Escuela JAE 2013

Teoría de números II: Primos y gran criba

Fernando Chamizo

Julio 2013

En estas notas el hincapié está en las ideas más que en las demostraciones. A cambio se dan indicaciones, referencias y por supuesto, se hace notar qué argumentos no son rigurosos.

Respecto a la notación, C denotará habitualmente una constante, no necesariamente siempre la misma y p denotará un número primo.

1. Heurística sobre la distribución de los primos

Partiendo de las sencillas sumas de progresiones geométricas:

$$1+\frac{1}{2^5}+\frac{1}{2^{10}}+\frac{1}{2^{15}}+\frac{1}{2^{20}}+\cdots=\frac{1}{1-2^{-5}},\qquad 1+\frac{1}{3^5}+\frac{1}{3^{10}}+\frac{1}{3^{15}}+\frac{1}{3^{20}}+\cdots=\frac{1}{1-3^{-5}},$$

si las multiplicamos obtenemos

$$\sum_{n\in C_{2,3}}\frac{1}{n^5}=\frac{1}{1-2^{-5}}\cdot\frac{1}{1-3^{-5}}$$

donde $C_{2,3}$ son los enteros positivos que sólo contiene 2 y 3 en su factorización en primos, añadiendo además el 1. Esta idea se generaliza fácilmente a un número mayor de factores primos y a otras potencias. El resultado cuando tomamos todos los primos y usamos que todo entero positivo (mayor que 1) tiene una descomposición única, es la *fórmula producto de Euler* [10] válida para $s>1$

$$\zeta(s)=\prod_p\frac{1}{1-p^{-s}}\qquad\text{con}\qquad \zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s} \tag{1.1}$$

y p recorriendo los primos. Esta fórmula es importante porque relaciona algo sin primos y algo con primos. Cuando $s\to 1^+$, se tiene $\zeta(s)\to\infty$. De hecho aproximando por la integral

parece que debería ser $\zeta(s) \sim 1/(s-1)$ donde el símbolo $\sim$ significa que el cociente de ambos miembros tiene límite 1. Es decir,

$$\lim_{s\to 1^+} (s-1)\zeta(s) = \lim_{s\to 1^+} (s-1)\int_1^\infty \frac{dx}{x^s} = 1.$$

La prueba es un ejercicio que requiere algunas explicaciones que caen dentro del Cálculo I, aunque eso no significa que sea inmediato. Un ejercicio más complicado que expresa la misma idea de manera más precisa es

$$\lim_{s\to 1^+} \Big(\zeta(s) - \frac{1}{s-1}\Big) = \textsf{constante}.$$

También se sabe que todas las derivadas de $\log\big(\zeta(s)(s-1)\big)$ tienen límite cuando $s \to 1^+$. La derivada primera da

$$\lim_{s\to 1^+} \Big(\frac{\zeta'(s)}{\zeta(s)} + \frac{1}{s-1}\Big) = \textsf{constante}. \tag{1.2}$$

Con la misma constante que en la fórmula anterior. Incluso esta constante es famosa y aparece en otros contextos. El misterio acerca de ella y las pruebas de las afirmaciones anteriores se pueden consultar por ejemplo en [6].

Volviendo a (1.1), la idea es que los primos no pueden crecer muy deprisa, por ejemplo el primo n-ésimo no puede cumplir $p_n \geq Cn^2$ con $C > 0$, porque ello contradiría $\zeta(1) = \infty$. Tampoco pueden crecer muy despacio porque entonces la singularidad en 1 sería más abrupta de lo que es, no se la "mataría" multiplicando sólo por $s-1$. Esto es técnicamente más complicado de manipular pero con trabajo permitiría descartar que que $p_n \leq Cn$ o incluso $p_n \leq Cn(\log n)^\alpha$ para cualquier $\alpha < 1$.

Aquí apenas nos adentraremos en tecnicismos y nos limitaremos a introducir algunos pasos iniciales para entender la distribución de los primos.

Partiendo de (1.1), para ver mejor la relación con los primos, aplicamos derivación logarítmica, el truco favorito de Euler [2]:

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_p \frac{p^{-s}\log p}{1-p^{-s}} = \sum_p \Big(\frac{\log p}{p^s} + \frac{\log p}{p^{2s}} + \frac{\log p}{p^{3s}} + \dots\Big) = \sum_{n=1}^\infty \frac{\psi(n)-\psi(n-1)}{n^s},$$

donde se ha usado la notación clásica

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

con $k \in \mathbb{Z}^+$ y por tanto p^k recorriendo las potencias de primos que no superan a x.

Manipulando la serie,

$$-\frac{\zeta'(s)}{s\zeta(s)} = \frac{1}{s}\sum_{n=1}^{\infty}\psi(n)\Big(\frac{1}{n^s} - \frac{1}{(n+1)^s}\Big) = \sum_{n=1}^{\infty}\int_n^{n+1}\frac{\psi(n)dx}{x^{s+1}} = \int_1^{\infty}\frac{\psi(x)dx}{x^{s+1}},$$

que también se puede escribir como

$$-\frac{\zeta'(s)}{s\zeta(s)} - \frac{1}{s-1} = \int_1^{\infty}\frac{\psi(x)/x - 1}{x^s}\,dx. \tag{1.3}$$

Sabemos por (1.2) que el primer término es finito cuando $s \to 1^+$, por tanto el límite de la integral debe serlo también. Una consecuencia bastante espectacular es:

$$\boxed{\text{Si } \lim_{x\to\infty}\frac{\psi(x)}{x} \text{ existe, entonces vale 1.}}$$

La explicación es que si el límite fuera $C \neq 1$ entonces la integral en (1.3) cuando $s \to 1^+$ divergiría como $\int_1^{\infty}(C-1)/x$.

Resulta que el problema "técnico" de probar la existencia del límite es más complicado de lo que parece y no admite una solución demasiado simple y bonita (aunque esto es cuestión de gustos).

Para los más interesados diremos que una de las pruebas utiliza que la única forma de que no exista el límite es que $\psi(x)/x$ oscile y eligiendo $s \to 1^+ + it_0$ con un $t_0 > 0$ adecuado se podría conseguir que la oscilación resonase con la de x^{1+it_0} en el denominador con lo cual la integral en (1.3) tendería a infinito. Sin embargo se sabe que $\lim \zeta'(s)/\zeta(s) < \infty$ cuando $s \to 1^+ + it_0$. La demostración de esto último no es complicada pero sí muy ingeniosa, mientras que la de la resonancia requiere análisis armónico (transformadas de Fourier y regularizaciones [3, §3.10]).

Creyéndonos la existencia del límite, tenemos

$$\boxed{\lim_{x\to\infty}\frac{1}{x}\sum_{p\le x}\log p = 1.} \tag{1.4}$$

Esto sugiere que la "probabilidad" de que un número n sea primo es como $1/\log n$. De manera precisa, con argumentos elementales (pero no inmediatos), se deduce

$$\lim_{x\to\infty}\frac{|\{p \le x\}|}{x}\log x = 1$$

donde p recorre los primos, que es un enunciado habitual del *teorema de los números primos*. Una de las pruebas más breves de estos resultados está en [8].

Una famosa y breve memoria de Riemann [4], su único trabajo en teoría de números, sugería fuertemente una relación exacta entre $\psi(x)$ y los ceros de $\zeta(s)$ cuando se extiende $\zeta(s)$ a una función holomorfa de variable compleja definida en $\mathbb{C} - \{1\}$. Estas ideas, que no lograron materializarse en una prueba de (1.4) hasta décadas después del trabajo de Riemann, son fundamentales para entender la distribución de los números primos.

2. Heurística sobre los primos gemelos

Se dice que n y $n+2$ son *primos gemelos* si tanto n como $n+2$ son primos. Por ejemplo, 3 y 5 o también 2027 y 2029. Una antigua y famosa conjetura sin resolver es la existencia de infinitos primos gemelos. Evidentemente, sin saber la prueba o la refutación de esta conjetura las preguntas acerca de la distribución de los primos gemelos están fuera de nuestro alcance. Sin embargo, vamos a desarrollar una heurística acerca de cómo deberían funcionar las cosas.

Comencemos con un cálculo sencillo de una expresión complicada:

$$A_p = \frac{\mathrm{Prob}\big(\{n+2 \text{ no es divisible por } p, \text{ sabiendo que } n \text{ no es divisible por } p\}\big)}{\mathrm{Prob}\big(\{n+2 \text{ no es divisible por } p\}\big)}.$$

Por ejemplo,

$$A_2 = \frac{\mathrm{Prob}\big(\{n+2 \text{ impar, sabiendo } n \text{ impar}\}\big)}{\mathrm{Prob}\big(\{n+2 \text{ impar}\}\big)} = \frac{1}{1-1/2} = 2.$$

El caso $p = 3$ es más ilustrativo:

$$A_3 = \frac{\mathrm{Prob}\big(\{3 \nmid n+2, \text{ sabiendo } 3 \nmid n\}\big)}{\mathrm{Prob}\big(\{3 \nmid n+2\}\big)} = \frac{1-1/2}{1-1/3} = \frac{3}{4},$$

porque una vez que sabemos $3 \nmid n$ sólo quedan $3-1$ clases de congruencia y debemos evitar una de ellas. La misma idea funciona en general para cualquier primo $p > 2$:

$$A_p = \frac{\mathrm{Prob}\big(\{p \nmid n+2, \text{ sabiendo } p \nmid n\}\big)}{\mathrm{Prob}\big(\{p \nmid n+2\}\big)} = \frac{1-1/(p-1)}{1-1/p} = \frac{p(p-2)}{(p-1)^2}.$$

Como los primos no son divisibles por "nada no trivial", esto hace sospechar que en intervalos grandes

$$\frac{\mathrm{Prob}\big(\{n+2 \text{ primo, sabiendo que } n \text{ es primo}\}\big)}{\mathrm{Prob}\big(\{n+2 \text{ primo}\}\big)} \to K = 2\prod_{p>2} \frac{p(p-2)}{(p-1)^2} = 1{,}32032\ldots$$

Lo cual sugiere, por la probabilidad de la intersección y el teorema de los números primos,

$$\mathrm{Prob}\big(\{n \text{ y } n+2 \text{ primos}\}\big) \sim \mathrm{Prob}\big(\{n \text{ primo}\}\big)\mathrm{Prob}\big(\{n+2 \text{ primo}\}\big)K \sim \frac{K}{\log^2 n}.$$

Si preferimos algo más tangible que estas probabilidades, la conjetura plausible es

$$\lim_{x\to\infty} \frac{|\{n \le x \ : \ n \text{ y } n+2 \text{ son primos}\}|}{x} \log^2 x = K. \tag{2.1}$$

Por supuesto, esto implicaría la existencia de infinitos los primos gemelos y por tanto está fuera del alcance de los conocimientos actuales. Ahora podemos avanzar que el objetivo de estas lecciones es obtener un resultado parcial en este sentido.

3. Una desigualdad combinatoria

Comencemos con una igualdad muy simple. Para a y q enteros positivos, se cumple

$$\frac{1}{q}\sum_{k=1}^{q} e^{2\pi i a k/q} = \begin{cases} 1 & \text{si } q \mid a \\ 0 & \text{si } q \nmid a. \end{cases} \tag{3.1}$$

Para $q \mid a$ esto es evidente y para $q \nmid a$ se deduce de que la suma de todas las raíces q-ésimas de la unidad es cero. A pesar de su sencillez, esta fórmula es la base del análisis de Fourier discreto, responsable de muchas aplicaciones de las matemáticas en el mundo digital.

En principio, en un contexto aritmético (3.1) puede parece inútil porque, por ejemplo, si $q = 100$, es más fácil ver si un número es divisible por 100 que sumar 100 números complejos raros. Las ventajas aparecen cuando a o q o ambos varían. En esos casos puede ser más sencillo estudiar las interferencias de ondas discretas que controlar muchas propiedades de divisibilidad.

Aunque nuestros argumentos son bastante más generales (sirven para cualquier conjunto que evite algunas clases de congruencia módulo cierto conjunto de primos), aquí nos fijaremos en un conjunto directamente relacionado con los primos gemelos:

$$\mathcal{Z}_N = \big\{n \in [N, 2N) \ : \ n \text{ y } n+2 \text{ son primos}\big\}. \tag{3.2}$$

Para cualquier primo p, que supondremos distinto de 2, si $p < N$ se tiene $p \nmid n$ y $p \nmid n+2$ para cada $n \in \mathcal{Z}_N$, por tanto

$$|\mathcal{Z}_N|^2 = \Big|\sum_{n\in\mathcal{Z}_N} 1\Big|^2 = \Big|\sum_{\substack{k=0\\ k\neq 0,p-2}}^{p-1} \sum_{\substack{n\in\mathcal{Z}_N\\ n\equiv k \ (p)}} 1\Big|^2 \le (p-2)\sum_{k=0}^{p-1}\Big|\sum_{\substack{n\in\mathcal{Z}_N\\ n\equiv k \ (p)}} 1\Big|^2$$

donde se ha aplicado la desigualdad de Cauchy-Schwarz.

Desarrollando el cuadrado y usando (3.1) con $q = p$, la última expresión es

$$(p-2)\sum_{\substack{n,m\in\mathcal{Z}_N\\ n\equiv m\ (p)}}\sum 1 = \frac{p-2}{p}\sum_{n,m\in\mathcal{Z}_N}\sum\sum_{k=1}^{p} e^{2\pi i(n-m)k/p} = \frac{p-2}{p}\sum_{k=1}^{p}\left|S_{k/p}\right|^2.$$

donde hemos abreviado

$$S_\alpha = \sum_{n\in\mathcal{Z}_N} e^{2\pi i n\alpha}.$$

Claramente $|S_{p/p}|^2 = |\mathcal{Z}_N|^2$ y al pasar su contribución al primer miembro y despejar, obtenemos

$$\frac{2}{p-2}|\mathcal{Z}_N|^2 \le \sum_{k=1}^{p-1}\left|S_{k/p}\right|^2.$$

Para $p = 2$ se tiene $e^{2\pi i n k/p} = (-1)^k$ porque todos los elementos de $\mathcal{Z}_N$ son impares y la desigualdad es cierta olvidándonos del factor $2/(p-2)$. Entonces, sin excepciones para $p < N$ se tiene

$$\frac{2}{p}|\mathcal{Z}_N|^2 \le \sum_{k=1}^{p-1}\left|S_{k/p}\right|^2.$$

Si en vez de un primo se escogen ahora dos primos p_1 y p_2 distintos y con $2 < p_j < N$, se puede probar (véanse indicaciones más adelante)

$$\frac{2^2}{q}|\mathcal{Z}_N|^2 = \frac{2}{p_1}\cdot\frac{2}{p_2}|\mathcal{Z}_N|^2 \le \sum_{\substack{k=1\\ \mathrm{mcd}(k,q)=1}}^{q}\left|S_{k/q}\right|^2 \qquad \text{con } q = p_1p_2.$$

En general, si $q = p_1p_2\ldots p_r$ con $p_j < N$ primos distintos

$$\frac{2^r}{q}|\mathcal{Z}_N|^2 \le \sum_{\substack{k=1\\ \mathrm{mcd}(k,q)=1}}^{q}\left|S_{k/q}\right|^2. \tag{3.3}$$

Toda la prueba de (3.3) radica en entender el paso de un primo a dos primos pero esto no es tan fácil, aunque sólo requiera argumentos elementales. La clave está en notar que si $q = p_1p_2$ entonces

$$\sum_{\substack{k=1\\ \mathrm{mcd}(k,q)=1}}^{p-1}\left|S_{k/q}\right|^2 = \sum_{k_1=1}^{p_1-1}\sum_{k_2=1}^{p_2-1}\Big|\sum_{n\in\mathcal{Z}_N} e^{2\pi i n k_1/p_1}e^{2\pi i n k_2/p_2}\Big|^2.$$

Una vez hecho esto, la suma en k_1 se trata como en el caso de un primo salvo que hay que reemplazar algunos unos por $e^{2\pi i n k_2/p_2}$. Es un desafío para el lector interesado completar la prueba de (3.3) empleando estas ideas u otras que se le ocurran.

4. Una desigualdad analítica

El análisis de Fourier nos dice que una onda (función) 1-periódica es superposición de tonos puros de la forma $e^{2\pi inx}$, $n \in \mathbb{Z}$. Para simplificar, consideramos sólo ondas que contienen un número finito de tonos puros, digamos los que tienen frecuencias entre 1 y N:

$$W(x) = \sum_{n=1}^{N} a_n e^{2\pi inx} \qquad \text{con} \qquad a_1, a_2, \ldots, a_n \in \mathbb{C}. \tag{4.1}$$

La *energía total* de W es $E_W = \int_0^1 |W|^2 = \sum_{n=1}^N |a_n|^2$. Aunque la energía esté acotada, al *intensidad* de W en un punto x, definida como $|W(x)|^2$, puede ser grande. Por ejemplo, si $a_j = 1/\sqrt{N}$, se tiene $E_W = 1$ pero $|W(0)|^2 = N$. Siempre podemos escoger los a_n con mala intención para que haya resonancias en algunos puntos. La *desigualdad de gran criba* (veremos más adelante la razón de este extraño nombre) afirma que es imposible que la intensidad media supere en mucho a la energía si el promedio se hace en suficientes puntos bien separados.

¿Por qué nos preocupamos ahora de estos problemas sobre ondas? El objetivo es acotar el segundo miembro de (3.3), o más bien un promedio suyo, utilizando que es una suma de intensidades de ondas.

Dado $\delta > 0$, se dice que $0 < x_1 < x_2 < \ldots < x_M \leq 1$ están *δ-espaciados* en $(0, 1]$ si $x_{j+1} - x_j \geq \delta$ para $1 \leq j \leq M$ y $1 - x_M + x_1 \geq \delta$.

Proposición 4.1 (Desigualdad de gran criba)**.** *Si $x_1, x_2, x_3, \cdots \in (0, 1]$ están δ-espaciados y $W(x)$ es como en* (4.1)*, entonces se tiene*

$$\sum_{\nu} \big|W(x_\nu)\big|^2 \leq (N + \delta^{-1}) \sum_{n=1}^{N} |a_n|^2.$$

Demostrar esta desigualdad tal y como está escrita, es duro [7], [5, §9.1] pero para nuestra aplicación no es muy importante perder una constante multiplicando al segundo miembro y nos contentaremos con esa versión débil que admite una prueba bastante elemental debida a P.X. Gallagher.

Si $I \subset \mathbb{R}$ es un intervalo, para todo $x, y \in I$ se tiene $|f(x)| \leq |f(y)| + \int_I |f'|$, simplemente por la regla de Barrow. Al tomar y de modo que $|f|$ alcance un mínimo, se sigue

$$|f(x)| \leq \frac{1}{|I|} \int_I |f| + \int_I |f'|$$

que puede considerarse como una desigualdad de Sobolev.

Subdividiendo $(0, 1]$ en intervalos semiabiertos I_k de longitud $\delta/2$, en cada uno de ellos a lo más hay un x_ν. Ahora elegimos $f = W^2$ en la desigualdad anterior para $I = I_k$ y sumamos

en k.

$$\sum_{\nu} \big|W(x_\nu)\big|^2 \le C \sum_{k} \left(\delta^{-1} \int_{I_k} |W|^2 + \int_{I_k} |WW'| \right) = C\delta^{-1} \int_0^1 |W|^2 + C \int_0^1 |WW'|.$$

La desigualdad de Cauchy-Schwarz implica

$$\sum_{\nu} \big|W(x_\nu)\big|^2 \le C\delta^{-1} \int_0^1 |W|^2 + C\Big(\int_0^1 |W'|^2 \Big)^{1/2} \Big(\int_0^1 |W|^2 \Big)^{1/2}$$

y la fórmula para la energía (la identidad de Parseval) concluye la prueba de la Proposición 4.1 salvo que hay cierta $C > 1$ que multiplica al segundo miembro.

5. El resultado

Con las herramientas de las dos secciones anteriores vamos a demostrar el siguiente resultado sobre primos gemelos:

Teorema 5.1. *Sea $\mathcal{Z}_N$ definido como en* (3.2)*. Existe una constante $C > 0$ tal que*

$$|\mathcal{Z}_N| \le \frac{CN}{\log^2 N} \qquad \text{para todo } N > 1.$$

Teniendo en cuenta la conjetura (2.1), esta cota es del orden esperado. Incluimos dos consecuencias. La primera es una forma equivalente más atractiva del enunciado, mientras que la segunda busca un poco la espectacularidad y se debe a V. Brun, que introdujo los métodos de criba modernos y dedujo el teorema anterior, inicialmente en una versión ligeramente más débil [5, §6.1], [9].

Corolario 5.2. *Existe una constante $C > 0$ tal que*

$$|\{n \le N \ : \ n \text{ y } n+2 \text{ son primos}\}| \le \frac{CN}{\log^2 N} \qquad \text{para todo } N > 1.$$

Demostración. Sea $K \in \mathbb{Z}^+$ tal que $2^K \le N < 2^{K+1}$, entonces aplicando el teorema

$$|\{n \le N \ : \ n \text{ y } n+2 \text{ son primos}\}| \le \sum_{k=1}^{K} \big|\mathcal{Z}_{2^k}\big| \le \frac{C}{\log^2 2} \sum_{k=1}^{K} k^{-2} 2^k.$$

Como $a_k = k^{-2}2^k$ es una sucesión creciente para $k \ge 3$,

$$\sum_{k=3}^{K} k^{-2}2^k \le 2 \sum_{K/2 \le k \le K} k^{-2}2^k \le 8K^{-2} \sum_{K/2 \le k \le K} 2^k \le 16 a_K$$

y basta notar que $a_K \le N/\log_2^2 N$. □

Corolario 5.3. *La siguiente serie converge:*

$$\sum_{n\in\mathcal{P}_2}\frac{1}{n} \qquad \textit{donde} \quad \mathcal{P}_2=\{n\ :\ n\ y\ n+2\ son\ primos\}.$$

Demostración. Reescribimos la suma del enunciado y la acotamos con el teorema:

$$\sum_{n\in\mathcal{P}_2}\frac{1}{n}=\sum_{k=1}^{\infty}\sum_{n\in\mathcal{Z}_{2^k}}\frac{1}{n}\le\sum_{k=1}^{\infty}\frac{\left|\mathcal{Z}_{2^k}\right|}{2^k}\le\frac{C}{\log^2 2}\sum_{k=1}^{\infty}\frac{1}{k^2}$$

y esta última serie converge. □

Para probar el Teorema 5.1 utilizaremos un resultado auxiliar que está dentro de la rutina cuando uno tiene experiencia con algunos resultados de la teoría de la estimación de funciones aritméticas pero cuyos detalles podrían asustar al principiante.

Lema 5.4. *Sea $\mathcal{Q}_M$ el conjunto de enteros en $[1,M]$ que no tienen factores primos repetidos. Existe una constante $C>0$ tal que*

$$\sum_{q\in\mathcal{Q}_M}\frac{d(q)}{q}\ge C\log^2 M \qquad \textit{para } M>1$$

donde $d(q)$ indica el número de divisores de q.

Vamos a hacer trampas y daremos una prueba olvidándonos de la condición de no tener factores primos repetidos, es decir, cambiando $q\in\mathcal{Q}_M$ por $1\le q\le M$.

Escribiendo $q=dm$ para cada divisor d de q,

$$\sum_{q=1}^{M}\frac{d(q)}{q}=\sum_{d=1}^{M}\sum_{m\le M/d}\frac{1}{md}\ge\sum_{d\le\sqrt{M}}\frac{1}{d}\sum_{m\le\sqrt{M}}\frac{1}{m}\ge C\log^2 M,$$

donde se ha usado que las sumas parciales de la serie armónica se aproximan por la integral.

Un truco para aprovechar esta idea sin hacer trampas es quitar todos los q que sean divisibles por el cuadrado de un primo. Así

$$\sum_{q\in\mathcal{Q}_M}\frac{d(q)}{q}\ge\sum_{q=1}^{M}\frac{d(q)}{q}-\sum_{p\le\sqrt{M}}\frac{1}{p^2}\sum_{q\le M/p^2}\frac{d(p^2q)}{q}.$$

Los sumatorios en q se pueden estimar con razonamientos similares a los anteriores pero hay que tener cuidado para obtener estimaciones precisas. En unas notas anteriores se probó el lema de esta forma considerando $p=2$ y $p=3$ aparte y empleando $d(p^2q)\le 3d(q)$ en el resto de los casos. Quizá algún lector interesado sepa encontrar una prueba alternativa más rápida e igualmente elemental.

Demostración del Teorema 5.1. En (3.3), se tiene $2^r = d(q)$ porque los divisores de q están en biyección con los subconjuntos de los r factores primos. Sumando en $q \in \mathcal{Q}_{\sqrt{N}}$ y aplicando el lema anterior, se deduce

$$|\mathcal{Z}_N|^2 \leq \frac{C}{\log^2 N} \sum_{k/q \in \mathcal{F}} \left|S_{k/q}\right|^2 \tag{5.1}$$

donde $\mathcal{F}$ es el conjunto de fracciones irreducibles en $(0,1]$ con denominador menor o igual que $\sqrt{N}$. Ahora bien, $\mathcal{F}$ es un conjunto N^{-1}-espaciado porque si $a/b, a'/b' \in \mathcal{F}$, entonces $|a/b - a'/b'| \geq 1/bb' \geq 1/N$. Además podemos escribir $S_{k/q}$ como

$$S_{k/q} = \sum_{n=1}^{2N} a_n e^{2\pi i n k/q} \qquad \text{con} \qquad a_n = \begin{cases} 1 & \text{si } n \in \mathbb{Z}_N \\ 0 & \text{si } n \notin \mathbb{Z}_N. \end{cases}$$

Estamos por tanto en condiciones de aplicar la desigualdad de gran criba con N^{-1} en lugar de δ y $2N$ en lugar de N. De esta forma, al utilizar la Proposición 4.1 en (5.1), se obtiene

$$|\mathcal{Z}_N|^2 \leq \frac{C}{\log^2 N} N \sum_{n=1}^{2N} |a_n|^2 = \frac{CN}{\log^2 N} |\mathcal{Z}_N|,$$

lo cual prueba el resultado. □

6. Epílogo

Si repasamos los argumentos anteriores, veremos que lo único que hemos empleado de $\mathcal{Z}_N$ es que sus elementos evitan dos clases de congruencias (la de 0 y la de $p-2$) al tomar módulo un primo $p \leq \sqrt{N}$. El número de clases de congruencia que se excluyen y el rango de los primos, no son críticos, y las ideas expuestas llevan al siguiente resultado general [5]:

Teorema 6.1. *Sea $\mathcal{Z}$ un conjunto de enteros $\mathcal{Z} \subset [K+1, K+N]$, $N, K \in \mathbb{Z}^+$. Supongamos que para cada primo $p \leq M$, hay $\omega(p)$ clases módulo p que no aparecen en $\mathcal{Z}$, entonces*

$$|\mathcal{Z}| \leq \frac{N + M^2}{\sum_{q \leq \mathcal{Q}_M} h(q)} \qquad \textit{con} \qquad h(q) = \prod_{p|q} \frac{\omega(p)}{p - \omega(p)}$$

donde se ha utilizado la notación del Lemma 5.4.

La famosa criba de Eratóstenes permite construir una tabla de primos excluyendo los números en la clase del cero módulo primos pequeños. Por extensión se llaman *métodos de criba* a las técnicas que permiten estimar el tamaño de un conjunto que excluye clases de

congruencia módulo ciertos primos. El nombre de gran criba para la Proposición 4.1 proviene de que históricamente sirvió para probar el método de criba dado por el teorema anterior, el cual fue especialmente útil para excluir un gran número de clases de congruencia, aunque en nuestra aplicación sólo hemos excluido dos.

Los métodos de criba alcanzan una complicación impresionante a pesar de proceder de problemas de enunciado elemental. No sólo proporcionan cotas superiores, aunque las cotas inferiores o las fórmulas asintóticas son más escasas y difíciles. Respecto a los primos gemelos, uno de los pináculos de la aplicación de los métodos de criba es el siguiente resultado [5, §25.6] probado por J.-R. Chen en 1973:

Teorema 6.2. *Sea*

$$\mathcal{Z}'_N = \big\{n \in [N, 2N) \ : \ n \text{ es primo y } n+2 \text{ tiene a lo más dos factores primos}\big\}.$$

Existe una constante $C > 0$ tal que

$$|\mathcal{Z}'_N| \geq \frac{CN}{\log^2 N} \qquad \text{para todo } N > 1.$$

Resultados anteriores permitían más posibles factores para $n + 2$, lo que aparentemente suscita una esperanza para una nueva mejora que pase de dos a uno y que por tanto pruebe la conjetura de los primos gemelos. Sin embargo, con los conocimientos actuales esta esperanza es remota. Hay una dificultad teórica muy fuerte, llamada *fenómeno de paridad* [5, §16.4], que impide el paso de a lo más dos factores primos a un factor primo. Esta dificultad implica que necesariamente hay que utilizar información intrínseca acerca del problema concreto más allá del número de clases de congruencia excluidas.

Hay varios libros dedicados por entero a los métodos de criba. Uno de los más recientes y completos es [5]. Aunque contiene temas avanzados, como el teorema anterior, los capítulos iniciales están escritos para principiantes y el libro es en gran medida autocontenido. Una introducción más ligera y muy bien escrita es [1].

Referencias

[1] A. C. Cojocaru and M. R. Murty. *An introduction to sieve methods and their applications*, volume 66 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2006.

[2] W. Dunham. *Euler: El maestro de todos los matemáticos*. Nivola libros y ediciones, S.L., 2000.

[3] H. Dym and H. P. McKean. *Fourier series and integrals*. Academic Press, New York, 1972. Probability and Mathematical Statistics, No. 14.

[4] H. M. Edwards. *Riemann's zeta function*. Academic Press, New York-London, 1974. Pure and Applied Mathematics, Vol. 58.

[5] J. Friedlander and H. Iwaniec. *Opera de cribro*, volume 57 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2010.

[6] A. Ivić. *The Riemann zeta-function*. Dover Publications Inc., Mineola, NY, 2003. Theory and applications, Reprint of the 1985 original [Wiley, New York; MR0792089 (87d:11062)].

[7] H. L. Montgomery. The analytic principle of the large sieve. *Bull. Amer. Math. Soc.*, 84(4):547–567, 1978.

[8] D. J. Newman. *Analytic number theory*, volume 177 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.

[9] H. Rademacher. *Lectures on elementary number theory*. A Blaisdell Book in the Pure and Applied Sciences. Blaisdell Publishing Co. Ginn and Co. New York-Toronto-London, 1964.

[10] C. E. Sandifer. *The early mathematics of Leonhard Euler*. MAA Spectrum. Mathematical Association of America, Washington, DC, 2007.