. Esto indica que la identidad en (2.1) tiene relevancia al nivel operacional y que no debe entenderse como un resultado puramente formal. Nuestra digresión llega a su fin con la observación anterior y lo que está en orden es la

Prueba de 2.1. Supongamos que $\mathbf{P} = \{p_1, p_2, \ldots\}$ y para $m \in \mathbb{N}$ denotemos con S_m al conjunto formado por los naturales que no tienen divisores entre los primeros m números primos. Probaremos a continuación, haciendo inducción sobre m, que la fórmula

$$\left(\sum_{n=1}^{\infty} \frac{1}{n^s}\right) \prod_{k=1}^{m} \left(1 - \frac{1}{p_k^s}\right) = \sum_{n \in S_m} \frac{1}{n^s} \tag{2.2}$$

es siempre justa. De la convergencia absoluta de la serie $\sum_{n=1}^{\infty} \frac{1}{n^s}$ en el semiplano $\sigma > 1$ se sigue que

$$\begin{aligned} \sum_{n \in S_1} \frac{1}{n^s} &= \sum_{n=1}^{\infty} \frac{1}{n^s} - \sum_{n=1}^{\infty} \frac{1}{(2n)^s} \\ &= \sum_{n=1}^{\infty} \frac{1}{n^s} - \frac{1}{2^s} \sum_{n=1}^{\infty} \frac{1}{n^s} \\ &= \left(1 - \frac{1}{2^s}\right) \sum_{n=1}^{\infty} \frac{1}{n^s}. \end{aligned}$$

Esta última igualdad establece la base de la inducción para el aserto en

[1] Si s está en un subconjunto compacto del semiplano $\sigma > 1$ entonces $\sigma \geq 1+\delta$ para algún $\delta > 0$. De la cadena de igualdades $|n^s| = |n^{\sigma+it}| = n^{\sigma}|e^{it \log n}| = n^{\sigma}$ se sigue que $\left|\frac{1}{n^s}\right| \leq \frac{1}{n^{1+\delta}}$ para cada $n \in \mathbb{N}$. La convergencia absoluta de $\sum_{n=1}^{\infty} \frac{1}{n^s}$ es ahora clara y una manera breve de establecer la convergencia uniforme es aplicando el criterio M de Weierstrass.

(2.2). Supongamos válido el resultado para m. La cadena de igualdades

$$\begin{aligned}\left(\sum_{n=1}^{\infty}\frac{1}{n^s}\right)\prod_{k=1}^{m+1}\left(1-\frac{1}{p_k^s}\right) &= \left(1-\frac{1}{p_{m+1}^s}\right)\left(\sum_{n=1}^{\infty}\frac{1}{n^s}\right)\prod_{k=1}^{m}\left(1-\frac{1}{p_k^s}\right)\\ &= \left(1-\frac{1}{p_{m+1}^s}\right)\sum_{n\in S_m}\frac{1}{n^s}\\ &= \sum_{n\in S_m}\frac{1}{n^s}-\sum_{n\in S_m}\frac{1}{(p_{m+1}\cdot n)^s}\\ &= \sum_{n\in S_{m+1}}\frac{1}{n^s}\end{aligned}$$

implica la veracidad del resultado para $m+1$ y la prueba de (2.2) termina. La idea es ahora deducir (2.1) a partir de (2.2) al hacer $m\to\infty$.

Sea $a_k=-\dfrac{1}{p_k^s}$. La convergencia absoluta de la serie $\displaystyle\sum_{k=1}^{\infty}\frac{1}{k^s}$ en el semiplano $\sigma>1$ implica la convergencia de $\displaystyle\sum_{k=1}^{\infty}|a_k|$ en esa misma región. Ahora bien, las desigualdades

$$\frac{|a_n|}{2}<|\log(1+a_n)|<\frac{3|a_n|}{2}$$

son válidas para n suficientemente grande y el criterio de comparación garantiza entonces la convergencia de $\displaystyle\sum_{n=1}^{\infty}\log(1+a_n)$.

Supongamos que $\mathfrak{S}=\displaystyle\sum_{n=1}^{\infty}\log(1+a_n)$. La continuidad de la función exponencial en $\mathbb{C}$ permite asegurar entonces que

$$\exp(\mathfrak{S})=\exp\left(\lim_{N\to\infty}\sum_{n=1}^{N}\log(1+a_n)\right)=\lim_{N\to\infty}\prod_{n=1}^{N}(1+a_n).$$

Así, el producto $\displaystyle\prod_{k=1}^{\infty}(1+a_k)$ es finito y distinto de cero y por tanto

$$\sum_{n=1}^{\infty}\frac{1}{n^s}=\lim_{m\to\infty}\frac{\sum_{n\in S_m}\frac{1}{n^s}}{\prod_{n=1}^{m}(1+a_n)}=\prod_{n=1}^{\infty}(1+a_n)^{-1}=\prod_{p\in\mathbf{P}}\left(1-\frac{1}{p^s}\right)^{-1},$$

como deseabamos establecer.

□

La serie que aparece en el lado derecho de la identidad que se acaba de probar se denota por $\zeta(s)$ y la función que se obtiene al variar el parámetro s en dicha serie se conoce como **función zeta de Riemann**. En la demostración de la proposición **2.1** obtuvimos de paso un dato adicional sobre la naturaleza de dicha función:

Proposición 2.2. La función ζ no se anula en el semiplano $\sigma > 1$.

□

Otra propiedad de la función ζ que ya estamos en posición de aseverar es su analiticidad en el semiplano $\sigma > 1$. Esto es en realidad una consecuencia directa de la analiticidad de las funciones $S_N(s) = \sum_{n=1}^{N} \frac{1}{n^s}$ en $\sigma > 1$ y del hecho que la sucesión $\{S_N\}_{N\in\mathbb{N}}$ converge uniformemente a la función ζ en los subconjuntos compactos de dicho dominio[2]. Un dato más fino sobre la naturaleza de la función zeta está dado por el siguiente resultado

Proposición 2.3. La función ζ admite una extensión analítica en el semiplano $\sigma > 0$, excepto por un polo simple en $s = 1$ con residuo 1.

Prueba. Sean N y M dos enteros positivos con $N < M$. Haciendo $a_n = 1$

[2]El resultado que respalda a esta conclusión se conoce en algunos textos como teorema de convergencia analítica. Esa es la denominación que emplearemos en lo sucesivo para referirnos al teorema en cuestión (cf. J. E. Marsden; M. J. Hoffman. *Análisis Básico de Variable Compleja.* Editorial Trillas, México, 1996, pág. 211.).

y $f(u) = \frac{1}{u^s}$ en la identidad de Abel[3] se llega a que

$$\begin{aligned}\sum_{N<n\leq M} \frac{1}{n^s} &= M^{1-s} - N^{1-s} + s\int_N^M \frac{\lfloor u\rfloor}{u^{s+1}}\,du \\ &= M^{1-s} - N^{1-s} + s\int_N^M \frac{u}{u^{s+1}}\,du + s\int_N^M \frac{\lfloor u\rfloor - u}{u^{s+1}}\,du \\ &= \frac{N^{1-s} - M^{1-s}}{s-1} + s\int_N^M \frac{\lfloor u\rfloor - u}{u^{s+1}}\,du\end{aligned}$$

siempre que $\Re(s) > 1$. Al tomar límite cuando $M \to \infty$ obtenemos

$$\zeta(s) = \sum_{n\leq N} \frac{1}{n^s} + \frac{N^{1-s}}{s-1} + s\int_N^\infty \frac{\lfloor u\rfloor - u}{u^{s+1}}\,du.$$

Cabe notar que la integral presente en el miembro derecho de la igualdad anterior existe para cada s en el semiplano $\sigma > 0$ y para todo natural N. En particular, si $N = 1$ la identidad nos permite concluir que

$$\zeta(s) = 1 + \frac{1}{s-1} + s\int_1^\infty \frac{\lfloor u\rfloor - u}{u^{s+1}}\,du = \frac{s}{s-1} + s\int_1^\infty \frac{\lfloor u\rfloor - u}{u^{s+1}}\,du \qquad (2.3)$$

y de aquí la extensión anunciada para la funcíon ζ.

□

§**2.** Vamos a utilizar ahora la identidad en (2.1) para relacionar la función ζ con la función ψ de Chebyshev estudiada en el capítulo anterior.

Para s en $\sigma > 1$ la proposición **2.1** indica que

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

Por otro lado, la no anulación de la función ζ en $\sigma > 1$ nos permite asegurar que

$$\log\zeta(s) = \sum_{p\in\mathbf{P}} \log\left(\frac{p^s}{p^s-1}\right) + 2i\pi k$$

[3]Estamos considerando la siguiente formulación de la identidad: si $f \in C^1[1,\infty)$ y $A(x)$ es la función suma de la función $a(n) = a_n$ entonces $\sum_{1\leq n\leq x} a_n f(n) = A(x)f(x) - \int_1^x A(t)f'(t)\,dt$.

para algún k fijo en $\mathbb{Z}$. Puesto que los términos de la serie en la identidad previa son funciones analíticas en $\sigma > 1$, la derivada de la función $\log \zeta(s)$ se puede obtener derivando término a término dicha serie. Así, llegamos a que

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_{p\in\mathbf{P}} \frac{d}{ds} \log\left(\frac{p^s}{p^s-1}\right) = -\sum_{p\in\mathbf{P}} \frac{p^{-s}}{1-p^{-s}} \log p = -\sum_{p\in\mathbf{P}} (p^{-s} + p^{-2s} + \ldots) \log p. \tag{2.4}$$

A fin de obtener una expresión más simple para la última serie de (2.4), introducimos la función Λ de von Mangoldt, la cual se define como sigue

$$\Lambda(n) = \begin{cases} \log p & \text{si } n = p^k \text{ donde } p \in \mathbf{P} \text{ y } k \in \mathbb{N} \\ 0 & \text{en otro caso.} \end{cases}$$

Resulta intuitivamente claro ahora que la serie $\displaystyle\sum_{n\in\mathbb{N}} \frac{\Lambda(n)}{n^s}$ coincide con $\displaystyle\sum_{p\in\mathbf{P}} (p^{-s} + p^{-2s} + \ldots) \log p$. Para *ver* esto basta con notar que en la serie donde aparece la función Λ de von Mangoldt abundan los términos nulos y cuando $n = p^k$ el sumando correspondiente se transforma en $\dfrac{\log p}{p^{ks}}$. La identidad en cuestión se tiene entonces al agrupar los términos que tienen a $\log p$ por factor común. Claramente, dichos términos son exactamente las potencias naturales de p^{-s}. La validez del reacomodo final queda justificada por el

Teorema 2.4. La serie $\displaystyle\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}$ converge absolutamente para cada s en el semiplano $\sigma > 1$.

Prueba. Supongamos que $s = \sigma + it$. Para cada $n \in \mathbb{N}$ se tiene que

$$\left|\frac{\Lambda(n)}{n^s}\right| = \frac{\Lambda(n)}{n^\sigma} \leq \frac{\log n}{n^\sigma}. \tag{2.5}$$

Consideremos ahora dos reales positivos A y B tales que $\sigma = 1 + A + B$. Puesto que

$$\lim_{n\to\infty} \frac{\log n}{n^A} = 0$$

aseguramos la existencia de $M \in \mathbb{R}^+$ tal que $0 \leq \dfrac{\log n}{n^A} \leq M$ para cada $n \in \mathbb{N}$. Luego,

$$\frac{\log n}{n^\sigma} = \frac{\log n}{n^{1+A+B}} = \left(\frac{\log n}{n^A}\right)\left(\frac{1}{n^{1+B}}\right) \leq \frac{M}{n^{1+B}}$$

y de aquí la convergencia de la serie de término n-ésimo $\dfrac{\log n}{n^\sigma}$. El teorema se sigue ahora de la desigualdad en (2.5) al aplicar el criterio de comparación. □

Antes de volver con la identidad en (2.4) vamos a mencionar una relación explícita entre la función ψ de Chebyshev y la función Λ de von Mangoldt. En el transcurso de la prueba de la proposición **1.8** se mencionó que el coeficiente del término $\log p$ en la serie que define a $\psi(x)$ es igual a $k_p = \text{máx}\{i \in \mathbb{N} : x^{1/i} \geq p\}$. Por otro lado, si consideramos la función suma de la función de von Mangoldt y la denotamos con $F(x)$ tenemos que

$$F(x) = \sum_{n \leq x} \Lambda(n) = \sum_{p \leq x} \sum_{p^m \leq x} \log p = \sum_{p \leq x} k_p \log p = \psi(x),$$

esto es, la función suma de la función de von Mangoldt es precisamente la función ψ de Chebyshev. Por tanto,

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = \lim_{k\to\infty}\left(\sum_{n=1}^{k} \frac{\Lambda(n)}{n^s}\right) = \lim_{k\to\infty} \int_{1^-}^{k} \frac{d\psi(t)}{t^s} = \int_{1^-}^{\infty} \frac{d\psi(t)}{t^s}.$$

El resultado siguiente será clave en la obtención de información relativa al comportamiento de la función ζ sobre la recta $\sigma = 1$.

Teorema 2.5. Sea $f(s)$ una función analítica en la cerradura del semiplano $\sigma > 1$, excepto por un polo simple en $s = 1$. Si $f(s)$ no se anula en $\sigma > 1$ y

$$-f'(s)/f(s) = \int_0^{\infty} e^{-sx}\, d\phi(x) \tag{2.6}$$

donde $\phi(x)$ es una función no decreciente cuando $x \geq 0$ entonces $f(s)$ no se anula en la recta $\sigma = 1$.

Prueba. Sea $s = \sigma + it$. Si a la función $g(\sigma, t) = -\Re\{f'(s)/f(s)\}$ se le aplica la hipótesis en (2.6) se obtiene

$$\begin{aligned} g(\sigma,t) &= \Re\left\{\int_0^\infty e^{-sx}\, d\phi(x)\right\} \\ &= \Re\left\{\int_0^\infty e^{-\sigma x}e^{-itx}\, d\phi(x)\right\} \\ &= \int_0^\infty e^{-\sigma x}\cos(tx)\, d\phi(x). \end{aligned}$$

La desigualdad de Cauchy-Schwarz indica entonces que

$$\begin{aligned} g(\sigma,t)^2 &= \left(\int_0^\infty e^{-\sigma\cdot x/2}[e^{-\sigma\cdot x/2}\cos(tx)]\, d\phi(x)\right)^2 \\ &\leq \int_0^\infty e^{-\sigma x}\, d\phi(x)\int_0^\infty e^{-\sigma x}\cos^2(tx)\, d\phi(x) \\ &= \int_0^\infty e^{-\sigma x}\, d\phi(x)\int_0^\infty e^{-\sigma x}\left(\frac{1+\cos(2tx)}{2}\right) d\phi(x) \\ &= \frac{g(\sigma,0)\{g(\sigma,0)+g(\sigma,2t)\}}{2}. \end{aligned} \tag{2.7}$$

Ahora bien, dado que f tiene un polo simple en $s = 1$ se tiene que los desarrollos de Laurent para f y f' alrededor de $s = 1$ son

$$f(s) = \frac{b_1}{s-1} + a_0 + a_1(s-1) + a_2(s-1)^2 + \ldots$$

y

$$f'(s) = -\frac{b_1}{(s-1)^2} + a_1 + 2a_2(s-1) + \ldots$$

Luego,

$$\begin{aligned} (\sigma-1)g(\sigma,0) &= (\sigma-1)(-1)\Re\left\{\frac{-\frac{b_1}{(\sigma-1)^2} + a_1 + 2a_2(\sigma-1)+\ldots}{\frac{b_1}{(\sigma-1)} + a_0 + a_1(\sigma-1) + a_2(\sigma-1)^2+\ldots}\right\} \\ &= \Re\left\{\frac{\frac{b_1}{(\sigma-1)} - a_1(\sigma-1) - 2a_2(\sigma-1)^2-\ldots}{\frac{b_1}{(\sigma-1)} + a_0 + a_1(\sigma-1) + a_2(\sigma-1)^2+\ldots}\right\} \\ &= \Re\left\{\frac{b_1 - a_1(\sigma-1)^2 - 2a_2(\sigma-1)^3-\ldots}{b_1 + a_0(\sigma-1) + a_1(\sigma-1)^2 + a_2(\sigma-1)^3+\ldots}\right\} \end{aligned}$$

y por tanto

$$\lim_{\sigma \to 1^+} (\sigma - 1) g(\sigma, 0) = 1.$$

Supongamos que ahora $f(s)$ tiene un cero de multiplicidad $m = m(t) \geq 0$ en $s = 1 + it$. El desarrollo de Taylor alrededor de s de la función f es

$$f(z) = a_m (z - s)^m + a_{m+1} (z - s)^{m+1} + \dots$$

y de aquí que

$$\begin{aligned} (\sigma - 1) g(\sigma, t) &= (\sigma - 1)(-1) \Re \left\{ \frac{a_m m (\sigma - 1)^{m-1} + a_{m+1} (m + 1)(\sigma - 1)^m + \dots}{a_m (\sigma - 1)^m + a_{m+1} (\sigma - 1)^{m+1} + \dots} \right\} \\ &= (-1) \Re \left\{ \frac{a_m m (\sigma - 1)^m + a_{m+1} (m + 1)(\sigma - 1)^{m+1} + \dots}{a_m (\sigma - 1)^m + a_{m+1} (\sigma - 1)^{m+1} + \dots} \right\} \end{aligned}$$

Por tanto, $(\sigma - 1) g(\sigma, t) \to -m$ cuando $\sigma \to 1^+$. Si en la desigualdad de (2.7) multiplicamos ambos lados por $(\sigma - 1)^2$, se observa que

$$g(\sigma, t)^2 (\sigma - 1)^2 \leq \frac{g(\sigma, 0)^2 (\sigma - 1)^2 + g(\sigma, 0) g(\sigma, 2t) (\sigma - 1)^2}{2}.$$

Al tomar límite cuando $\sigma \to 1^+$ en la anterior desigualdad se obtiene que

$$m(t)^2 \leq \frac{1 - m(2t)}{2} \leq \frac{1}{2}$$

y de esto que $m(t) = 0$. Luego, la función f no tiene ceros sobre la línea $\sigma = 1$.

□

Nótese que la extensión analítica de la función zeta de Riemann que se obtuvo en la proposición **2.3** satisface todas las condiciones del teorema anterior. Luego, como corolario inmediato de dicho resultado se tiene que **la función ζ no se anula a lo largo del conjunto que queda al remover el punto $s = 1$ de la recta $\sigma = 1$.** Este resultado será sumamente importante en ambos enfoques que en este capítulo se emplean para establecer la conjetura de Gauss-Legendre.

§**3.** La primer prueba que presentaremos de la conjetura de Gauss y Legendre la obtendremos del

Teorema 2.6. (D. J. Newman) Sea $f : [0, \infty) \to \mathbb{R}$ una función acotada y localmente integrable. Para $z \in \sigma > 0$ sea $g(z) = \int_0^\infty f(t)e^{-zt}\, dt$. Si $g(z)$ admite una extensión analítica en la cerradura de $\sigma > 0$ entonces

$$\lim_{T\to\infty} \int_0^T f(t)\, dt = g(0).$$

Prueba. Para $T > 0$ la función $g_T(z) = \int_0^T f(t)e^{-zt}dt$ es entera. Sea R un real positivo y C la frontera de la región

$$\{z \in \mathbb{C} : |z| \leq R \quad \text{y} \quad \Re(z) \geq -\delta\}$$

donde $\delta > 0$ es lo suficientemente pequeño como para que $g(z)$ resulte analítica sobre C y en el interior de la región acotada por C.

Una manera de ver que siempre es posible dar un δ tal es como sigue: la función g es analítica en el segmento de recta L que une a $-Ri$ con Ri y por tanto para cada $z \in L$ existe una bola $B(z, r_z)$ donde g es analítica. La colección de todas estas bolas proporciona una cubierta abierta de L. Al ser L compacto garantizamos la existencia de un recubrimiento finito para L por bolas centradas en puntos sobre él. Ordenemos los centros de dichas bolas en orden decreciente con respecto a su parte imaginaria. Digamos que la lista resultante de centros es: $z_1, z_2, \ldots, z_k$. Sin pérdida de generalidad podemos suponer que $z_1 = iR$ y que $z_k = -iR$. Otra suposición que podemos hacer es que no hay bolas tangentes y que ninguna de ellas queda contenida dentro de alguna otra. Así, si $\{i_1, \cdots, i_k\}$ son los puntos de intersección en el semiplano izquierdo de las bolas del recubrimiento, basta con tomar $\delta = \frac{1}{2} \min_{1\leq j\leq k} |\Re(i_j)|$.

La fórmula integral de Cauchy permite asegurar entonces que

$$g(0) - g_T(0) = \frac{1}{2\pi i} \int_C (g(z) - g_T(z))e^{zT} \left(1 + \frac{z^2}{R^2}\right) \frac{dz}{z}. \tag{2.8}$$

Vamos a utilizar la identidad anterior para estimar el valor absoluto de $g(0)-g_T(0)$. La idea aquí es mostrar que el valor absoluto de dicha diferencia se puede hacer **arbitrariamente** pequeño.

En el semicírculo $C_+ = C \cap \{z \in \mathbb{C} : \Re(z) > 0\}$ se tiene que

$$|g(z) - g_T(z)| = \left| \int_T^\infty e^{-zt} f(t)\, dt \right| \leq M \int_T^\infty |e^{-zt}|\, dt = M \int_T^\infty e^{-\Re(z)t}\, dt = \frac{M e^{-\Re(z)T}}{\Re(z)}$$

donde M es alguna cota finita para la función f. Por otra parte, el factor $k(z) = e^{zT}\left(1 + \dfrac{z^2}{R^2}\right)\left(\dfrac{1}{z}\right)$ es tal que

$$\begin{aligned} |k(z)| &= |e^{zT}| \left| \left(1 + \frac{z^2}{R^2}\right)\left(\frac{1}{z}\right) \right| \\ &= e^{\Re(z)T} \left| \frac{1}{z} + \frac{z}{R^2} \right| \\ &= e^{\Re(z)T} \left| \frac{R}{z} + \frac{z}{R} \right| \left(\frac{1}{R}\right). \end{aligned}$$

Luego, si $z = Re^{i\theta}$ se sigue que $|k(z)| = e^{\Re(z)T}|e^{-i\theta} + e^{i\theta}| \cdot \frac{1}{R} = e^{\Re(z)T} \cdot 2\cos\theta \cdot \frac{1}{R}$ y en definitiva

$$|k(z)| = \left| e^{zT}\left(1 + \frac{z^2}{R^2}\right)\left(\frac{1}{z}\right) \right| = \frac{2\Re(z)}{R^2} e^{\Re(z)T}.$$

Así, el aporte a $g(0) - g_T(0)$ que se obtiene de C_+ está acotado en valor absoluto por

$$\left(\frac{1}{2\pi}\right)\left(\frac{M e^{-\Re(z)T}}{\Re(z)}\right)\left(\frac{2\Re(z)}{R^2} e^{\Re(z)T}\right)(\pi R) = \frac{M}{R}. \tag{2.9}$$

Para estimar la integral sobre $C_- = C \cap \{z \in \mathbb{C} : \Re(z) < 0\}$, estudiamos los términos en $g(z)$ y $g_T(z)$ por separado. Como g_T es una función entera, el principio de deformación de caminos asegura que integrar sobre C_- es equivalente a integrar sobre el semicírculo $D_- = \{z \in \mathbb{C} : |z| = R \quad \text{y} \quad \Re(z) < 0\}$. La acotación que en este caso se obtiene es

$$|g_T(z)| = \left| \int_0^T f(t) e^{-zT}\, dt \right| \leq M \int_0^T |e^{-zt}|\, dt \leq M \int_{-\infty}^T |e^{-zt}|\, dt = \frac{M e^{-\Re(z)T}}{|\Re(z)|}.$$

Así,

$$\left|\frac{1}{2\pi i}\int_{C_-} g_T(z)e^{zt}\left(1+\frac{z^2}{R^2}\right)\frac{dz}{z}\right| \leq \left(\frac{1}{2\pi}\right)\left(\frac{Me^{-\Re(z)T}}{|\Re(z)|}\right)(e^{\Re(z)T})\left(\frac{2|\Re(z)|}{R^2}\right)(\pi R)$$
$$= \frac{M}{R}. \tag{2.10}$$

Procedemos a estimar ahora la integral $\mathrm{I}(T,R) = \displaystyle\int_{C_-} g(z)e^{zT}\left(\frac{1}{z}+\frac{z}{R^2}\right)dz$. De la analiticidad de la función g sobre C_- se asegura la existencia de una constante $N = N(\delta, R)$ tal que

$$\left|g(z)\left(\frac{1}{z}+\frac{z}{R^2}\right)\right| \leq N$$

para cada $z \in C_-$. Luego, $\left|g(z)e^{zT}\left(\frac{1}{z}+\frac{z}{R^2}\right)\right| \leq Ne^{\Re(z)T}$. Si consideramos la sección C_1 del contorno C_- donde $\Re(z) \leq -\delta_1 < 0$, las estimaciones anteriores permiten concluir que en C_1 el integrando converge uniformemente a 0 cuando $T \to \infty$. En las porciones restantes del contorno se cumple que el integrando se encuentra acotado por la constante N y de aquí que $\lim_{T\to\infty} \mathrm{I}(T,R) = 0$.

Estamos listos para concluir nuestra demostración. Dado $\epsilon > 0$, hagamos $R = \frac{1}{\epsilon}$. Como $\mathrm{I}(T,R) \to 0$ cuando $T \to \infty$, existe $T_0 \in \mathbb{R}$ tal que $|\mathrm{I}(T,R)|$ es menor que $M \cdot \epsilon$ cuando $T > T_0$. De (2.8), (2.9) y (2.10) se obtiene que

$$|g(0) - g_T(0)| \leq M\cdot\epsilon + M\cdot\epsilon + M\cdot\epsilon = 3M\epsilon$$

cuando T es suficientemente grande y de ahí que

$$\lim_{T\to\infty} g_T(0) = g(0),$$

que es justamente lo que se deseaba establecer.

□

Presentamos a continuación dos importantes consecuencias del teorema de Newman que se acaba de demostrar:

a) **La integral** $I_\vartheta = \displaystyle\int_1^\infty \frac{\vartheta(x)-x}{x^2}\,dx$ **converge.** En efecto, si definimos $f(t) = e^{-t}\vartheta(e^t)-1$ se cumple que f es una función acotada y localmente integrable en $[0,\infty)$. Así, al tenerse que

$$\int_0^\infty f(t)e^{-st}\,dt = \int_1^\infty \frac{\vartheta(x)-x}{x^{s+2}}dx$$

es necesario detenernos a analizar con cierto detalle a la función

$$g(s) = \int_1^\infty \frac{\vartheta(x)-x}{x^{s+2}}dx$$

Para s en el semiplano $\sigma > 1$ definimos la función $\Phi(s)$ mediante la serie $\displaystyle\sum_{p\in\mathbf{P}} \frac{\log p}{p^s}$. De las estimaciones hechas en la prueba del teorema **2.4** se desprende de inmediato que la serie anterior converge absoluta y uniformemente en los subconjuntos compactos del semiplano $\sigma > 1$. El teorema de convergencia analítica implica entonces que $\Phi(s)$ representa una función analítica en $\sigma > 1$. Ahora bien, de

$$\begin{aligned}
\Phi(s+1) &= \sum_{p\in\mathbf{P}} \frac{\log p}{p^{s+1}} \\
&= \sum_{n=2}^{\infty} \frac{1}{n^{s+1}}(\vartheta(n)-\vartheta(n-1)) \\
&= -\sum_{n=2}^{\infty}\left(\frac{1}{n^{s+1}} - \frac{1}{(n-1)^{s+1}}\right)\vartheta(n-1) \\
&= (s+1)\sum_{n=2}^{\infty}\vartheta(n-1)\int_{n-1}^{n}\frac{dx}{x^{s+2}} \\
&= (s+1)\sum_{n=2}^{\infty}\int_{n-1}^{n}\frac{\vartheta(x)}{x^{s+2}}dx \\
&= (s+1)\int_1^\infty \frac{\vartheta(x)}{x^{s+2}}dx
\end{aligned}$$

se tiene que

$$\frac{\Phi(s+1)}{s+1} - \frac{1}{s} = \int_1^\infty \frac{\vartheta(x)}{x^{s+2}}dx - \int_1^\infty \frac{x}{x^{s+2}}dx = g(s).$$

Con ayuda de estas igualdades probaremos ahora que la función g admite una extensión analítica en $\sigma \geq 0$. Una vez que establezcamos eso podremos aplicar libremente el teorema de Newman a las funciones g y $\vartheta(e^t)$ a fin de obtener el resultado en cuestión. Veamos.

De la cadena de igualdades en (2.4) se sabe que

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{p\in\mathbf{P}} \frac{\log p}{p^s - 1} = \sum_{p\in\mathbf{P}} \frac{(p^s-1)\log p + \log p}{p^s(p^s-1)}$$

y de ahí que

$$-\frac{\zeta'(s)}{\zeta(s)} = \Phi(s) + \sum_{p\in\mathbf{P}} \frac{\log p}{p^s(p^s-1)}.$$

Afirmamos ahora que la serie $\displaystyle\sum_{p\in\mathbf{P}} \frac{\log p}{p^s(p^s-1)}$ converge absolutamente siempre que $\sigma > \dfrac{1}{2}$. En efecto, si p es suficientemente grande se cumple que $9p^\sigma > 10$ y por tanto, para dichos p, se observa que

$$|p^s - 1| \geq |p^s| - 1 = p^\sigma - 1 > \frac{p^\sigma}{10}.$$

De las acotaciones anteriores se sigue que

$$\left|\frac{\log p}{p^s(p^s-1)}\right| = \frac{\log p}{|p^s||p^s-1|} < \frac{10\log p}{p^{2\sigma}}$$

para los primos *suficientemente* grandes. La convergencia de la serie $\displaystyle\sum_{p\in\mathbf{P}} \frac{\log p}{p^{2\sigma}}$ implica entonces la convergencia absoluta de la serie $\displaystyle\sum_{p\in\mathbf{P}} \frac{\log p}{p^s(p^s-1)}$. Del teorema de convergencia analítica se desprende entonces que

$$F(s) = \sum_{p\in\mathbf{P}} \frac{\log p}{p^s(p^s-1)}$$

representa una función analítica en el semiplano $\sigma > \frac{1}{2}$. Luego, de

$$\Phi(s) - \frac{1}{s-1} = -\frac{\zeta'(s)}{\zeta(s)} - \frac{1}{s-1} - F(s)$$

colegimos que la función $g(s) = \dfrac{\Phi(s+1)}{s+1} - \dfrac{1}{s}$ puede extenderse analíticamente al semiplano $\sigma \geq 0$.

El teorema de Newman nos permite concluir entonces que

$$g(0) = \int_0^\infty f(t)\,dt = \int_0^\infty [e^{-t}\vartheta(e^t) - 1]\,dt = \int_1^\infty \frac{\vartheta(x) - x}{x^2}\,dx.$$

b) **La convergencia de la integral I_ϑ implica que** $\vartheta(x) \sim x$. Para convencernos de esto notamos en primer lugar que la relación $\vartheta(x) \sim x$ es equivalente a la conjunción del siguiente par de afirmaciones

(**1**) El conjunto $\{x > 0 : \vartheta(x) \geq \lambda x\}$ está acotado siempre que $\lambda > 1$;

(**2**) El conjunto $\{x > 0 : \vartheta(x) \leq \lambda x\}$ está acotado siempre que $\lambda \in (0, 1)$.

Así, a fin de establecer la veracidad de $\vartheta(x) \sim x$ lo que haremos será probar, por contradicción, el cumplimiento de **1** y **2**.

Si la afirmación en **1** es falsa, existe $\lambda > 1$ tal que $\vartheta(x)/x \geq \lambda$ para x arbitrariamente grandes. Puesto que ϑ es no decreciente se sigue que $\vartheta(t) \geq \vartheta(x) \geq \lambda x$ siempre que $x \leq t \leq \lambda x$ y por lo tanto

$$\int_x^{\lambda x} \frac{\vartheta(t) - t}{t^2}\,dt \geq \int_x^{\lambda x} \frac{\lambda x - t}{t^2}\,dt.$$

Haciendo el cambio de variable $t = xu$ en la segunda integral de la línea previa se obtiene que

$$\int_x^{\lambda x} \frac{\vartheta(t) - t}{t^2}\,dt \geq \int_1^{\lambda} \frac{\lambda - u}{u^2}\,du > 0,$$

lo que entra en contradicción con el criterio de Cauchy para la convergencia de integrales impropias.

De manera similar, si $\vartheta(x) \leq \lambda x$ para valores arbitrariamente grandes de x y $\lambda < 1$ entonces

$$\int_{\lambda x}^{x} \frac{\vartheta(x) - t}{t^2}\, dt \leq \int_{\lambda x}^{x} \frac{\lambda x - t}{t^2}\, dt = \int_{\lambda}^{1} \frac{\lambda - u}{u^2}\, du < 0$$

lo que contradice también al criterio de Cauchy mencionado recientemente.

La discusión precedente indica que $\vartheta(x) \sim x$ y lo mencionado en la discusión debajo de la prueba de la proposición **1.15** nos permite concluir entonces que

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

§4. Hemos presentado la ruta más corta que se conoce hacia la prueba de la notable conjetura de Gauss-Legendre. La primera demostración de dicho resultado fue presentada de modo independiente y casi simultáneamente en 1896 por J. Hadamard y C. J. de la Vallée Poussin. Comenta Donald J. Newman en [12] que

> ... Las pruebas originales de Hadamard y de de la Vallée Poussin se basaron en la no anulación de $\zeta(z)$ en $\Re(z) \geq 1$, pero requerían además de molestos estimados para $\zeta(z)$ en el infinito. La razón de eso es que las fórmulas para los coeficientes de las series de Dirichlet involucran integrales sobre contornos infinitos (a diferencia de lo que ocurre con las series de potencias) y por tanto su evaluación efectiva requiere de estimados en ∞.
>
> Las pruebas modernas, debidas principalmente a Wiener y a Ikehara, eluden la necesidad de estimaciones en ∞ y dependen tan sólo del correspondiente resultado de no anulación para $\zeta(z)$, pero involucran algunos resultados sobre transformadas de Fourier.

El aporte principal de D. J. Newman fue retomar los métodos de integración en contornos (para evitar el uso del Análisis de Fourier) y restringir su atención sólo a contornos finitos (para evitar las mencionadas estimaciones en el infinito). La demostración de Newman puede considerarse entonces como una mejora natural al tratamiento presentado por Norbert Wiener y S. Ikehara.

El trabajo de Wiener es digno de mención también pues él fue de los primeros en hacer depender la demostración de la conjetura de Gauss-Legendre únicamente en la no anulación de la función zeta en $\sigma = 1$. De hecho, Wiener mostró que el cumplimiento de la conjetura de Gauss-Legendre es completamente equivalente a la no anulación de la función ζ sobre dicha recta.

La denominación estándar para la conjetura otrora conocida como de Gauss-Legendre es TEOREMA DEL NÚMERO PRIMO. De ahora en adelante, será así como nos referimos a dicho resultado. Vamos a mostrar ahora la deducción a la Wiener-Ikehara del TEOREMA DEL NÚMERO PRIMO. El ingrediente clave de este enfoque es un teorema similar al que se tiene en **2.6**.

Antes de pasar a la prueba de dicho resultado vamos a enunciar y demostrar un lema técnico al que confinaremos ciertos cálculos que serán de utilidad en lo sucesivo.

Lema 2.7.

a) Sea g una función analítica en la cerradura del semiplano $\sigma > 0$ y λ una constante real positiva. Entonces $g(\sigma + it)$ converge uniformemente a $g(it)$ en $[-\lambda, \lambda]$ cuando $\sigma \to 0$.

b) Para cada $\alpha \in (0, \infty)$ se tiene que

$$\frac{1}{2}\int_{-2\alpha}^{2\alpha}\left(1 - \frac{|t|}{2\alpha}\right)e^{ixt}\,dt = \frac{\operatorname{sen}^2 \alpha x}{\alpha x^2}.$$

c) Si $\Re(s) > 1$, la integral $I_s = \displaystyle\int_0^\infty e^{-(s-1)x}\,dx$ converge a $\dfrac{1}{s-1}$.

d) La integral $I = \displaystyle\int_0^\infty \frac{\operatorname{sen}^2 v}{v^2}\, dv$ converge a $\dfrac{\pi}{2}$.

Prueba.

a) Sin pérdida de generalidad podemos suponer que $\lambda = 1$. Sea $\epsilon > 0$. Para $t \in [-1, 1]$ hagamos $\theta_t = \left\{s : |s - it| < \dfrac{\delta_t}{2}\right\}$ en donde δ_t es tal que $|s - it| < \delta_t$ implica que $|g(s) - g(it)| < \dfrac{\epsilon}{2}$. Sea $\{\theta_{t_1}, \ldots, \theta_{t_n}\}$ una cubierta finita del segmento de recta que une a $-i$ con i. Sea $\delta = \min\limits_{1\le j\le n}\{\delta_{t_j}/2\}$. Si $0 \le \sigma \le \delta$ y para $t \in [-1, 1]$, t_j es tal que $it \in \theta_{t_j}$ se sigue que

$$|\sigma + it - it_j| = \sqrt{\sigma^2 + (t - t_j)^2} < \sqrt{\delta^2 + \left(\frac{\delta_{t_j}}{2}\right)^2} \le \sqrt{\frac{\delta_{t_j}^2}{2}} = \frac{\delta_{t_j}}{\sqrt{2}} < \delta_{t_j}$$

y por tanto

$$|g(\sigma + it) - g(it)| \le |g(\sigma + it) - g(it_j)| + |g(it_j) - g(it)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

b) Empezamos por encontrar una fórmula general para la integral indefinida $\displaystyle\int (a + bt)e^{ixt}\, dt$. Dado que

$$\int te^{ixt}\, dt = \frac{e^{ixt}}{x^2} + \frac{e^{ixt}t}{ix} + C,$$

se sigue que $\displaystyle\int (a + bt)e^{ixt}\, dt = \frac{a}{ix}e^{ixt} + \frac{b}{ix}(e^{ixt}t) + \frac{b}{x^2}e^{ixt} + C$. Ahora bien, de la descomposición

$$\begin{aligned}\int_{-2\alpha}^{2\alpha}\left(1 - \frac{|t|}{2\alpha}\right)e^{ixt}\, dt &= \left[\int_{-2\alpha}^{0} + \int_0^{2\alpha}\right]\left(1 - \frac{|t|}{2\alpha}\right)e^{ixt}\, dt \\ &= \int_{-2\alpha}^{0}\left(1 + \frac{t}{2\alpha}\right)e^{ixt}\, dt + \int_0^{2\alpha}\left(1 - \frac{t}{2\alpha}\right)e^{ixt}\, dt\end{aligned}$$

y la fórmula previamente obtenida se sigue que

$$\int_{-2\alpha}^{2\alpha}\left(1 - \frac{|t|}{2\alpha}\right)e^{ixt}\, dt = \frac{1}{\alpha x^2} - \frac{e^{ix(2\alpha)} + e^{-ix(2\alpha)}}{2\alpha x^2} = \frac{2 - 2\cos 2\alpha x}{2\alpha x^2} = \frac{2\operatorname{sen}^2 \alpha x}{\alpha x^2},$$

tal como se había previsto.

c) En este caso se tiene que

$$\int_0^T e^{-(s-1)x}\,dx = \frac{1}{s-1} - \frac{e^{-(s-1)T}}{s-1}.$$

La conclusión deseada es ahora una consecuencia inmediata del hecho que $e^{-(s-1)T} \to 0$ cuando $T \to \infty$.

d) Consideremos la función $F : \mathbb{C} \to \mathbb{C}$ definida como

$$F(z) = \begin{cases} -2 & \text{si } z = 0 \\ \frac{e^{2iz}-2iz-1}{z^2} & \text{en otro caso.} \end{cases}$$

Si denotamos ahora con L_R al intervalo $(-R, R)$ y con C_R a la intersección del semiplano $\Im(z) \geq 0$ con el círculo de radio R y centro en el origen, el teorema de Cauchy nos permite asegurar que

$$\int_{L_R \cup C_R} F(z)\,dz = 0.$$

La identidad anterior es cierta para cada $R > 0$, pues la función F es entera. Por otra parte, de

$$\int_{C_R} F(z)\,dz = \int_{C_R} \frac{e^{2iz}-1}{z^2}\,dz - 2i\int_{C_R} \frac{1}{z}\,dz = \int_{C_R} \frac{e^{2iz}-1}{z^2}\,dz + 2\pi$$

y del hecho que

$$\left|\int_{C_R} \frac{e^{2iz}-1}{z^2}\,dz\right| \leq (\pi R) \max_{z \in C_R} \frac{|e^{2iz}-1|}{|z^2|} \leq (\pi R)\left(\frac{2}{R^2}\right) = \frac{2\pi}{R}$$

se sigue que $\lim\limits_{R\to\infty} \int_{C_R} F(z)\,dz = 2\pi$ y por tanto

$$\begin{aligned} -2\pi &= \lim_{R\to\infty} \int_{L_R} F(z)\,dz \\ &= \lim_{R\to\infty} \int_{-R}^{R} \frac{e^{2ix}-1-2ix}{x^2}\,dx \\ &= \int_{-\infty}^{\infty} \frac{e^{2ix}-1-2ix}{x^2}\,dx. \end{aligned}$$

Al extraer la parte real de la integral anterior se llega a que

$$\pi = \int_{-\infty}^{\infty} \frac{\operatorname{sen}^2 x}{x^2}\, dx = 2I.$$

Luego, $I = \pi/2$ y la demostración termina. □

Pasamos ahora al análisis de la ruta alternativa hacia el Teorema del Número Primo. La similitud con el teorema **2.5** es más que evidente.

Teorema 2.8. (N. Wiener - S. Ikehara) Sea $A(x)$ una función no negativa y no decreciente definida en el intervalo $(0, \infty)$. Sea $s = \sigma + it$. Suponga que

$$f(s) = \int_0^{\infty} A(x)e^{-xs}\, dx$$

converge para $\sigma > 1$ y que $g(s) = f(s) - \dfrac{1}{s-1}$ se puede extender analíticamente a la cerradura del semiplano $\sigma > 1$. Se cumple entonces que

$$\lim_{x\to\infty} e^{-x}A(x) = 1.$$

Prueba. La prueba es en dos partes. Poniendo $B(x) = e^{-x}A(x)$ probaremos primero que para $\lambda > 0$ se tiene que

$$\lim_{y\to\infty} \int_{-\infty}^{\lambda y} B\left(y - \frac{v}{\lambda}\right) \frac{\operatorname{sen}^2 v}{v^2}\, dv = \pi. \tag{2.11}$$

Usaremos después (2.11) para probar que

$$\lim_{x\to\infty} B(x) = 1. \tag{2.12}$$

Primera parte. En la luz del apartado c del lema **2.7** tenemos que para $\sigma > 1$,

$$\begin{aligned} g(s) &= f(s) - \frac{1}{s-1} \\ &= \int_0^{\infty} A(x)e^{-sx}\, dx - \int_0^{\infty} e^{-(s-1)x}\, dx \\ &= \int_0^{\infty} [B(x) - 1]e^{-(s-1)x}\, dx. \end{aligned}$$

Ahora bien, para λ y ϵ positivos y

$$g_\epsilon(t) = g(1 + \epsilon + it)$$

se cumple que

$$\frac{1}{2}\int_{-2\lambda}^{2\lambda} g_\epsilon(t)\left(1 - \frac{|t|}{2\lambda}\right)e^{iyt}\,dt = \frac{1}{2}\int_{-2\lambda}^{2\lambda}\left(1 - \frac{|t|}{2\lambda}\right)e^{iyt}\left(\int_0^\infty [B(x) - 1]e^{-(\epsilon+it)x}\,dx\right)dt. \tag{2.13}$$

Al aplicar el inciso b del lema y el teorema de Fubini al lado derecho de (2.13) llegamos a que

$$\begin{aligned}\frac{1}{2}\int_{-2\lambda}^{2\lambda} g_\epsilon(t)\left(1 - \frac{|t|}{2\lambda}\right)e^{iyt}\,dt &= \int_0^\infty [B(x) - 1]e^{-\epsilon x}\left(\int_{-2\lambda}^{2\lambda}\frac{1}{2}e^{i(y-x)t}\left[1 - \frac{|t|}{2\lambda}\right]dt\right)dx \\ &= \int_0^\infty [B(x) - 1]e^{-\epsilon x}\frac{\operatorname{sen}^2\lambda(y-x)}{\lambda(y-x)^2}\,dx.\end{aligned}$$

Luego, si en ambos extremos de la identidad anterior hacemos que $\epsilon \to 0$ se obtiene

$$\frac{1}{2}\int_{-2\lambda}^{2\lambda} g(1 + it)\left(1 - \frac{|t|}{2\lambda}\right)e^{iyt}\,dt = \lim_{\epsilon\to 0}\int_0^\infty [B(x) - 1]e^{-\epsilon x}\frac{\operatorname{sen}^2\lambda(y-x)}{\lambda(y-x)^2}\,dx. \tag{2.14}$$

Por otro lado, el teorema de convergencia dominada de Lebesgue implica que

$$\lim_{\epsilon\to 0}\int_0^\infty e^{-\epsilon x}\frac{\operatorname{sen}^2\lambda(y-x)}{\lambda(y-x)^2}\,dx = \int_0^\infty \frac{\operatorname{sen}^2\lambda(y-x)}{\lambda(y-x)^2}\,dx. \tag{2.15}$$

Este último resultado, en conjunción con el teorema de convergencia monótona, permiten afirmar que

$$\lim_{\epsilon\to 0}\int_0^\infty B(x)e^{-\epsilon x}\frac{\operatorname{sen}^2\lambda(y-x)}{\lambda(y-x)^2}\,dx = \int_0^\infty B(x)\frac{\operatorname{sen}^2\lambda(y-x)}{\lambda(y-x)^2}\,dx. \tag{2.16}$$

De (2.14), (2.15) y (2.16) se tiene entonces que

$$\frac{1}{2}\int_{-2\lambda}^{2\lambda} g(1 + it)\left(1 - \frac{|t|}{2\lambda}\right)e^{iyt}\,dt = \int_0^\infty B(x)\frac{\operatorname{sen}^2\lambda(y-x)}{\lambda(y-x)^2}\,dx - \int_0^\infty \frac{\operatorname{sen}^2\lambda(y-x)}{\lambda(y-x)^2}\,dx.$$

Si en la igualdad previa se hace $y \to \infty$, el lema de Riemann-Lebesgue asegura que el lado izquierdo tiende a cero. Esto implica a su vez que

$$\begin{aligned}\lim_{y\to\infty}\int_{-\infty}^{\lambda y} B\left(y-\frac{v}{\lambda}\right)\frac{\operatorname{sen}^2 v}{v^2}\,dv &= \lim_{y\to\infty}\int_0^{\infty} B(x)\frac{\operatorname{sen}^2 \lambda(y-x)}{\lambda(y-x)^2}\,dx\\ &= \lim_{y\to\infty}\int_0^{\infty} \frac{\operatorname{sen}^2 \lambda(y-x)}{\lambda(y-x)^2}\,dx\\ &= \lim_{y\to\infty}\int_{-\infty}^{\lambda y} \frac{\operatorname{sen}^2 v}{v^2}\,dv\\ &= \pi,\end{aligned}$$

tal como se había anunciado en (2.11).

Segunda parte. Para probar (2.12) basta con mostrar que

$$\limsup_{x\to\infty} B(x) \le 1 \quad \text{y} \quad \liminf_{x\to\infty} B(x) \ge 1.$$

Empezamos con la desigualdad de más a la izquierda. Si a y λ son dos números positivos y $y > a/\lambda$, la identidad en (2.11) asegura que para $\lambda > 0$

$$\limsup_{y\to\infty}\int_{-a}^{a} B\left(y-\frac{v}{\lambda}\right)\frac{\operatorname{sen}^2 v}{v^2}\,dv \le \pi$$

pues el integrando es no negativo. Además, al ser $A(u) = B(u)e^u$ una función no decreciente se sigue que para $v \in [-a, a]$,

$$e^{y-a/\lambda}B\left(y-\frac{a}{\lambda}\right) \le e^{y-v/\lambda}B\left(y-\frac{v}{\lambda}\right)$$

y por lo tanto

$$B\left(y-\frac{v}{\lambda}\right) \ge B\left(y-\frac{a}{\lambda}\right)e^{(v-a)/\lambda} \ge B\left(y-\frac{a}{\lambda}\right)e^{-2a/\lambda}.$$

Tenemos así que

$$\pi \ge \limsup_{y\to\infty}\int_{-a}^{a} B\left(y-\frac{v}{\lambda}\right)\frac{\operatorname{sen}^2 v}{v^2}\,dv \ge \limsup_{y\to\infty} B\left(y-\frac{a}{\lambda}\right)e^{-2a/\lambda}\int_{-a}^{a}\frac{\operatorname{sen}^2 v}{v^2}\,dv. \tag{2.17}$$

Notemos ahora que para cada a y λ fijos se cumple que $\limsup_{y\to\infty} B\left(y - \frac{a}{\lambda}\right)$ es igual a $\limsup_{y\to\infty} B(y)$. Luego, si $a \to \infty$ y $\lambda \to \infty$ de tal manera que $a/\lambda \to 0$ se sigue de (2.17) que

$$\left(\int_{-\infty}^{\infty} \frac{\operatorname{sen}^2 v}{v^2}\, dv\right) \limsup_{y\to\infty} B(y) \leq \pi.$$

De esto último y del apartado d del lema **2.7** se concluye que

$$\limsup_{y\to\infty} B(y) \leq \frac{\pi}{\int_{-\infty}^{\infty} \frac{\operatorname{sen}^2 v}{v^2}\, dv} = \frac{\pi}{\pi} = 1.$$

Pasamos ahora a la prueba de la cota inferior para $\liminf_{x\to\infty} B(x)$. A fin de establecer dicho punto empezamos por notar que el hecho de que la función B sea no negativa con límite superior no mayor a 1 implica de inmediato que B está acotada por alguna constante C sobre todo el intervalo $(0, \infty)$. Luego, si a y λ son números positivos tales que $y < a/\lambda$ se sigue que

$$\int_{-\infty}^{\lambda y} B\left(y - \frac{v}{\lambda}\right) \frac{\operatorname{sen}^2 v}{v^2}\, dv \leq C\left[\int_{-\infty}^{-a} + \int_{a}^{\infty}\right] \frac{\operatorname{sen}^2 v}{v^2}\, dv + \int_{-a}^{a} B\left(y - \frac{v}{\lambda}\right) \frac{\operatorname{sen}^2 v}{v^2}\, dv. \tag{2.18}$$

Como antes, para $-a \leq v \leq a$, tenemos que

$$e^{y - \frac{a}{\lambda}} B\left(y - \frac{v}{\lambda}\right) \leq e^{y + a/\lambda} B\left(y + \frac{a}{\lambda}\right)$$

y por tanto

$$\int_{-a}^{a} B\left(y - \frac{v}{\lambda}\right) \frac{\operatorname{sen}^2 v}{v^2}\, dv \leq B\left(y + \frac{a}{\lambda}\right) e^{2a/\lambda} \int_{-a}^{a} \frac{\operatorname{sen}^2 v}{v^2}\, dv. \tag{2.19}$$

Ahora bien, al tenerse que $\liminf_{y\to\infty} B\left(y + \frac{a}{\lambda}\right) = \liminf_{y\to\infty} B(y)$, los resultados en (2.11), (2.18) y (2.19) permiten concluir que

$$\begin{aligned} \pi &\leq C\left[\int_{-\infty}^{-a} + \int_{a}^{\infty}\right] \frac{\operatorname{sen}^2 v}{v^2}\, dv + \liminf_{y\to\infty} \int_{-a}^{a} B\left(y - \frac{v}{\lambda}\right) \frac{\operatorname{sen}^2 v}{v^2}\, dv \\ &\leq C\left[\int_{-\infty}^{-a} + \int_{a}^{\infty}\right] \frac{\operatorname{sen}^2 v}{v^2}\, dv + \left(e^{2a/\lambda} \int_{-a}^{a} \frac{\operatorname{sen}^2 v}{v^2}\, dv\right) \liminf_{y\to\infty} B(y). \end{aligned}$$

Si en la última expresión de la anterior cadena de desigualdades hacemos $a \to \infty$, $\lambda \to \infty$ y $a/\lambda \to 0$ llegamos a que $\pi \leq \pi \liminf_{y\to\infty} B(y)$, con lo que la prueba termina.

□

La deducción del TEOREMA DEL NÚMERO PRIMO a partir del teorema de Wiener-Ikehara es directa ahora. En el §**2** de este capítulo se estableció una importante identidad integral para la función $-\dfrac{\zeta'(s)}{\zeta(s)}$ en el semiplano $\sigma > 1$. Dicha identidad asegura que

$$-\frac{\zeta'(s)}{\zeta(s)} = \int_{1^-}^{\infty} \frac{d\psi(x)}{x^s} = s\int_1^{\infty} \frac{\psi(x)}{x^{s+1}}\,dx = s\int_0^{\infty} e^{-xs}\psi(e^x)\,dx \tag{2.20}$$

siempre que $s \in \sigma > 1$. Por otro lado, la proposición **2.3** indica que en el semiplano $\sigma > 0$ la función ζ puede escribirse como

$$1 + \frac{h(s)}{s-1} = \frac{H(s)}{s-1}$$

donde $h(s)$ es analítica en $\sigma > 0$ y $H(1) = 1$. Luego,

$$-\frac{\zeta'(s)}{\zeta(s)} = -\frac{d}{ds}\log\zeta(s) = \frac{1}{s-1} + g(s) \tag{2.21}$$

donde $g(s) = -H'(s)/H(s)$. De la igualdad en (2.21) se obtiene que g es una función analítica en la cerradura del semiplano $\sigma > 1$. Las relaciones en (2.20) y (2.21) nos permiten asegurar entonces que

$$\int_0^{\infty} \psi(e^x)e^{-xs}\,dx = \frac{1}{s(s-1)} + \frac{g(s)}{s}.$$

Puesto que $g(s)$ es analítica en la cerradura de $\sigma > 1$, se tienen todas las hipótesis del teorema de Wiener-Ikehara. Por tanto,

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

y la demostración termina.

§**5.** En la parte final del apartado anterior se presentó una prueba del TEOREMA DEL NÚMERO PRIMO en base al teorema de Wiener-Ikehara y a la no

anulación de la función zeta en la recta $\sigma = 1$. A continuación mostraremos que la no anulación de la función ζ sobre dicha recta se puede obtener como consecuencia del TEOREMA DEL NÚMERO PRIMO.

Para $s \in \sigma > 1$ consideremos la función

$$F(s) = -\frac{\zeta'(s)}{s\zeta(s)} - \frac{1}{s-1} = \int_1^\infty \frac{\psi(x) - x}{x^{s+1}}\, dx.$$

Esta función F es analítica en $\sigma > 0$ salvo en los ceros de la función ζ. Luego, para $\epsilon > 0$, el TEOREMA DEL NÚMERO PRIMO asegura la existencia de un real x_0 tal que $|\psi(x) - x| < \epsilon x$ cuando $x > x_0$. Así, para $s \in \sigma > 1$ se cumple que

$$|F(s)| < \int_1^{x_0} \frac{|\psi(x) - x|}{x^2}\, dx + \int_{x_0}^\infty \frac{\epsilon}{x^\sigma}\, dx < \int_1^{x_0} \frac{|\psi(x) - x|}{x^2}\, dx + \frac{\epsilon}{\sigma - 1}$$

donde $\sigma = \Re(s)$. Si multiplicamos ambos extremos de la cadena de desigualdades en la línea anterior por $(\sigma - 1)$ y tomamos límite cuando $\sigma \to 1^+$ se obtiene que $\lim_{\sigma \to 1^+}(\sigma - 1)F(\sigma + it) = 0$ para cada t fijo. Por otra parte, si $1 + it$ es un cero de la función ζ, entonces el límite de $(\sigma - 1)F(\sigma + it)$ cuando $\sigma \to 1^+$ tiene que ser igual al residuo de $F(s)$ en el polo simple $s = 1 + it$. Como dicho residuo no puede ser cero, se ha entablado una contradicción con el valor obtenido para el límite en cuestión.

De todo lo anterior se colige que la no anulación de la función ζ sobre la línea $\sigma = 1$ es un hecho equivalente al TEOREMA DEL NÚMERO PRIMO. Otra equivalencia notable se presenta en el siguiente

Escolio 2.9. $\pi(x) \sim \dfrac{x}{\log x}$ si y sólo si $p_n \sim n \log n$.

Prueba. De la relación $\pi(x) \sim \dfrac{x}{\log x}$ se desprende de inmediato que

$$\lim_{x \to \infty}(\log \pi(x) + \log\log x - \log x) = 0.$$

Esta igualdad implica a su vez que

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

y por consiguiente

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

Si en esta última línea hacemos $x = p_n$, llegamos a que $\dfrac{n\log n}{p_n} \to 1$ cuando $n \to \infty$.

Recíprocamente, si $p_n \sim n\log n$ se sigue que

$$\frac{p_n}{p_{n+1}} \sim \frac{n\log n}{(n+1)\log(n+1)} \sim 1.$$

Para cada $n \in \mathbb{N}$, fíjese $x(n)$ de tal manera que $p_n \leq x(n) < p_{n+1}$. Se cumple entonces que $x(n) \sim p_n$ y por tanto $x(n) \sim n\log n$. De esta última relación se desprende que $\log n + \log\log n \sim \log x(n)$ y de ahí que $\log n \sim \log x(n)$. Luego

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

y la prueba termina.

□

Para concluir mencionaremos algunas conexiones entre el Teorema del Número Primo y ciertos puntos de la discusión efectuada en el capítulo anterior.

La primera de ellas es con relación al problema de la representación de **P** mediante polinomios. Veamos. Es claro, por ejemplo, que la imagen de $\mathbb{N}$ bajo el polinomio $2x + 1$ contiene, prácticamente, a todo **P**. Una vez que se ha notado esto, hay una pregunta que surge de modo natural: ¿existe un polinomio cuadrático f de tal manera que $f(\mathbb{N}) \supseteq \mathbf{P}$?

Una manera de responder a esta interrogante es utilizando la divergencia de la serie de los recíprocos de los números primos. Si suponemos que $f(n) = n^2 + bn + c$ es un polinomio tal que $f(\mathbb{N}) \supseteq \mathbf{P}$ entonces

$$\sum_{p\in\mathbf{P}} \frac{1}{p} \leq \sum_{n\in\mathbb{N}} \frac{1}{f(n)}[f(n) > 0]. \tag{2.22}$$

Por otra parte, al ser $f(n) \sim n^2$ se tiene que

$$\sum_{n\in\mathbb{N}} \frac{1}{f(n)}[f(n) > 0]$$

es una serie convergente. Esto último contradice, en vista de la desigualdad (2.23), la divergencia de la serie $\sum_{p\in\mathbf{P}} \frac{1}{p}$.

Otra forma de abordar el problema es precisamente con ayuda del Teorema del Número Primo. Sean f un polinomio cuadrático tal que $f(\mathbb{N}) \supseteq \mathbf{P}$ y A una constante (positiva) para la cual se cumple que $f(n) \leq An^2$ para cada $n \in \mathbb{N}$. De lo anterior se tiene que, para $x > 1$, el intervalo $[1, x]$ contiene no más de $\sqrt{\frac{x}{A}}$ números primos. La estimación anterior implica que

$$\lim_{x\to\infty} \frac{\pi(x)\log x}{x} = 0$$

lo que es claramente absurdo en la luz del Teorema del Número Primo.

Estos argumentos pueden modificarse para mostrar que si

$$\{f_1, \ldots, f_n\} \subseteq \mathbb{Z}[x]$$

y $\deg[f_i] \geq 2$ para cada $i \in \{1, \ldots, n\}$ entonces no puede ser el caso que

$$\mathbf{P} \subseteq \bigcup_{i=1}^{n} f_i(\mathbb{N}).$$

La segunda conexión tiene que ver con la desigualdad en (1.5), a saber

$$\prod_{\substack{p\leq n,\\ p\in\mathbf{P}}} p < 4^n.$$

Esta desigualdad es uno de los ingredientes esenciales en la prueba de Erdős del postulado de Bertrand.

Como bien sabemos ya, el Teorema del Número Primo es equivalente a que $\lim_{x\to\infty} \frac{\vartheta(x)}{x} = \lim_{x\to\infty} \frac{\sum_{p\leq x} \log p}{x} = 1$. Lo anterior implica a su vez que

$$\lim_{n\to\infty} \left(\prod_{\substack{p\leq n,\\ p\in\mathbf{P}}} p \right)^{\frac{1}{n}} = e.$$

Luego, para cada $\epsilon > 0$ se tiene que $\prod_{\substack{p\leq n,\\ p\in\mathbf{P}}} p < (e+\epsilon)^n$ siempre que n es suficientemente grande. Esto indica que el número 4 en el extremo derecho de la desigualdad (1.5) se encuentra un tanto distante de la mejor constante posible para la desigualdad en cuestión.

Capítulo 3

Primos en progresiones aritméticas

§**1.** El objetivo de este apartado es estudiar un importante teorema debido a Johann Peter Gustav Lejeune Dirichlet. La prueba de dicho resultado fue el tema central de una memoria de Dirichlet publicada en 1837. Muchos estudiosos de la Aritmética consideran a ese escrito como el principal referente en el nacimiento de la Teoría Analítica de Números. Que esto sea así se debe, básicamente, a la notabilidad del teorema y a lo novedoso de las técnicas introducidas por Dirichlet en su demostración. Presentamos a continuación, sin mayor preámbulo, el enunciado del teorema a cuyo análisis dedicaremos nuestro esfuerzo en lo sucesivo:

Teorema 3.1. Sean a y m dos números naturales coprimos. Se tiene entonces que en la progresión aritmética

$$a, a+m, a+2m, a+3m, \ldots$$

aparecen infinitos números primos.

De acuerdo con L. E. Dickson (ver [6], página 415), Euler demostró el resultado en el caso en que $a = 1$. Dickson agrega que A. M. Legendre creyó haber tenido una prueba para otras elecciones generales de a y m,

pero que su argumento se basaba en un lema que posteriormente sería encontrado erróneo.

Para poner las aportaciones de Dirichlet en contexto, abordaremos a continuación la labor de presentar pruebas para varias instancias especiales del teorema del momento.

Ejemplo 3.2. Si $m = 2$ y $a = 2k + 1$ para algún $k \in \mathbb{N}$, la progresión aritmética toma la forma

$$2k + 1, 2(k + 1) + 1, 2(k + 2) + 1, \ldots$$

Luego, es claro que la progresión consiste de todos los números impares positivos salvo los k primeros. La veracidad del aserto es consecuencia entonces de la sola infinitud del conjunto de primos.

□

Ejemplo 3.3. Si $m = 3$ y $a = 2$, la progresión aritmética en cuestión consiste de los números

$$2, 5, 8, 11, \ldots$$

Para mostrar que hay infinitos primos dentro de la progresión, la técnica a emplear será constructiva. A saber, probaremos que dado un conjunto $A_k = \{p_1, \ldots, p_k\}$ de primos en la progresión, siempre podemos detectar otro primo (en la progresión aritmética) que coincida con ningún elemento de A_k. El punto clave aquí es la consideración, una vez que A_k se ha fijado, de la expresión

$$a_k = 3(p_1 \cdots p_k) - 1.$$

Claramente, el número determinado por dicha receta pertenece a la progresión aritmética que se esta considerando. Lo que resta ver es que la expresión nos permite derivar, de algún modo, un número primo en la progresión no contenido en A_k. Veamos. Al ser a_k un número positivo distinto de 1 se sigue que a_k es un producto de primos congruentes con 1 ó 2 en

módulo 3. Dado que el producto de números congruentes con 1 módulo 3 es otro número de ese tipo, se sigue que a_k siempre posee un factor primo, q_k, congruente con 2 en módulo 3. La observación crucial ahora es que q_k no puede ser un elemento de A_k, pues de serlo se tendría a q_k como divisor de $3(p_1 \cdots p_k)$ y de $3(p_1 \cdots p_k) - 1$ y por tanto de $3(p_1 \cdots p_k) - [3(p_1 \cdots p_k) - 1] = 1$, lo que es ciertamente absurdo. El argumento anterior ilustra cómo obtener más primos en la progresión aritmética $\{3k + 2\}_{k \in \mathbf{Z}^+}$ una vez que se ha identificado a uno o más de ellos.

□

Ejemplo 3.4. Consideremos ahora el caso en que $m = 4$. Si a es algún número natural coprimo con m se tiene que a es congruente con 1 ó 3 en módulo 4. Luego, estudiar la aparición de primos en una progresión arimética con diferencia común igual a 4 se reduce a estudiar solamente los casos en que $a = 1$ ó $a = 3$. La prueba de la aparición de infinitos primos en $\{4k + 3\}_{k \in \mathbb{Z}^+}$ queda como la hecha en el ejemplo **3.3**, *mutatis mutandis*. Discutiremos ahora la prueba del resultado en el caso en que $a = 1$. El deseo de hacer nuestro argumento totalmente autocontenido justifica la siguiente digresión en el tema de los residuos cuadráticos.

Recordemos[1] que a es un residuo cuadrático módulo un primo impar p si existe un número entero x tal que $x^2 \equiv a$ mód p. Una observación importante que se tiene es que en el conjunto $A = \{1, 2, \ldots, p-1\}$ hay tantos residuos cuadráticos como no cuadráticos (como es de esperarse, a es un residuo no cuadrático si a no es un residuo cuadrático). La manera usual de probar esa observación es notando que a es residuo cuadrático módulo p si y sólo si $x^2 \equiv a$ mód p para algún $x \in A$. Luego, una manera de determinar el número de residuos cuadráticos en $\{1, 2, \ldots, p - 1\}$ es formando una partición del conjunto

$$\{1^2, 2^2, \ldots, (p-1)^2\}$$

con clases cuyos elementos sean congruentes entre sí, pero incongruentes

[1] El entero a es coprimo con p a lo largo de esta digresión.

con los elementos de cualquier otra clase. No es difícil ver que una partición tal resulta ser

$$\left\{1^2,(p-1)^2\right\}\cup\left\{2^2,(p-2)^2\right\}\cup,\ldots,\cup\left\{\left(\frac{p-1}{2}\right)^2,\left(\frac{p+1}{2}\right)^2\right\}.$$

De esto se sigue que en A hay $(p-1)/2$ residuos cuadráticos y $(p-1)/2$ residuos no cuadráticos. Con ayuda de este hecho vamos a derivar una caracterización útil de la noción de residuo cuadrático.

Si a es residuo cuadrático módulo p se tiene que $x^2 \equiv a$ mód p para algún entero x coprimo con p. Al ser p un primo impar, podemos elevar ambos miembros de la congruencia anterior a la potencia $(p-1)/2$ y obtener así que $x^{p-1} \equiv a^{(p-1)/2}$ mód p. Al aplicar el pequeño teorema de Fermat al miembro izquierdo de la congruencia previa llegamos a que $a^{(p-1)/2}$ es congruente con 1 módulo p siempre que a se supone residuo cuadrático módulo p. Nos aventuramos a estudiar ahora la validez de la implicación recíproca. Consideremos para ello la ecuación

$$x^{(p-1)/2} \equiv 1 \text{ mód } p.$$

Un resultado clásico de Lagrange asegura que dicha ecuación posee a lo más $(p-1)/2$ soluciones entre 1 y $p-1$. Como vimos anteriormente, cada residuo cuadrático es solución de la ecuación. Dado que en el conjunto $A = \{1,\ldots,p-1\}$ hay exactamente $(p-1)/2$ de ellos se concluye que el conjunto solución de la ecuación consiste precisamente de los enteros que son residuos cuadráticos módulo p. De todo lo anterior se desprende que *el entero a es residuo cuadrático módulo p si y sólo si* $a^{(p-1)/2} \equiv 1$ mód p. La caracterización esta es conocida como criterio de Euler y nos permite asegurar, por ejemplo, que -1 es residuo cuadrático de los primos congruentes con 1 módulo 4 y residuo no cuadrático de los primos congruentes con 3 módulo 4. Veamos ahora como es que este dato muestra injerencia en la determinación del número de primos que aparecen en la progresión aritmética $\{4k+1\}_{k\in\mathbb{Z}^+}$.

La técnica a emplear será la misma que la utilizada en el ejemplo **3.3**. Sea $A_k = \{p_1, \ldots, p_k\}$ un subconjunto de la progresión $\{4k+1\}_{k\in\mathbb{Z}^+}$ conformado exclusivamente de números primos. Considérese a continuación la expresión $a_k = 4(p_1 \cdots p_k)^2 + 1$. Dado que a_k es un entero positivo distinto de 1, a_k posee un divisor primo q_k. La existencia de ese divisor primo implica de inmediato que -1 es residuo cuadrático módulo q_k y por tanto $q_k \equiv 1$ mód 4. Claramente, el entero q_k es un número que no pertenece a A_k. Por tanto, nuestro argumento proporciona una manera de encontrar infinitos primos congruentes con 1 módulo 4 siempre que seamos capaces de determinar uno o más ejemplos concretos de primos de tal forma. Dado que $5 \equiv 1$ mód 4, concluimos que la progresión $\{4k+1\}_{k\in\mathbb{Z}^+}$ contiene infinitos números primos, tal como se esperaba establecer.

□

Los ejemplos anteriores proporcionan evidencia a favor del resultado anunciado. Cabe mencionar, además, que con ayuda de argumentos similares a los aquí expuestos se puede establecer la existencia de infinitos primos en otras progresiones aritméticas particulares. Sin embargo, nosotros optamos por dirigir inmediatamente la atención hacia la prueba de Dirichlet del caso general. Nuestra decisión queda respaldada por el hecho de que los enfoques *ad hoc* en la línea de los presentados previamente están condenados al fracaso. La justificación de tal aseveración viene en la última sección de este capítulo.

La prueba del teorema de Dirichlet sobre primos en progresiones aritméticas consta de dos ingredientes principales: el conocimiento de las propiedades de ciertos homomorfismos de $\mathbb{Z}_m^*$ y algunos resultados de no anulación para series creadas en base a dichos homomorfismos. Desarrollaremos a continuación, paso a paso, todo el material que en la prueba se requerirá. En el caso de los homomorfismos, el estudio se hace en general para grupos abelianos finitos.

§**2.** Sea G un grupo abeliano finito. Un carácter χ del grupo G es un homomorfismo de G en $\mathbb{C}^*$ (el grupo de complejos distinto de cero bajo

la multiplicación usual de números complejos). Una observación directa indica que el conjunto de caracteres de un grupo G queda dotado con estructura de grupo al considerar el producto usual de funciones. El grupo así obtenido se denomina **grupo dual** de G y se denota por $\widehat{G}$. El neutro del grupo $\widehat{G}$ recibe el nombre de carácter principal de G y se simboliza como χ_0.

Vamos a enunciar y demostrar ahora mismo dos propiedades básicas de los caracteres que resultarán importantes a la postre.

Escolio 3.5. Sea G un grupo finito. Se tiene entonces que:

a) Si $|G| = n$ entonces $\chi(g)$ es una raíz n-ésima de la unidad para cada $g \in G$.

b) Dado $H \leq G$ y $\chi \in \widehat{H}$, existen $[G : H]$ caracteres de G cuya restricción a H coincide con χ.

Prueba. La prueba del primer inciso es simple. Una consecuencia del teorema de Lagrange (en Grupos) indica que para cada $g \in G$ se cumple que $g^{|G|} = e$ y por tanto

$$1 = \chi(e) = \chi(g^{|G|}) = [\chi(g)]^{|G|} = [\chi(g)]^n$$

como se deseaba probar.

Para demostrar el segundo inciso notamos, en primer lugar, que si H es igual G entonces el resultado es trivialmente cierto. Ahora bien, si el subgrupo H es propio entonces $G = \langle H \cup \{g_1, \ldots, g_k\}\rangle$ para algunos elementos $g_1, \ldots, g_k \in G \setminus H$.

Si $k = 1$ y $\mathbf{m}$ lo usamos para denotar al índice de H en G entonces $g_1^{\mathbf{m}} = h_1$ para cierto $h_1 \in H$. Más aún, cada elemento de G puede escribirse en la forma hg_1^r donde $h \in H$ y $r \in [0, \mathbf{m}) \cap \mathbb{Z}$. Así, si $\chi_1 \in \widehat{G}$ es tal que $\chi_1|_H = \chi$ se sigue que

$$\chi_1(g) = \chi_1(hg_1^r) = \chi_1(h)\chi_1(g_1^r) = \chi(h)\chi_1^r(g_1).$$

La igualdad anterior indica que la manera de distinguir dos caracteres de G con idéntica restricción a H es por medio de la imagen que asignan a g_1. Dado que $\chi_1(g_1)$ es siempre una raíz $\mathbf{m}$-ésima de $\chi(h_1)$ se concluye que a lo más hay $\mathbf{m}$ caracteres de G cuya restricción a H es precisamente χ. La prueba concluye una vez que observamos que si $\epsilon_1, \ldots, \epsilon_{\mathbf{m}}$ son las raíces $\mathbf{m}$-ésimas de la unidad entonces las $\mathbf{m}$ funciones de G en $\mathbb{C}^*$ de la forma

$$\chi_k(hg_1^r) = \chi(h)\epsilon_k^r$$

determinan $\mathbf{m}$ elementos distintos de $\widehat{G}$ que al ser restringidos a H coinciden con χ.

Supongamos ahora que el resultado se cumple para k fijo (pero arbitrario). Probémosle entonces para el caso cuando G se obtiene de H por adjunción de $k+1$ elementos en $G \setminus H$.

Consideremos la sucesión de subgrupos de G determinados por las relaciones

$$G_0 = H, \qquad G_j = \langle G_{j-1} \cup \{g_j\} \rangle \quad \text{y} \quad G_{k+1} = G.$$

Los subgrupos $G_0, G_1, \ldots, G_k, G_{k+1}$ así contruidos son tales que

$$G_0 \leq G_1 \leq \ldots \leq G_k \leq G_{k+1}.$$

La hipótesis de inducción implica ahora que, dado $\chi_j \in G_j$, el número de caracteres de G_{j+1} que restringidos a G_j actuán como χ_j es $[G_{j+1} : G_j]$. Por tanto, dado $\chi \in \widehat{H}$, el número de elementos de $\widehat{G}$ cuya a restricción a H es igual a χ es

$$[G_1 : G_0][G_2 : \cdot G_1] \cdots [G_k : G_{k-1}][G_{k+1} : G_k] = \frac{|G_1| \cdots |G_{k+1}|}{|G_0| \cdots |G_k|} = \frac{|G_{k+1}|}{|G_0|} = [G : H],$$

con lo que la demostración culmina.

□

El siguiente teorema es una consecuencia notable de lo obtenido recientemente:

Teorema 3.6. Todo grupo abeliano finito es isomorfo al producto directo de un número finito de grupos cíclicos.

Prueba. Sea m el mínimo común múltiplo de los órdenes de todos los elementos del grupo G y $p_1^{a_1} \cdots p_r^{a_r}$ la descomposición en factores primos de m. Si $g_1, \ldots, g_r$ son elementos de G de órdenes respectivos $p_1^{a_1}, \ldots, p_r^{a_r}$ se sigue que

$$g_0 = g_1 \cdots g_r$$

tiene órden igual a m.

Sea ahora $H = \langle g_0 \rangle$ y χ el carácter de H que manda el elemento g_0^a en el complejo $(e^{\frac{2\pi i}{m}})^a$. Por el inciso b del escolio anterior se asegura la existencia de $\chi_1 \in \widehat{G}$ tal que $\chi_1|_H = \chi$. Luego, al tenerse que $g^m = e$ para cada $g \in G$ se cumple que

$$\mathrm{Im}(\chi_1) = \{z \in \mathbb{C} : z^m = 1\}.$$

Por otro lado, si hacemos $K = \mathrm{Ker}(\chi_1)$, el teorema fundamental de homomorfismos de grupos nos permite asegurar que

$$G/K \simeq \mathrm{Im}(\chi_1)$$

y de aquí que $|K| = |G|/m$. Ahora bien, al cumplirse que $H \cap K = \langle e \rangle$ se obtiene que

$$|HK| = \frac{|H||K|}{|H \cap K|} = (m)\left(\frac{|G|}{m}\right) = |G|$$

y por tanto $G = HK \simeq H \times K$. Hemos probado entonces que todo subgrupo cíclico de G de orden máximo es necesariamente un factor directo del grupo. La prueba termina ahora al notar que el argumento puede reiterarse sobre K, el cual es un grupo de orden menor a $|G|$.

□

Vamos a enunciar (y demostrar) a continuación un teorema donde resumiremos las propiedades más importantes de $\widehat{G}$. Dichas propiedades

desplegarán el mayor beneficio para nuestro estudio cuando las aterricemos al grupo que hemos distinguido desde un principio: $\mathbb{Z}_m^*$.

Teorema 3.7.

a) Un grupo abeliano finito G posee exactamente $|G|$ caracteres distintos. Esto es, $|G| = |\widehat{G}|$.

b) El grupo dual de G es siempre isomorfo a G.

c) Valen las siguientes identidades

$$\sum_{\chi\in\widehat{G}} \chi(g) = \begin{cases} |G| & \text{si } g = e \\ 0 & \text{en otro caso} \end{cases} \quad \text{y} \quad \sum_{g\in G} \chi(g) = \begin{cases} |G| & \text{si } \chi = \chi_0 \\ 0 & \text{en otro caso.} \end{cases}$$

Prueba. Empezamos, naturalmente, con la prueba del inciso a. En el teorema anterior se demostró que existen $g_1, \ldots, g_s \in G$ tales que los elementos de G pueden escribirse como $g_1^{r_1} g_2^{r_2} \cdots g_s^{r_s}$ donde $0 \leq r_i < h_i$ para $1 \leq i \leq s$ y los enteros h_i dependen solamente de G y satisfacen que $h_1 \cdots h_s = |G|$.

Así, para $\chi \in \widehat{G}$ se tiene que

$$\chi^{h_i}(g_i) = \chi(e) = 1$$

y por tanto, $\chi(g_i)$ es una raíz h_i-ésima de la unidad para cada $i \in \{1, \ldots, s\}$. Esta observación nos permite concluir que a lo más hay $h_1 \cdots h_s = |G|$ elementos en $\widehat{G}$. Un argumento análogo al usado en la prueba del inciso b del escolio **3.5.** nos permite concluir que en realidad hay exactamente $|G|$ caracteres en $\widehat{G}$ y la demostración de esta parte culmina.

La prueba de la parte b es en tres pasos:

1) Primero mostramos que el resultado vale cuando G es un grupo cíclico. En realidad, no hay mucho que hacer aquí pues si $G = \langle g \rangle$ y $|G| = n$ entonces basta con elegir $z \in \mathbb{C}^*$ de orden n. Se tiene entonces que el carácter χ de G que manda a g en z es tal que $\widehat{G} = \langle \chi \rangle$.

2) En esta parte lo que establecemos es que si $G_1, \ldots, G_r$ son grupos abelianos finitos entonces $\widehat{G_1} \times \cdots \times \widehat{G_k} \simeq \widehat{G_1 \times \cdots \times G_k}$. Dado que ambos grupos tienen el mismo orden, la prueba se reduce a exhibir un monomorfismo de un grupo a otro. Supongamos entonces que $\chi_i \in \widehat{G_i}$ y que χ es la aplicación de $G_1 \times \cdots \times G_k$ en $\mathbb{C}^*$ tal que $\chi(a_1, \ldots, a_k) = \chi_1(a_1) \cdots \chi_k(a_k)$. La construcción implica de inmediato que $\chi \in \widehat{G_1 \times \cdots \times G_k}$.

   Así, si $F : \widehat{G_1} \times \cdots \times \widehat{G_k} \to \widehat{G_1 \times \cdots \times G_k}$ es tal que $F(\chi_1, \ldots, \chi_k) = \chi$, afirmamos que F es monomorfismo. Para ver que F es homomorfismo denotemos con χ al carácter que F asocia a $(\chi_1\chi_1' \ldots, \chi_k\chi_k')$. Se tiene entonces que para $(a_1, \ldots, a_k)$ arbitrario en $G_1 \times \cdots \times G_k$ se cumple que

$$\begin{aligned} \chi(a_1, \ldots, a_k) &= (\chi_1\chi_1')(a_1) \cdots (\chi_k\chi_k')(a_k) \\ &= [\chi_1(a_1) \cdots \chi_k(a_k)][\chi_1'(a_1) \cdots \chi_k'(a_k)] \\ &= \chi'(a_1, \ldots, a_k)\chi''(a_1, \ldots a_k) \end{aligned}$$

   donde $\chi' = F(\chi_1, \ldots, \chi_k)$ y $\chi'' = F(\chi_1', \ldots, \chi_k')$. De todo esto se deriva que $F(\chi_1\chi_1' \ldots, \chi_k\chi_k') = F(\chi_1, \ldots, \chi_k)F(\chi_1', \ldots, \chi_k')$, tal como se deseaba establecer.

   La prueba de la inyectividad la efectuamos de manera similar. Si $F(\chi_1, \ldots, \chi_k) = \psi$, $F(\chi_1', \ldots, \chi_k') = \psi'$ y $\psi = \psi'$ se sigue que

$$\psi(a_1, \ldots, a_k) = \psi'(a_1, \ldots, a_k)$$

   para cada $(a_1, \ldots, a_k) \in G_1 \times \cdots \times G_k$. Esto último implica que $\chi_j = \chi_j'$ para cada $j \in \{1, \ldots, k\}$ y nuestra prueba termina.

3) Lo que resta hacer es conjuntar lo hecho en los pasos anteriores. Por lo establecido en el teorema **3.6** tenemos que si G es un grupo abeliano finito entonces $G \simeq G_1 \times \cdots \times G_k$ para ciertos grupos cíclicos G_i. Lo probado en 1 y 2 nos permite concluir entonces que

$$\widehat{G} \simeq \widehat{G_1 \times \cdots \times G_k} \simeq \widehat{G_1} \times \cdots \times \widehat{G_k} \simeq G_1 \times \cdots \times G_k \simeq G.$$

Vamos a proceder ahora con la prueba de las identidades que aparecen en el inciso c. Iniciamos con la identidad de más a la derecha.

Si $\chi = \chi_0$ entonces $\chi(g) = 1$ para cada $g \in G$ y por tanto

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

Ahora bien, si χ no es el carácter principal de G entonces $\chi(g')$ es distinto de 1 para algún $g' \in G$ y por tanto

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

Así,

$$[1 - \chi(g')] \sum_{g \in G} \chi(g) = 0$$

y de aquí que $\sum_{g \in G} \chi(g) = 0$, como se deseaba obtener.

La prueba del primer caso de la identidad de más a la izquierda depende de lo probado en el inciso a. En efecto, al ser $|G| = |\widehat{G}|$, se sigue que si $g = e$ entonces

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

Para calcular la suma en el caso en que g es diferente de e lo que hacemos es probar, en primer lugar, que hay un carácter en $\widehat{G}$ que no manda a g en 1.

Consideremos el subgrupo H de G generado por g. Al ser g distinto del elemento neutro de G se sigue que $|\widehat{H}| = |H| > 1$. En particular, existe $\psi \in \widehat{H}$ diferente del carácter principal. Como g genera al grupo es claro que $\psi(g) \neq 1$. Al aplicar la parte b del escolio **3.5** deducimos la existencia de $\chi' \in \widehat{G}$ cuya restricción a H es ψ. Lo anterior implica, en especial, que $\chi'(g) \neq 1$ y de aquí que

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

De lo anterior se desprende inmediatamente que $\sum_{\chi \in \widehat{G}} \chi(g) = 0$, tal como deseabamos establecer.

□

Antes de seguir con la exposición haremos una observación importante con respecto a las identidades establecidas en el inciso c del teorema previo. La denominación común en la literatura para dichas identidades es *relaciones de ortogonalidad.* Una manera de ver que la designación anterior es justa se indica a continuación.

Denotemos con $\mathbb{C}^G$ al conjunto de funciones de G en $\mathbb{C}$. Este conjunto puede ser dotado con estructura de espacio vectorial (sobre $\mathbb{C}$). La suma de funciones en $\mathbb{C}^G$ es la usual y lo mismo vale para la multiplicación por escalar. Si hacemos ahora

$$\langle f, g \rangle = \frac{1}{|G|} \sum_{x \in G} f(x)\overline{g(x)}$$

tenemos que $\mathbb{C}^G$ es un espacio vectorial complejo con producto interior hermítico.

Luego, si χ_1 y χ_2 son caracteres distintos de G, tenemos que

$$\langle \chi_1, \chi_2 \rangle = \frac{1}{|G|} \sum_{x \in G} \chi_1(x)\overline{\chi_2(x)} = \frac{1}{|G|} \sum_{x \in G} (\chi_1 \chi_2^{-1})(x) = 0.$$

Por otro lado, si $\chi_1 = \chi_2$ entonces

$$\langle \chi_1, \chi_2 \rangle = \frac{1}{|G|} \sum_{x \in G} \chi_1(x)\overline{\chi_1(x)} = \frac{1}{|G|} \sum_{x \in G} |\chi_1(x)|^2 = 1.$$

Los cálculos efectuados nos permiten concluir que $\widehat{G}$, el grupo dual de G, es un conjunto ortonormal del espacio vectorial $\mathbb{C}^G$ con respecto al producto interior $\langle , \rangle$.

Procedemos ahora a calcular la dimensión del espacio vectorial considerado. Supongamos que $G = \{g_1, \ldots, g_n\}$. La manera rápida de efectuar el cálculo es identificando primero a la función f con la $n-$ada

$(f(g_1), \ldots, f(g_n)) \in \mathbb{C}^n$. La identificación induce, de hecho, un isomorfismo entre $\mathbb{C}^G$ y $\mathbb{C}^{|G|}$ y por consiguiente $\dim(\mathbb{C}^G) = |G|$. Dado que $\widehat{G}$ es un subconjunto linealmente independiente de $\mathbb{C}^G$ con cardinal igual a la dimensión del espacio concluimos que

Escolio 3.8. El grupo dual de G es una base ortonormal para $\mathbb{C}^G$ y por tanto toda función $f \in \mathbb{C}^G$ puede expresarse de manera única como combinación lineal de caracteres, a saber

$$f = \frac{1}{|G|} \sum_{x \in \widehat{G}} \langle f, \chi \rangle \chi.$$

□

Sin duda alguna, la denominación mencionada resulta ahora más natural a la lectura.

Sea $\chi \in \widehat{\mathbb{Z}_m^*}$. Dicho carácter induce una aplicación χ' de $\mathbb{N}$ en $\mathbb{C}$. La aplicación se define por partes: si $(n, m) > 1$ entonces $\chi'(n) = 0$. Por otro lado, si $(n, m) = 1$ entonces $\chi'(n) = \chi([n])$ donde $[n]$ es clase de equivalencia de n en la congruencia módulo m. Abusando de la notación, nos permitiremos en lo sucesivo usar la letra χ tanto para el carácter como para la función aritmética que **induce**. Las funciones aritméticas obtenidas se conocen como **caracteres de Dirichlet módulo** m. En la siguiente proposición se dará cuenta de sus principales propiedades.

Proposición 3.9.

a) Si $(n, m) = 1$ entonces $\chi(n)$ es una raíz $\phi(m)-$ésima de la unidad.

b) Los caracteres de Dirichlet módulo m son funciones completamente multiplicativas con período m.

Prueba. La prueba de a es inmediata: como $(n, m) = 1$ entonces $[n] \in \mathbb{Z}_m^*$ y por tanto

$$\chi^{\phi(m)}(n) = \chi^{\phi(m)}([n]) = \chi([n]^{\phi(m)}) = \chi([1]) = 1.$$

La prueba del inciso b es en dos partes. Probaremos primero que los caracteres de Dirichlet módulo m son funciones completamente multiplicativas. Tomamos entonces k y n en $\mathbb{N}$ y procedemos por casos:

Si $(k, m) = 1$ y $(n, m) = 1$ entonces $(kn, m) = 1$ y por tanto

$$\chi(kn) = \chi([kn]) = \chi([k])\chi([n]) = \chi(k)\chi(n).$$

Si $(k, m) = 1$ y $(n, m) > 1$ entonces $(kn, m) > 1$ y de ahí que

$$\chi(kn) = 0 = \chi(n) = \chi(k)\chi(n).$$

El caso en que $(k, m) > 1$ y $(n, m) = 1$ es análogo al anterior. Resta analizar entonces el caso en que $(k, m) > 1$ y $(n, m) > 1$. En este escenario se cumple también que $(kn, m) > 1$ y por ende $\chi(kn) = 0 = \chi(k) = \chi(k)\chi(n)$.

La periodicidad es una consecuencia directa del hecho que $(n + m, m)$ es igual a (n, m). En efecto, si $(n, m) > 1$ entonces $(n + m, m) > 1$ y por tanto

$$\chi(n) = 0 = \chi(n + m).$$

Si $(n, m) = 1 = (n + m, m)$ entonces $\chi(n + m) = \chi([n + m]) = \chi([n]) = \chi(n)$ y la prueba termina.

□

Una pregunta que surge de modo natural en este momento es si el entero m es el mínimo período que un carácter de Dirichlet módulo m puede tener. La respuesta es negativa y el contraejemplo más simple se tiene en $m = 8$. En este caso, $\mathbb{Z}_8^* = \{[1], [3], [5], [7]\}$ y el carácter de Dirichlet módulo 8 inducido por el carácter $\{([1], 1), ([3], -1), ([5], 1), ([7], -1)\}$ es tal que $\chi(n) = 0$ si n es par, $\chi(n) = 1$ si $n \equiv 1$ mód 8, $\chi(n) = -1$ si $n \equiv 3$ mód 8, $\chi(n) = 1$ si $n \equiv 5$ mód 8 y $\chi(n) = -1$ si $n \equiv 7$ mód 8. Luego, si $m' = 4$ se tiene que $\chi(n + m') = \chi(n)$ para cada $n \in \mathbb{N}$ y de aquí que $m = 8$ no sea el período mínimo en este caso.

Es importante añadir que la relaciones de ortogonalidad derivadas en el teorema **3.7** dan lugar a identidades análogas para caracteres de Dirichlet.

Específicamente, si χ es el carácter de Dirichlet módulo m inducido por χ_0 entonces $\chi(n) = 1$ si $(n, m) = 1$ y $\chi(n) = 0$ cuando $(n, m) > 1$ y por tanto

$$\sum_{n=1}^{m} \chi(n) = \phi(m).$$

Si χ no es inducido por el carácter principal de $\widehat{\mathbb{Z}_m^*}$ entonces $\chi(n) \neq 1$ para algún $n \in [1, m)$ y de ahí que

$$\sum_{n=1}^{m} \chi(n) = 0.$$

Consideremos ahora las sumas de la forma $\sum_{\chi} \chi(n)$ donde χ varia en el conjunto de caracteres de Dirichlet módulo m. Claramente, si $n \equiv 1$ mód m entonces $\chi(n) = \chi([1])$ para cada $\chi \in \widehat{\mathbb{Z}_m^*}$ y por consiguiente

$$\sum_{\chi} \chi(n) = \sum_{\chi \in \widehat{\mathbb{Z}_m^*}} \chi([1]) = |\widehat{\mathbb{Z}_m^*}| = \phi(m).$$

Si $(n, m) > 1$ entonces la suma es trivialmente igual a 0. Si $(n, m) = 1$ y $[n]$ no es la clase del neutro de $\widehat{\mathbb{Z}_m^*}$ entonces la relación de ortogonalidad correspondiente nos permite concluir que

$$\sum_{\chi} \chi(n) = \sum_{\chi \in \widehat{\mathbb{Z}_m^*}} \chi([n]) = 0.$$

§**3.** Por cada carácter de Dirichlet módulo m consideremos la serie

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

Dado que $|\chi(n)| \leq 1$ se tiene que la serie es absolutamente convergente en el semiplano $\sigma > 1$. La función determinada por la suma de la serie en tal semiplano se denota por $L(s, \chi)$ y se denomina **función L de Dirichlet** o simplemente **función L**. En la proposición que sigue estableceremos la

analiticidad en $\sigma > 0$ de ciertas funciones L. Específicamente, lo que se tiene es que

Proposición 3.10. Si χ es un carácter de Dirichlet módulo m no inducido por el carácter principal de $\widehat{\mathbb{Z}_m^*}$ entonces $L(s,\chi)$ es analítica en el semiplano $\sigma > 0$.

Prueba. Sea $H(x) = \sum_{n\leq x} \chi(n)$. De la fórmula de integración por partes para la integral de Riemann-Stieltjes se tiene que

$$\sum_{n\leq x} \frac{\chi(n)}{n^s} = \int_{1^-}^{x} \frac{dH(t)}{t^s} = \frac{H(x)}{x^s} + s\int_1^x \frac{H(t)}{t^{s+1}}dt$$

Ahora bien, al ser $H(x)$ una función acotada[2], se sigue que

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = s\int_1^{\infty} \frac{H(t)}{t^{s+1}}dt.$$

[2]Sea χ un carácter de Dirichlet módulo m no inducido por el principal de $\widehat{\mathbb{Z}_m^*}$. Para cada $N \in \mathbb{N}$ se cumple que

$$\left|\sum_{n=1}^{N} \chi(n)\right| \leq \frac{\phi(m)}{2}.$$

Prueba. Por el algoritmo de la división en $\mathbb{Z}$ se tiene que $N = qm + r$ donde $q \geq 0$ y $1 \leq r < m$. Luego, dado que $a \equiv b$ mód m implica $\chi(a) = \chi(b)$ se sigue que

$$\sum_{n=1}^{N} \chi(n) = \left(\sum_{n=1}^{m} + \sum_{n=m+1}^{2m} + \ldots + \sum_{n=(q-1)m+1}^{qm}\right)\chi(n) + \sum_{n=qm+1}^{qm+r} \chi(n) = \sum_{n=1}^{r} \chi(n) = -\sum_{n=r+1}^{m} \chi(n).$$

Así, $2\left|\sum_{n=1}^{N} \chi(n)\right| = \left|\sum_{n=1}^{r} \chi(n) - \sum_{n=r+1}^{m} \chi(n)\right| \leq \sum_{n=1}^{m} |\chi(n)| = \phi(m)$ y la prueba termina. □

El teorema previo proporciona sólo una manera de acotar la suma $\sum_{n=1}^{N} \chi(n)$. Claramente, siempre se tiene a m como cota trivial. En 1918, G. Pólya e I. M. Vinogradov obtuvieron, de modo independiente, uno de los resultados más finos que se tienen en este rubro. El teorema obtenido por ellos establece que $\sum_{n=1}^{N} \chi(n) = O(\sqrt{m}\log m)$ siempre que χ es un carácter de Dirichlet módulo m diferente al inducido por el principal (cf. H. Davenport. *Multiplicative Number Theory.* Springer Verlag, New York, 1980, págs. 135-137).

Como la integral de la derecha es uniformemente convergente en todo semiplano $\sigma \geq \delta$ donde $\delta > 0$, el teorema de convergencia analítica nos permite concluir la analiticidad de $L(s, \chi)$ en el semiplano $\sigma > 0$.

□

Vamos a derivar ahora una consecuencia importante de la multiplicatividad de los caracteres de Dirichlet. El resultado es análogo al desarrollo en producto obtenido para la función zeta de Riemann en el teorema **2.1**.

Teorema 3.11. $L(s, \chi) = \prod_{p \in \mathbf{P}} \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}$ para $\sigma > 1$.

Prueba. Consideremos el producto finito

$$P(x) = \prod_{p \leq x} \left\{1 + \frac{\chi(p)}{p^s} + \frac{\chi(p^2)}{p^{2s}} + \ldots\right\}$$

Dado que $\sigma > 1$, cada factor es convergente. Luego, al haber sólo un número finito de factores podemos multiplicar las series involucradas y reordenar los términos obtenidos sin alterar la suma. Un término típico es de la forma

$$\frac{\chi(p_1^{a_1} \cdots p_r^{a_r})}{(p_1^{a_1} \cdots p_r^{a_r})^s}.$$

Por el teorema fundamental de la Aritmética podemos escribir

$$P(x) = \sum_{n \in A} \frac{\chi(n)}{n^s}$$

en donde A consta de todos los naturales n cuya descomposición sólo tiene factores primos menores o iguales a x. Por consiguiente

$$\sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} - P(x) = \sum_{n \in B} \frac{\chi(n)}{n^s}$$

en donde B es el conjunto de todos los n que tienen por lo menos un factor primo mayor a x. Así,

$$\left|\sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} - P(x)\right| \leq \sum_{n \in B} \frac{|\chi(n)|}{|n^s|} \leq \sum_{n > x} \frac{|\chi(n)|}{|n^s|}.$$

La convergencia absoluta de la serie $\sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$ implica que la serie de más a la derecha tiende a 0 cuando $x \to \infty$. Luego, $P(x) \to \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$ cuando $x \to \infty$ y de aquí que

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \prod_{p\in\mathbf{P}} \left\{1 + \frac{\chi(p)}{p^s} + \frac{\chi(p^2)}{p^{2s}} + \ldots\right\} = \prod_{p\in\mathbf{P}} \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}.$$

Lo único que resta mostrar es que el límite de $P(x)$ es distinto de 0. Sea X un número mayor o igual a 1 que cumple que $\sum_{n>X} \frac{|\chi(n)|}{|n^s|} < 1$. Se tiene entonces que

$$\left|\prod_{p>X} \left\{1 + \frac{\chi(p)}{p^s} + \frac{\chi(p^2)}{p^{2s}} + \ldots\right\} - 1\right| \leq \sum \frac{|\chi(n)|}{|n^s|}$$

donde la suma de la derecha se efectúa sobre todos los naturales cuyos factores primos son mayores a X. Claramente, la suma de dicha serie es menor o igual a $\sum_{n>X} \frac{|\chi(n)|}{|n^s|} < 1$. Por tanto,

$$\prod_{p>X} \left\{1 + \frac{\chi(p)}{p^s} + \frac{\chi(p^2)}{p^{2s}} + \ldots\right\} \neq 0$$

y la demostración culmina.

□

Como una consecuencia notable del teorema anterior se tiene que si χ_0 es el carácter de Dirichlet inducido por el neutro de $\widehat{\mathbb{Z}_m^*}$ entonces

$$L(s,\chi_0) = \prod_{p\in\mathbf{P}} \left(1 - \frac{\chi_0(p)}{p^s}\right)^{-1} = \frac{\prod_{p\in\mathbf{P}} \left(1 - \frac{1}{p^s}\right)^{-1}}{\prod_{p|m} \left(1 - \frac{\chi_0(p)}{p^s}\right)^{-1}} = \zeta(s) \prod_{p|m} \left(1 - \frac{1}{p^s}\right) \quad (3.1)$$

siempre que $\sigma > 1$.

§**4.** La estrategia usada por *Lejeune* Dirichlet para establecer la infinitud de los primos en la progresión aritmética $a, a+m, a+2m, \ldots$ fue probar que

$$\lim_{s\to 1^+} \sum_{p\equiv a \text{ mód } m} \frac{1}{p^s} = +\infty. \tag{3.2}$$

A continuación nos daremos a la tarea de justificar dicha igualdad.

Del desarrollo en producto obtenido en el teorema **3.11** se tiene que

$$L(s,\chi) = \prod_{p\in\mathbf{P}} \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}$$

siempre que $\sigma > 1$. De esto se sigue que

$$\begin{aligned}
\log L(s,\chi) &= \sum_{p\in\mathbf{P}} \log\left(1 - \frac{\chi(p)}{p^s}\right)^{-1} + 2\mathbf{k}_\chi i\pi \\
&= \sum_{p\in\mathbf{P}} \sum_{k=1}^{\infty} \frac{1}{k}\frac{1}{p^{ks}}\chi(p^k) + 2\mathbf{k}_\chi i\pi \\
&= \sum_{p\in\mathbf{P}} \frac{\chi(p)}{p^s} + \sum_{p\in\mathbf{P}} \sum_{k=2}^{\infty} \frac{1}{k}\frac{\chi(p^k)}{p^{ks}} + 2\mathbf{k}_\chi i\pi
\end{aligned}$$

para algún entero fijo $\mathbf{k}_\chi$. Mostraremos ahora que la expresión que aparece en el segundo término tiene módulo acotado siempre que $\sigma \geq 1$. En efecto, de

$$\begin{aligned}
\left|\sum_{p\in\mathbf{P}} \sum_{k=2}^{\infty} \frac{1}{k}\frac{\chi(p^k)}{p^{ks}}\right| &\leq \sum_{p\in\mathbf{P}} \sum_{k=2}^{\infty} \frac{1}{k|p^{ks}|} \\
&= \sum_{p\in\mathbf{P}} \sum_{k=2}^{\infty} \frac{1}{kp^{k\sigma}} \\
&\leq \frac{1}{2} \sum_{p\in\mathbf{P}} \sum_{k=2}^{\infty} \frac{1}{p^{k\sigma}} \\
&= \frac{1}{2} \sum_{p\in\mathbf{P}} \frac{p^{-2\sigma}}{1-p^{-\sigma}} \\
&\leq \frac{1}{2} \cdot \frac{1}{1-2^{-\sigma}} \cdot \sum_{p\in\mathbf{P}} p^{-2\sigma}
\end{aligned}$$

se obtiene que $\left|\sum_{p\in\mathbf{P}}\sum_{k=2}^{\infty}\frac{1}{k}\frac{\chi(p^k)}{p^{ks}}\right| \leq \frac{1}{2}\cdot\frac{1}{1-2^{-\sigma}}\cdot\zeta(2) \leq \zeta(2)$. Es importante agregar que la acotación hecha es independiente del carácter de Dirichlet χ. Luego, si b es un representante de la clase inversa de $[a] \in \mathbb{Z}_m^*$ se tiene que

$$\sum_{\chi}\chi(b)\log L(s,\chi) = \sum_{\chi}\chi(b)\sum_{p\in\mathbf{P}}\frac{\chi(p)}{p^s} + R(s) \tag{3.3}$$

donde $R(s) = \sum_{\chi}\chi(b)\sum_{p\in\mathbf{P}}\sum_{k=2}^{\infty}\frac{1}{k}\frac{\chi(p^k)}{p^{ks}} + 2i\pi\sum_{\chi}\chi(b)\mathbf{k}_{\chi}$. La acotación obtenida líneas arriba para $\sum_{p\in\mathbf{P}}\sum_{k=2}^{\infty}\frac{1}{k}\frac{\chi(p^k)}{p^{ks}}$ implica de inmediato que $R(s)$ es una función acotada en el semiplano $\sigma \geq 1$. Más aún, la identidad en (3.3) puede reescribirse como

$$\sum_{\chi}\chi(b)\log L(s,\chi) = \sum_{p\in\mathbf{P}}\frac{1}{p^s}\sum_{\chi}\chi(bp) + R(s) = \phi(m)\sum_{p\equiv a \text{ mód } m}\frac{1}{p^s} + R(s) \tag{3.4}$$

y de aquí que el problema de la determinación de

$$\lim_{s\to 1^+}\sum_{p\equiv a \text{ mód } m}\frac{1}{p^s}$$

se reduzca al estudio de las funciones $L(s,\chi)$ cuando s es cercano a 1.

Si χ es el carácter de Dirichlet inducido por el principal de $\widehat{\mathbb{Z}_m^*}$ se sigue de (3.1) que $L(s,\chi)$ es analítica en $\sigma > 0$ excepto por un polo simple en $s = 1$ con residuo $\phi(m)/m$. De lo anterior se obtiene que $\lim_{s\to 1^+} L(s,\chi_0) = +\infty$. Mostraremos a continuación que para cada χ distinto del principal se cumple que

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

es siempre finito. Dicha **observación** nos permitirá concluir, en la luz de (3.4), que $\lim_{s\to 1^+}\sum_{p\equiv a \text{ mód } m}\frac{1}{p^s} = +\infty$.

La prueba de la observación dependerá fuertemente del siguiente resultado clásico de Edmund Landau:

Teorema 3.12. (Lema de Landau) Supóngase que β y γ son números reales con $\beta < \gamma$. Si $\{c_n\}_{n\in\mathbb{N}}$ es una sucesión de reales no-negativos, la serie $\sum_{n=1}^{\infty} \frac{c_n}{n^s}$ converge en el semiplano $\sigma > \gamma$ y la función

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

admite una extensión analítica en algún dominio que contiene a $[\beta, \gamma]$, entonces la serie

$$\sum_{n=1}^{\infty} \frac{c_n}{n^\beta}$$

converge.

Prueba. Dado que f es analítica en $\sigma > \gamma$ podemos elegir $\delta > 0$ tal que f resulte analítica en el disco $|s - (\gamma + \delta)| < 2\delta$. Se tiene así que la serie de Taylor de f converge en dicho disco. Ahora bien, para $\sigma > \gamma$, la derivada $m-$ésima de f es

$$f^{(m)}(s) = (-1)^m \sum_{n=1}^{\infty} c_n (\log n)^m n^{-s}$$

En particular, si $\gamma - \delta < \sigma < \gamma + \delta$ se tiene que

$$\begin{aligned}
f(\sigma) &= \sum_{n=0}^{\infty} (\sigma - \gamma - \delta)^n \frac{f^{(n)}(\gamma + \delta)}{n!} \\
&= \sum_{n=0}^{\infty} (\sigma - \gamma - \delta)^n \frac{(-1)^n \sum_{k=1}^{\infty} c_k (\log k)^n k^{-(\gamma+\delta)}}{n!} \\
&= \sum_{k=1}^{\infty} \sum_{n=0}^{\infty} \frac{(\gamma + \delta - \sigma)^n c_k (\log k)^n k^{-(\gamma+\delta)}}{n!} \\
&= \sum_{k=1}^{\infty} e^{(\gamma+\delta-\sigma)\log k} c_k k^{-(\gamma+\delta)} \\
&= \sum_{k=1}^{\infty} \frac{c_k}{k^\sigma}
\end{aligned}$$

lo que implica la convergencia de la serie $\sum_{n=1}^{\infty} \frac{c_n}{n^s}$ para s real mayor que $\gamma - \delta$.

Sea $\kappa = \text{ínf}\left\{ s \in (\beta, \gamma) : \sum_{n=1}^{\infty} \frac{c_n}{n^s} \text{ converge} \right\}$. Se cumple entonces que la serie es convergente en el semiplano $\sigma > \kappa$ y define ahí una función analítica. Dado que dicha función coincide con f en $\sigma > \gamma$ se tiene que

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

en $\sigma > \kappa$. Más aún, en virtud de este último dato y de lo hecho en el párrafo anterior colegimos incluso que $\kappa = \beta$. A fin de culminar la demostración, procederemos a mostrar (por contradicción) que β es mayor que la abscisa de convergencia[3] de la serie $\sum_{n=1}^{\infty} \frac{c_n}{n^s}$.

Supongamos que la función $f(s)$ admite una extensión analítica en el disco $|s - \beta| < \epsilon$ para un cierto $\epsilon > 0$. Eligiendo $c \in \left(\beta, \beta + \frac{\epsilon}{2}\right)$ y haciendo

$$f(s) = \sum_{k=1}^{\infty} c_k k^{-c} k^{c-s} = \sum_{k=1}^{\infty} c_k k^{-c} e^{(c-s)\log k} = \sum_{k=1}^{\infty} \sum_{n=0}^{\infty} \frac{c_k k^{-c} (\log k)^n}{n!} (c-s)^n$$

queda una serie doble que converge absolutamente en el disco $|s - c| < \frac{\epsilon}{2}$. En particular, vista como serie de Taylor en $(c - s)$ se trata de la serie de Taylor para $f(s)$ alrededor de c. De la analiticidad de f en el disco $|s - c| < \frac{\epsilon}{2}$ se concluye que la serie de Taylor es convergente en dicho dominio. Por tanto, para cada $s \in \left(\beta - \frac{\epsilon}{2}, \beta\right)$ la serie $\sum_{n=1}^{\infty} \frac{c_n}{n^s}$ converge y de aquí el resultado.

□

[3]La abscisa de convergencia de la serie $\sum_{n=1}^{\infty} \frac{c_n}{n^s}$ es el número real σ que satisface lo siguiente: la serie es convergente si $\Re(s) > \sigma$ y divergente si $\Re(s) < \sigma$. Los detalles sobre la *existencia* de dicho σ se dilucidan en el capítulo 9 del clásico texto de E. C. Titchmarsh sobre Teoría de Funciones.

Siguiendo a Paul T. Bateman probaremos a continuación la no anulación de las funciones L de Dirichlet (para χ no principal) a lo largo de la recta $\sigma = 1$. En realidad, probaremos algo un tanto más fuerte que eso, a saber

Teorema 3.13. Para $\sigma > 1$ sea

$$g(s) = g(s, \epsilon) = \prod_{p \in \mathbf{P}} \left(1 - \frac{\epsilon(p)}{p^s}\right)^{-1} = \sum_{n=1}^{\infty} \frac{\epsilon(n)}{n^s}$$

donde ϵ es una función aritmética completamente multiplicativa y acotada. Supongamos además que g admite una extensión analítica a algún dominio que contenga al intervalo $[1/2, 1]$. Se cumple entonces que

$$g(1) \neq 0.$$

Prueba. Supongamos que $g(1) = 0$. Probaremos que tal supuesto nos lleva a una contradicción. Consideremos, para $\sigma > 1$, la función

$$F(s) = \zeta^2(s) g(s) g^*(s)$$

donde $g^*(s) = g(s, \overline{\epsilon})$. Dado que $g(1) = 0$ se tiene también que $g^*(1) = 0$. Ahora bien, la hipótesis de analiticidad de g alrededor de $[1/2, 1]$ implica que la función $g^*(s)$ es analítica en el mismo dominio[4] alrededor de $[1/2, 1]$. De esto se desprende que la función F es analítica alrededor de $[1/2, 1]$ pues el polo doble de ζ^2 en $s = 1$ se cancela con el cero de $g \cdot g^*$ en dicho punto. Por otra parte, de la identidad

$$(1 - z)^{-1} = \exp\left(\sum_{n=1}^{\infty} \frac{z^k}{k}\right)$$

[4]Notar que $g^*(s) = \overline{g(\overline{s})}$.

se obtiene que

$$\begin{aligned} F(s) &= \prod_{p\in\mathbf{P}}\left(1-\frac{1}{p^s}\right)^{-2}\left(1-\frac{\epsilon(p)}{p^s}\right)^{-1}\left(1-\frac{\overline{\epsilon}(p)}{p^s}\right)^{-1} \\ &= \prod_{p\in\mathbf{P}}\exp\left(\sum_{k=1}^{\infty}\frac{2+\epsilon^k(p)+\overline{\epsilon}^k(p)}{kp^{ks}}\right) \\ &= \prod_{p\in\mathbf{P}}\left\{1+\left(\sum_{k=1}^{\infty}\frac{2+\epsilon^k(p)+\overline{\epsilon}^k(p)}{kp^{ks}}\right)+\frac{1}{2!}\left(\sum_{k=1}^{\infty}\frac{2+\epsilon^k(p)+\overline{\epsilon}^k(p)}{kp^{ks}}\right)^2+\ldots\right\} \end{aligned}$$

siempre que $\sigma > 1$. Así, si $\displaystyle\sum_{n=1}^{\infty}\frac{a_n}{n^s}$ es la expansión en serie de Dirichlet para la función F en el semiplano $\sigma > 1$ se tiene que $a_n \geq 0$ para cada $n \in \mathbb{N}$ pues

$$\frac{2+\epsilon^k(p)+\overline{\epsilon}^k(p)}{k} = \frac{2+2\Re\{\epsilon^k(p)\}}{k} \geq 0$$

para cada natural k. Por otro lado, al ser

$$F(s) = \prod_{p\in\mathbf{P}}\left(1+\frac{2}{p^s}+\frac{3}{p^{2s}}+\ldots\right)\left(1+\frac{\epsilon(p)}{p^s}+\frac{\epsilon^2(p)}{p^{2s}}+\ldots\right)\left(1+\frac{\overline{\epsilon}(p)}{p^s}+\frac{\overline{\epsilon}^2(p)}{p^{2s}}+\ldots\right)$$

se concluye que

$$\begin{aligned} a_{p^2} &= 3+2\epsilon(p)+2\overline{\epsilon}(p)+\epsilon^2(p)+\epsilon(p)\overline{\epsilon}(p)+\overline{\epsilon}^2(p) \\ &= 2-\epsilon(p)\overline{\epsilon}(p)+\{1+\epsilon(p)+\overline{\epsilon}(p)\}^2 \\ &\geq 2-|\epsilon(p)|^2 \\ &\geq 1. \end{aligned}$$

Si en el lema de Landau hacemos ahora $c_n = a_n$, $\beta = 1/2$ y $\gamma = 1$ tenemos que

$$\sum_{p\in\mathbf{P}}\frac{1}{p} \leq \sum_{p\in\mathbf{P}}\frac{a_{p^2}}{p} \leq \sum_{n=1}^{\infty}\frac{a_n}{n^{1/2}} < \infty$$

lo que entra en contradicción con la bien conocida divergencia de la serie de los recíprocos de los números primos (ver proposición **1.1**). La contradicción muestra que el supuesto de que $g(1) = 0$ es absurdo y de ahí la veracidad del aserto en cuestión. □

El teorema recién probado apareció por vez primera en el artículo

A. E. Ingham. *Note on Riemann's ζ–function and Dirichlet's L–functions*. J. London Math. Soc. **5** (1930), 107-112.

No obstante, es importante recalcar que la elegante demostración que aquí se ha dado fue presentada por Paul T. Bateman en una nota de 1997 para la revista *L'Enseignement Mathématique* (la referencia completa es como se indica en [2]).

Una de las aplicaciones más interesantes de este teorema de Ingham se obtiene al hacer $\epsilon(n) = \chi(n)n^{-i\alpha}$ donde χ es un carácter de Dirichlet no inducido por el principal del grupo dual respectivo. En dicho caso se tiene que

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

y la tesis del teorema nos permite afirmar que $L(1 + i\alpha, \chi) \neq 0$. En particular se cumple que, para χ no principal, la expresión $L(1, \chi)$ es un número diferente de 0. Este dato será crucial en la culminación de la prueba del teorema de Dirichlet.

Veamos. Dado que $L(s, \chi)$ es analítica y distinta de 0 en $s = 1$, existe $\epsilon > 0$ tal que $L(s, \chi) \neq 0$ siempre que $|L(s, \chi) - L(1, \chi)| < \epsilon$. Así, si f es una rama analítica del logaritmo para el disco $|L(s, \chi) - L(s, 1)| < \epsilon$ y $\delta(\epsilon) > 0$ es como en la definición de continuidad para la función L en $s = 1$, se sigue que la función $f(L(s, \chi))$ es continua en el disco $|s - 1| < \delta(\epsilon)$. Por tanto, si χ es un carácter de Dirichlet módulo m diferente al inducido por el principal de $\widehat{Z_m^*}$, se tiene que

$$\lim_{s\to 1^+} \log L(s, \chi) = \lim_{s\to 1^+} [f(L(s, \chi)) + 2ki\pi] = f(L(1, \chi)) + 2ki\pi = O(1)$$

para algún entero k fijo que sólo depende de χ. La identidad (3.4) nos permite concluir entonces que

$$\lim_{s\to 1^+} \sum_{p\equiv a \text{ mód } m} \frac{1}{p^s} = \lim_{s\to 1^+} \left[\frac{\log L(s, \chi_0)}{\phi(m)} + O(1) \right] = +\infty \tag{3.5}$$

con lo que damos por culminada la demostración del célebre teorema de Dirichlet sobre primos en progresiones aritméticas.

□

Haremos a continuación algunas anotaciones sobre el resultado que se acaba de probar. La **primera** de ellas es, en realidad, una observación que se desprende directamente de la prueba anterior; a saber, la divergencia de la serie de los recíprocos de los números primos que pertenecen a la progresión $a, a+m, a+2m, \ldots$ para a y m naturales coprimos. La verificación es por contradicción. Supongamos que la serie $\sum_{p \equiv a \text{ mód } m} \frac{1}{p}$ converge a P. Se tiene entonces que para $s > 1$

$$\sum_{p \equiv a \text{ mód } m} \frac{1}{p^s} < \sum_{p \equiv a \text{ mód } m} \frac{1}{p} = P$$

lo cual es absurdo en vista de (3.5).

La **segunda** nota proporciona información cualitativa sobre la manera en que los números primos se distribuyen en las clases de $\mathbb{Z}_m^*$. El teorema de Dirichlet asegura que cada clase contiene un número infinito de primos. La pregunta que surge entonces es: ¿hay alguna manera de determinar si los números primos quedan repartidos equitativamente entre las distintas clases de $\mathbb{Z}_m^*$? La respuesta es afirmativa y está dada por la siguiente

Definición 3.14. Sea A un subconjunto de $\mathbf{P}$. Se dice que A tiene **densidad analítica** (o **densidad de Dirichlet**) relativa a $\mathbf{P}$ igual a d si

$$\lim_{s \to 1^+} \frac{\sum_{p \in A} \frac{1}{p^s}}{\sum_{p \in \mathbf{P}} \frac{1}{p^s}} = d.$$

La noción anterior puede verse, intuitivamente, como una medida de la proporción de los primos contenidos en A. Veamos como es que esta *medida* echa luz sobre la pregunta que hemos planteado recientemente.

En la cadena de igualdades (3.5) se mencionó que para a y m coprimos

persiste la identidad

$$\sum_{p\equiv a \text{ mód } m} \frac{1}{p^s} = \frac{\log L(s,\chi_0)}{\phi(m)} + O(1).$$

Por otro lado, sabemos que si $s > 1$ entonces $L(s,\chi_0) = \zeta(s)\prod_{p|m}\left(1-\frac{1}{p^s}\right)$ y por tanto

$$\begin{aligned}\log L(s,\chi_0) &= \log\zeta(s) + \log\prod_{p|m}\left(1-\frac{1}{p^s}\right)\\ &= \sum_{p\in\mathbf{P}}\sum_{k=1}^{\infty}\frac{1}{kp^{ks}} + \log\prod_{p|m}\left(1-\frac{1}{p^s}\right)\\ &= \sum_{p\in\mathbf{P}}\frac{1}{p^s} + O(1),\end{aligned}$$

de donde se concluye que

$$\frac{\sum_{p\equiv a \text{ mód } m}\frac{1}{p^s}}{\sum_{p\in\mathbf{P}}\frac{1}{p^s}} = \frac{1}{\phi(m)} + \frac{O(1)}{\sum_{p\in\mathbf{P}}\frac{1}{p^s}}.$$

Así, si en la igualdad anterior se toma el límite cuando $s \to 1^+$, concluimos que la densidad analítica relativa de los primos que aparecen en la clase $[a] \in \mathbb{Z}_m^*$ es igual a $\frac{1}{\phi(m)}$. Dado que el número de clases en el grupo es exactamente igual a $\phi(m)$ y cada clase contiene infinitos primos, la (sorprendente) conclusión es que los números primos se equidistribuyen entre las $\phi(m)$ clases del grupo.

Vamos a ahondar un poco más en el fenómeno anterior. Recordemos para ello, que en el §**1** del primer capítulo de nuestro trabajo se habló sobre la noción de densidad asintótica de un subconjunto de los números naturales. A continuación probaremos una importante relación que existe entre las nociones de densidad asintótica y analítica (relativas) de un subconjunto de **P**.

Teorema 3.15. Sea A un subconjunto de $\mathbf{P}$. Si A tiene densidad asintótica relativa a $\mathbf{P}$ igual a α entonces A también tiene densidad analítica relativa a $\mathbf{P}$ igual a α.

Prueba. Sean $A(x)$ la función contadora de A en $[1, x]$. Para $s > 1$ y $N \in \mathbb{N}$ se tiene que

$$\begin{aligned}\sum_{n\in A\cap[1,N]} \frac{1}{n^s} &= \sum_{n=1}^{N} n^{-s}[A(n) - A(n-1)] \\ &= \sum_{n=1}^{N-1} A(n)[n^{-s} - (n+1)^s] + A(N)N^{-s} \\ &= \sum_{n=1}^{N-1} A(n) \int_n^{n+1} sx^{-s-1}\, dx + A(N)N^{-s} \\ &= s\int_1^N x^{-s-1}A(x)\, dx + A(N)N^{-s}.\end{aligned}$$

Tomando límite, cuando $N \to \infty$, obtenemos que $\sum_{n\in A} n^{-s} = s\int_1^\infty x^{-s-1}A(x)\, dx$ y por tanto

$$\left|\frac{\sum_{n\in A}\frac{1}{n^s}}{\sum_{n\in\mathbf{P}}\frac{1}{n^s}} - \alpha\right| = \left|\frac{\int_1^\infty x^{-s-1}[A(x) - \alpha\pi(x)]\, dx}{\int_1^\infty x^{-s-1}\pi(x)\, dx}\right|. \tag{3.6}$$

Ahora bien, dado $\epsilon > 0$, sea X un número mayor que 1 tal que

$$\left|\frac{A(x)}{\pi(x)} - \alpha\right| < \epsilon \quad \text{siempre que} \quad x \geq X.$$

Utilizando esta X se obtiene, de acuerdo con la igualdad en (3.6), que

$$\begin{aligned}\left|\frac{\sum_{n\in A}\frac{1}{n^s}}{\sum_{n\in\mathbf{P}}\frac{1}{n^s}} - \alpha\right| &\leq \frac{\int_1^X x^{-s-1}|A(x) - \alpha\pi(x)|\, dx + \epsilon\int_X^\infty x^{-s-1}\pi(x)\, dx}{\int_1^\infty x^{-s-1}\pi(x)\, dx} \\ &\leq \frac{(1+\alpha)X^2}{\int_1^\infty x^{-s-1}\pi(x)\, dx} + \epsilon\end{aligned}$$

El resultado se sigue ahora al notar que el denominador del primer término en la última expresión tiende a $+\infty$ cuando $s \to 1^+$. □

Ciertamente, la idea de densidad asintótica resulta más cercana a lo que uno comúnmente asocia con el concepto de *densidad*. No obstante, el enfásis se ha hecho en este apartado en la noción de densidad analítica relativa porque hay ejemplos de subconjuntos de **P** que carecen de densidad asintótica relativa. Más aún, el teorema **3.15** nos asegura que las densidades coinciden cuando ambas existen. Luego, la noción de densidad analítica relativa puede verse como una generalización conveniente de la idea de densidad asintótica relativa. Otra consecuencia interesante del teorema anterior es como sigue. No es difícil ver que todo subconjunto de **P** con densidad analítica relativa mayor que cero contiene infinitos números primos. Es natural entonces preguntarse si vale también el converso de dicho resultado. Es decir, ¿será cierto que todo subconjunto de **P** que cuente con infinitos primos tiene densidad analítica relativa mayor que 0?

Una manera en que podemos darnos una idea sobre la naturaleza de la respuesta es por medio del siguiente argumento condicional: una creencia sostenida[5] en Aritmética indica que el conjunto **M** de primos de Mersenne es infinito. Por otro lado, al tenerse que en el intervalo $[1, n]$ hay no más de $\log n$ primos de Mersenne se concluye que

$$\frac{|\mathbf{M} \cap [1, n]|}{\pi(n)} \leq 2\frac{\log^2 n}{n} \tag{3.7}$$

para n suficientemente grande y de aquí que las densidades asintótica y analítica relativas a **P** del *potencial* conjunto infinito de primos de Mersenne sean iguales a 0. Lo interesante de este argumento condicional es que nos proporciona, de paso, una manera efectiva de construir contraejemplos concretos de subconjuntos infinitos de **P** con densidad analítica relativa igual a 0: basta con fijar un número natural $a > 1$ y elegir al n-ésimo elemento del subconjunto como el primer primo mayor a a^n. La construcción

[5]Hendrik Lenstra, Carl Pomerance y Samuel Wagstaff tienen incluso conjeturas sobre el comportamiento asintótico de la función que cuenta primos de Mersenne en el intervalo $[1, x]$. Para mayor información al respecto véase el artículo *Recent developments in primality testing* de Pomerance.

garantiza que el número de elementos del subconjunto que aparecen en el intervalo $[1, n]$ es menor o igual a $\log_a n$. El cálculo de la densidad se reduce entonces a una situación análoga a la efectuada en (3.7).

§**5.** Una vez discutido lo anterior procederemos a buscar estimados asintóticos para las funciones contadoras de primos en progresiones aritméticas donde $(a; m) = 1$. Básicamente lo que haremos aquí es seguir las consideraciones que en el capítulo 2 nos llevaron a la primer prueba del Teorema del Número Primo.

Sean $\pi(x; m, a) = \sum_{p \leq x, p \equiv a \text{ mód } m} 1$ y $\psi(x; m, a) = \sum_{n \leq x, n \equiv a \text{ mód } m} \Lambda(n)$. La idea es mostrar que la función $\psi(x; m, a)$ es asintóticamente equivalente a la función $\dfrac{x}{\phi(m)}$. La prueba de que dicha relación implica a su vez que

$$\pi(x; m, a) \sim \frac{x}{\phi(m) \log x}$$

puede efectuarse siguiendo lineamientos análogos a los que se emplearon en el capítulo 1 para vincular a las funciones $\vartheta(x), \psi(x)$ y $\pi(x)$. Una consecuencia notable de esta última expresión asintótica es que la progresión aritmética $a, a + m, a + 2m, \ldots$ tiene densidad asintótica relativa a **P** igual a la constante $\dfrac{1}{\phi(m)}$.

Antes de entrar de lleno con las consideraciones asintóticas vamos a demostrar un lema que será crucial en lo sucesivo.

Lema 3.16. Para $\sigma > 1$ se cumple que $-L'(s, \chi)/L(s, \chi) = \sum_{n=1}^{\infty} \dfrac{\chi(n)\Lambda(n)}{n^s}$.

Prueba. De la prueba del teorema **2.4** se sigue que la serie

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

converge absoluta y uniformemente en todo subconjunto compacto del

semiplano $\sigma > 1$. Así

$$\begin{aligned} L(s,\chi)K(s,\chi) &= \sum_{j=1}^{\infty} \frac{\chi(j)}{j^s} \sum_{k=1}^{\infty} \frac{\chi(k)\Lambda(k)}{k^s} \\ &= \sum_{n=1}^{\infty} \frac{1}{n^s} \sum_{jk=n} \chi(j)\chi(k)\Lambda(k) \\ &= \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} \sum_{d|n} \Lambda(n) \\ &= \sum_{n=1}^{\infty} \frac{\chi(n)\log n}{n^s} \\ &= -L'(s,\chi) \end{aligned}$$

y nuestra demostración termina.

□

Consideremos entonces las funciones

$$\begin{aligned} g(s) &= -\zeta'(s)/\zeta(s) - L'(s,\chi)/2L(s,\chi) - L'(s,\overline{\chi})/2L(s,\overline{\chi}) \\ h(s) &= -\zeta'(s)/\zeta(s) - L'(s,\chi)/2iL(s,\chi) + L'(s,\overline{\chi})/2iL(s,\overline{\chi}). \end{aligned}$$

Para $\sigma > 1$, el lema anterior nos permite afirmar que

$$g(s) = \sum_{n=1}^{\infty} [1 + \Re\chi(n)]\Lambda(n)n^{-s} \quad \text{y} \quad h(s) = \sum_{n=1}^{\infty} [1 + \Im\chi(n)]\Lambda(n)n^{-s}.$$

Ahora bien, si χ no es principal, la analiticidad de la función

$$-\frac{\zeta'(s)}{\zeta(s)} - \frac{1}{s-1}$$

en $\sigma \geq 1$ y la no anulación de las funciones L en el mismo semiplano implican que las funciones $g(s) - \dfrac{1}{s-1}$ y $h(s) - \dfrac{1}{s-1}$ son analíticas en $\sigma \geq 1$. Luego, si $G(x)$ es la función suma de $[1 + \Re\chi(n)]\Lambda(n)$ y $H(x)$ es la función suma de $[1 + \Im\chi(n)]\Lambda(n)$, se tiene que

$$g(s) = s\int_1^{\infty} \frac{G(x)}{x^{s+1}}dx \quad \text{y} \quad h(s) = s\int_1^{\infty} \frac{H(x)}{x^{s+1}}dx$$

El cambio de variable $x \mapsto e^x$ nos permite reescribir las integrales anteriores como

$$g(s) = s\int_0^\infty G(e^x)e^{-xs}\,dx \qquad h(s) = s\int_0^\infty H(e^x)e^{-xs}\,dx.$$

El teorema de Wiener-Ikehara nos asegura entonces que

$$\lim_{x\to\infty}\frac{G(e^x)}{e^x} = \lim_{x\to\infty}\frac{G(x)}{x} = 1 = \lim_{x\to\infty}\frac{H(x)}{x} = \lim_{x\to\infty}\frac{H(e^x)}{e^x}$$

y de aquí que

$$\sum_{n\le x}[1+\Re\chi(n)]\Lambda(n) \sim x \quad \text{y} \quad \sum_{n\le x}[1+\Im\chi(n)]\Lambda(n) \sim x. \tag{3.8}$$

Por otra parte, si $\chi = \chi_0$ entonces $g(s) - \dfrac{2}{s-1}$ y $h(s) - \dfrac{1}{s-1}$ son analíticas en $\sigma \ge 1$ donde

$$g(s) = \sum_{n=1}^\infty[1+\chi_0(n)]\Lambda(n)n^{-s} \quad \text{y} \quad h(s) = \sum_{n=1}^\infty \Lambda(n)n^{-s}.$$

Aplicando el teorema de Wiener-Ikehara una vez más concluimos que

$$\sum_{n\le x}[1+\chi_0(n)]\Lambda(n) \sim 2x \quad \text{y} \quad \sum_{n\le x}\Lambda(n) \sim x. \tag{3.9}$$

Las relaciones en (3.8) y (3.9) son la esencia del siguiente

Teorema 3.17. Valen los estimados siguientes

$$\sum_{n\le x}\chi_0(n)\Lambda(n) \sim x \quad \text{y} \quad \sum_{n\le x}\chi(n)\Lambda(n) = o(x).$$

Prueba. La primera relación es consecuencia de (3.9). Las equivalencias ahí obtenidas implican que

$$\lim_{x\to\infty}\frac{\sum_{n\le x}\chi_0(n)\Lambda(n)}{2x} = \frac{1}{2}.$$

Pasamos ahora con la prueba de la segunda relación. De lo establecido en (3.8) se desprende que

$$\sum_{n\le x}[2+\chi(n)]\Lambda(n) \sim 2x.$$

Luego, al tenerse que $\sum_{n\leq x}\Lambda(n)\sim x$, concluimos que

$$\lim_{x\to\infty}\frac{\sum_{n\leq x}\chi(n)\Lambda(n)}{x}=\lim_{x\to\infty}\frac{\sum_{n\leq x}[2+\chi(n)]\Lambda(n)-2\sum_{n\leq x}\Lambda(n)}{x}=0.$$

□

Estamos en condiciones de presentar la prueba de la fórmula

$$\psi(x;m,a)\sim\frac{x}{\phi(m)}.$$

En efecto, de

$$\psi(x;m,a)=\sum_{n\leq x,n\equiv a \text{ mód } m}\Lambda(n)$$

y la relación de ortogonalidad

$$\sum_{\chi}\overline{\chi}(a)\chi(n)=\begin{cases}\phi(m) & \text{si } a\equiv n \text{ mód } m\\ 0 & \text{en otro caso}\end{cases}$$

se sigue que

$$\begin{aligned}\psi(x;m,a) &= \sum_{n\leq x,n\equiv a \text{ mód } m}\Lambda(n)\\ &= \frac{1}{\phi(m)}\sum_{n\leq x}\Lambda(n)\left(\sum_{\chi}\overline{\chi}(a)\chi(n)\right)\\ &= \frac{1}{\phi(m)}\sum_{\chi}\overline{\chi}(a)\sum_{n\leq x}\chi(n)\Lambda(n).\end{aligned}$$

La deducción termina al notar que la última expresión es asintóticamente equivalente a $\frac{x}{\phi(m)}$.

§**6.** Una pregunta obligada que surge al analizar las pruebas de los casos particulares del teorema de Dirichlet discutidos al inicio de este capítulo es si el argumento básico detrás de ellas puede extenderse para probar el resultado en general. M. Ram Murty, en su artículo de [11], probó que, en

realidad, la extensión del argumento no es posible en todos los casos y que una condición necesaria para que haya una prueba de este tipo para la progresión aritmética de término inicial a y diferencia común m (con a y m coprimos) es que $a^2 \equiv 1$ mód m. De acuerdo con el mismo Murty la demostración de la suficiencia de tal condición fue el tema de un escrito de I. Schur de alrededor de 1912.

Lo anterior indica, en cierto modo, que las funciones L son un recurso cuya introducción en la prueba del teorema de Dirichlet es consecuencia de la dificultad inherente al resultado. Por otra parte, es importante mencionar que las vicisitudes enfrentadas al establecer dicho teorema quedan más que compensadas por la información adicional que sobre **P** derivamos a partir de él.

Así, por ejemplo, el teorema implica que si a y b son números naturales coprimos y $\mathcal{P}$ es un conjunto infinito de primos dentro la progresión aritmética $a, a + m, a + 2m, a + 3m, \ldots$, entonces $\mathcal{P}$ contiene progresiones aritméticas arbitrariamente largas.

Otra ilustración de los alcances del resultado estrella de este capítulo se tiene al notar que el teorema es, en realidad, una garantía teórica de la posibilidad de generar infinitos números primos mediante polinomios de grado uno (con coeficientes coprimos). En 1857 Bunyakosvky fue más allá en esta línea al conjeturar que si $f \in \mathbb{Z}[x]$ es de grado mayor que 1, irreducible sobre $\mathbb{Q}[x]$, con coeficiente principal positivo y el máximo común divisor de los números en $\{f(n) : n \in \mathbb{Z}\}$ es 1, entonces $f(n) \in \mathbf{P}$ para un número infinito de enteros n. De acuerdo con R. Murty, *esta conjetura es un importante problema de la Teoría de Números que todavía se encuentra en espera de solución.*

Conclusiones

En nuestra exposición del Teorema del Número Primo se ha puesto especial atención en el desarrollo histórico de las ideas.

El estilo discursivo de la exposición ha sido deliberado y, en cierta medida, una consecuencia de nuestra convicción de que los escritos en Matemáticas no tienen porque reducirse a la verificación de pruebas.

Aún cuando la mayoría de los temas aquí tratados son clásicos, el enfoque dado a la discusión es, en la mejor de nuestras opiniones, novedoso. El cotejo de la literatura usual puede constatar la veracidad de dicha afirmación.

Como trabajo futuro se contempla, entre otros puntos, el profundizar en el estudio de la topología τ que proporciona la correspondencia biyectiva entre los subconjunto sustanciosos de $\mathbb{N}$ y los subconjuntos densos de $(\mathbb{N}, \tau)$. La consideración de una topología tal para $\mathbb{N}$ fue inspirada, esencialmente, por la lectura de la prueba topológica de Hillel Furstenberg de la infinitud de primos (cf. H. Furstenberg. *On the infinitude of primes.* Amer. Math. Monthly **62** (1955), pág. 353.).

Bibliografía

[1] E. P. Balanzario. *Breviario de Teoría Analítica de los Números.* Sociedad Matemática Mexicana. Serie: Textos. 2003.

[2] P. T. Bateman. *A theorem of Ingham implying that Dirichlet's L–functions have no zeros with real part one.* L'Enseignement Math. **43** (1997), págs. 281-284.

[3] P. L. Chebyshev. *Mémoire sur les nombres premiers.* Mémoires de l'Acad. Imp. Sci. de St. Pétersbourg, VII, 1850.

[4] N. Costa Pereira. *A short proof of Chebyshev's theorem.* Amer. Math. Monthly **92** (1985), págs. 494-495.

[5] H. G. Diamond; P. Erdős. *On sharp elementary prime number estimates.* L'Enseignement Math. **26** (1980), págs 313-321.

[6] L. E. Dickson. *History of the Theory of Numbers, Vol. I.* Chelsea Publishing Company, New York, 1952.

[7] H. M. Edwards. *Riemann's Zeta Function.* Dover Publications Inc., 1974.

[8] L. J. Goldstein. *A history of the prime number theorem.* Amer. Math. Monthly **80** (1973), págs. 599-615.

[9] M. Hardy; C. Woodgold. *Prime simplicity.* Math. Intelligencer **31** 4 (2009), págs. 44-52.

[10] J. Korevaar. *On Newman's quick way to the prime number theorem.* Math. Intelligencer **4**-3 (1982), págs. 108-115.

[11] M. R. Murty. *Primes in certain arithmetic progressions.* Journal of the Madras University, (1988), págs. 161-169.

[12] D. J. Newman. *Simple analytic proof of the prime number theorem.* Amer. Math. Monthly **87** (1980), págs. 693-696.

[13] H. S. Wilf. *What is an answer?* Amer. Math. Monthly **89** (1982), págs. 289-292.

[14] D. Zagier. *The first 50 million prime numbers.* Math. Intelligencer **0** (1977), págs. 7-19.

[15] D. Zagier. *Newman's short proof of the prime number theorem.* Amer. Math. Monthly **104** (1997), págs. 705-708.