Ocho lecciones de teoría de números

Fernando Chamizo Lorente

Lorente Fernando
Chamizo

2011

Índice general

Prefacio

Estas notas contienen ocho lecciones que impartí a mis estudiantes de doctorado Dulcinea Raboso y Serafín Ruiz (ocasionalmente hubo algún otro asistente), a quienes agradezco su interés y la corrección de algunas erratas. Las variaciones respecto al material escrito que les entregué en aquellas sesiones no son muy sustanciales. Básicamente consisten en algunos añadidos, la inclusión de figuras, cambios estilísticos menores y la corrección de las erratas detectadas.

Un vistazo somero basta para percatarse de que estas lecciones difieren en su presentación de lo que se encuentra habitualmente bajo epígrafes similares en una monografía. Es mi opinión que el rigor matemático se extrema con mucha frecuencia en la literatura, hasta el punto de que los enunciados y sus demostraciones desplazan toda referencia a los significados y las ideas. Aquí se intenta proceder en sentido contrario, eliminando casi todas las demostraciones en favor de una explicación de las ideas subyacentes. Me sentiría satisfecho si estas lecciones sirvieran como paso previo antes de adentrarse en la bibliografía, o como una forma de adquirir someramente conocimientos de los temas seleccionados sin pasar por los detalles.

Finalmente, respecto a los contenidos, la finalidad de las lecciones era cubrir conocimientos que pudieran necesitar mis estudiantes en su doctorado bajo mi dirección. Por ello, casi todos los temas caen dentro de la teoría analítica de números clasificable en cierto modo como clásica. La ordenación de los temas difiere de la original de las lecciones y creo que es más coherente.

Madrid, 26 de septiembre de 2011

Fernando Chamizo Lorente

Capítulo 1

El teorema de los números primos

1.1. Formulación

El *teorema de los números primos* es quizá el resultado más emblemático de la teoría analítica de números apareciendo con frecuencia en textos de divulgación científica. Se enuncia habitualmente como

$$\pi(x) \sim \frac{x}{\log x} \tag{1.1}$$

donde $\pi(x) = \#\{p \text{ primo } \leq x\}$ y $f \sim g$ significa $\lim_{x\to+\infty} f(x)/g(x) = 1$. Al revisar algunas tablas es fácil percatarse de que $x/\log x$ no es numéricamente una buena aproximación de $\pi(x)$ aunque el error relativo tienda a cero. De hecho se conoce que a la larga $\pi(x) - x/\log x \geq x/\log^2 x$ por tanto el error relativo no decae más rápido que $1/\log x$ que es comparable al inverso del número de cifras. Otra forma de enunciar el teorema más complicada pero más natural desde el punto de vista histórico, computacional y teórico, requiere introducir la función *logaritmo integral* Li, y es

$$\pi(x) \sim \mathrm{Li}(x) \qquad \text{donde } \mathrm{Li}(x) = \int_2^x \frac{dt}{\log t}. \tag{1.2}$$

Equivale a (1.1) porque $x/\log x \sim \mathrm{Li}(x)$ aplicando la regla de L'Hôpital. Las tablas muestran que la aproximación es sorprendentemente buena en este caso y además se conoce que $\pi(x) - \mathrm{Li}(x)$ toma valores negativos y positivos para x arbitrariamente grandes, por tanto no muestra el sesgo positivo de la aproximación anterior. La siguiente tabla compara los errores relativos en ambas aproximaciones

	$x = 10^2$	$x = 10^4$	$x = 10^6$	$x = 10^8$	$x = 10^{10}$
$\pi(x)/(x/\log x) - 1$	0.151292	0.131950	0.084489	0.061299	0.047797
$\pi(x)/\mathrm{Li}(x) - 1$	-0.140331	-0.012924	-0.001634	-0.000130	$-6.81552\cdot10^{-6}$

Teniendo en cuenta que $\mathrm{Li}(x+\epsilon) - \mathrm{Li}(x) = \epsilon/\log\xi$ con $\xi \in [x, x+\epsilon]$, el teorema de los números primos sugiere que la probabilidad de que un número de tamaño N sea primo es $1/\log N$. Sin embargo hay que tomar con precaución esta idea probabilística sobre todo cuando se combina con hipótesis de independencia pues aplicada sin cuidado conduce a conclusiones equivocadas.

1.2. Notas históricas

C.F. Gauss fue el primero en notar que $\mathrm{Li}(x)$ aproxima a $\pi(x)$ pero no escribió ningún resultado teórico a este respecto. P.L. Chebyshev realizó a mediados del siglo XIX los primeros grandes avances (véase [50]) mediante ingeniosos argumentos con números combinatorios. Por ejemplo, después de simplificar está claro que $\binom{2n}{n}$ es divisible por el producto de los primos entre $n+1$ y $2n$ y es de esperar que no tenga muchísimos factores más. Sus razonamientos le permitieron deducir, entre otras cosas,

$$C_1 \frac{x}{\log x} \leq \pi(x) \leq C_2 \frac{x}{\log x} \qquad \text{para } x \geq C_3 \tag{1.3}$$

con ciertas constantes C_1, C_2 y C_3 explícitas. Con ello logró demostrar el *postulado de Bertrand* (para todo $n > 1$ existe un primo entre n y $2n$) que era uno de sus objetivos. Además los argumentos elementales de Chebyshev le permitieron probar que si $\lim(\pi(x)\log x)/x$ existe, debe valer uno. Aunque este enunciado esté formalmente muy cerca de (1.1), la profundidad que alberga la existencia del límite aleja fundamentalmente sus demostraciones.

El punto de inflexión vino en 1859 con la famosa memoria de B. Riemann, su único trabajo en teoría de números ([17] contiene una traducción al inglés). En ella se da un esquema de la prueba del teorema de los números primos pero hay lagunas fundamentales que se tardaron 40 años en cubrir y la llamada Hipótesis de Riemann todavía sigue sin resolver. Incluso sin constituir una prueba del teorema, la memoria de Riemann es un trabajo importantísimo que introduce la función ζ, una función meromorfa ligada desde entonces a la distribución de los primos.

En 1896 J. Hadamard y C. de la Vallée Poussin demostraron independientemente el teorema de los números primos siguiendo el esquema de la memoria de Riemann. Un punto fundamental fue la no anulación de la función ζ en cierta región.

Hay una relación en los dos sentidos entre la acotación del error $\pi(x) - \mathrm{Li}(x)$ y la distribución de los ceros de ζ estudiada por métodos de variable compleja, por ello fue una gran sorpresa que en 1948 P. Erdős y A. Selberg encontraran una prueba elemental (pero no sencilla) del teorema de los números primos. Su interés teórico no es comparable al de la prueba clásica porque da menos información acerca del error. Por cierto, la mejor acotación para el orden del error está lejos de ser elemental y se debe a N.M. Korobov e I.M. Vinogradov. No ha habido ningún avance en este sentido desde 1958.

1.3. La función ζ de Riemann

Definimos en primer lugar la *función ζ de Riemann*[1] en $\Re s > 1$ como

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}. \tag{1.4}$$

Es una función holomorfa en esta región porque la serie converge uniformemente sobre compactos (su módulo está acotado por $\sum n^{-\alpha}$ con $\alpha = \Re(s)$).

A partir de la fórmula $(1-x)^{-1} = 1 + x + x^2 + \dots$ para $|x| < 1$ y el teorema fundamental de la aritmética (factorización única en primos) se deduce el *producto de Euler*

$$\zeta(s) = \prod_{p} \left(1 - p^{-s}\right)^{-1} \tag{1.5}$$

donde p recorre los primos. Tal fórmula fue probada por L. Euler en 1737 y su interés radica en que a la izquierda tenemos una expresión que depende de simples números naturales y a la derecha otra que depende de la misteriosa sucesión de primos. La derivada logarítmica de (1.5) y de nuevo el desarrollo de Taylor de $(1-x)^{-1}$ dan lugar a la fórmula

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} \tag{1.6}$$

donde $\Lambda(n)$ es la función llamada *símbolo de von Mangoldt*, que vale $\log p$ si $n = p^k$, $k \in \mathbb{Z}^+$ y cero en otro caso.

En cierto modo para demostrar el teorema de los números primos se busca "despejar" $\pi(x)$ en función de $\zeta(s)$ a partir de (1.5) para lo cual se muestra conveniente pasar por (1.6). Curiosamente este esquema requiere extender la función ζ más alla del semiplano $\Re s > 1$ donde (1.4) no es válida. Está claro que $\lim_{x\to 1^+} \zeta(x) = \infty$ por tanto hay una singularidad en $s = 1$. No es difícil probar sumando por partes que para $\Re s > 1$

$$\zeta(s) - \frac{1}{s-1} = 1 + s \int_1^{\infty} \big([t] - t\big) t^{-s-1}\, dt \tag{1.7}$$

pero la integral converge si $\Re s > 0$ y define una función holomorfa en esta región. Integrando por partes sucesivas veces se demuestra que $\zeta(s) - 1/(s-1)$ se extiende a una función entera (holomorfa en todo $\mathbb{C}$).

Nótese que (1.7) es numéricamente más eficiente que (1.4) en $\Re(s) > 1$ donde ambas son aplicables, sobre todo cuando s está cerca de 1. Esta eficiencia se acentúa tras la integración por partes. La generalización de este método, con consecuencias en la evaluación numérica se series, es la *fórmula de sumación de Euler-Maclaurin* [51] [31].

[1] En inglés la letra ζ se escribe *zeta* y se pronuncia ['zi:tə]. Según el diccionario de la RAE en castellano deberíamos decir *dseda* pero casi toda la comunidad científica dice *zeta*. El caso es similar a μ y ν que según el diccionario son *mi* y *ni*.

1.4. El esquema de la prueba clásica

Una fórmula analítica para ψ

Aunque en principio deseamos obtener $\pi(x)$ a partir de (1.5) y es posible proceder de este modo, resulta técnicamente más cómodo obtener la función introducida por Chebyshev

$$\psi(x) = \sum_{n \le x} \Lambda(n) \tag{1.8}$$

a partir de (1.6). No es difícil probar $\pi(x) = \sum_{2 \le n \le x} \Lambda(n)/\log n + O\big(x^{1/2}\big)$. De aquí si $\Lambda(n)$ es 1 en promedio se tiene el teorema de los números primos, más precisamente sumando por partes

$$\pi(x) - \mathrm{Li}(x) = \frac{\psi(x) - x}{\log x} + \int_2^x \frac{\psi(t) - t}{t \log^2 t}\, dt + O\big(x^{1/2}\big). \tag{1.9}$$

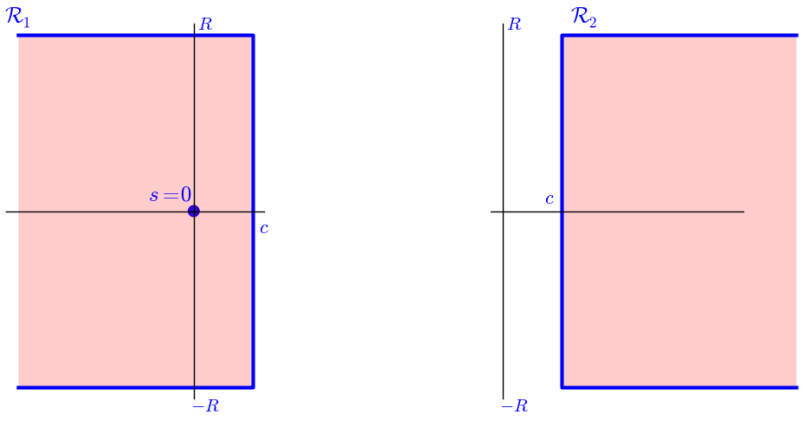
Todo lo que hay que probar es entonces

$$\psi(x) \sim x. \tag{1.10}$$

La manera más rápida de expresar $\psi(x)$ en términos de (1.6) pasa por aproximar la integral

$$I(a) = \frac{1}{2\pi i} \int_{(c \pm iR)} \frac{a^s}{s}\, ds \tag{1.11}$$

mediante el teorema de los residuos, donde $(c \pm iR)$ indica el segmento $\Re s = c$, $|\Im s| \le R$ y $a, c \in \mathbb{R}^+$. Si $a > 1$ se considera la región $\mathcal{R}_1 = \{\Re s \le c,\ |\Im s| \le R\}$ que contiene el polo en $s = 0$ del integrando con residuo 1, mientras que si $a < 1$ se considera $\mathcal{R}_2 = \{\Re s \ge c,\ |\Im s| \le R\}$ que no contiene polos.



De aquí

$$I(a) = \begin{cases} 1 + \texttt{error} & \text{si } a > 1 \\ \texttt{error} & \text{si } a < 1 \end{cases} \tag{1.12}$$

donde $\texttt{error} = O\big((R \log a)^{-1}\big)$ es la contribución de las fronteras horizontales. Así pues para cada $c > 1$ empleando (1.6)

$$\psi(x) = \frac{1}{2\pi i} \int_{(c \pm iR)} F(s)\, ds + \texttt{Error} \qquad \text{donde } F(s) = -\frac{x^s \zeta'(s)}{s\zeta(s)} = \sum_{n=1}^{\infty} \Lambda(n) \frac{(x/n)^s}{s} \tag{1.13}$$

con $\lim_{R\to\infty} \texttt{Error} = 0$ para $x > 1$ no entero. Esta última condición es para huir del caso $I(1)$ que no hemos calculado. Como $\psi(x) - \psi(x - \delta) \leq \log x$ si $\delta \leq 1$, siempre podemos suponer por ejemplo que x es semientero ($2x$ impar) a la hora de probar (1.10). En esa situación algunas cotas básicas para la función ζ (véase [13]) permiten deducir que $\texttt{Error} = O(xR^{-1} \log^2(Rx))$.

La relación con los ceros

Con (1.13) ya hemos despejado $\psi(x)$ en términos de ζ y ahora tenemos que aproximar la integral. Aplicamos de nuevo el teorema de los residuos en la región $\mathcal{R}_1$. La función F tiene polos en $s = 1$, $s = 0$ y en $s = z$ donde $\zeta(z) = 0$. Los residuos son

$$\operatorname{Res}(F, 1) = x, \qquad \operatorname{Res}(F, 0) = -\frac{\zeta'(0)}{\zeta(0)} \quad \text{y} \quad \operatorname{Res}(F, z) = -k\frac{x^z}{z} \tag{1.14}$$

con $k = k(z)$ la multiplicidad de z. Este cálculo se reduce a emplear $\zeta(s) \sim (s-1)^{-1}$ si $s \to 1$ y $\zeta(s) \sim K(s - z)^k$ si $s \to z$.

Eligiendo R tal que $\Im z \neq \pm R$ se deduce

$$\psi(x) = x - \frac{\zeta'(0)}{\zeta(0)} - \sum_{|\Im z| < R} \frac{x^z}{z} + \texttt{Error} \tag{1.15}$$

donde convenimos repetir cada cero de acuerdo con su multiplicidad.

Riemann demostró una simetría de ζ (la *ecuación funcional* por antonomasia) de la que se deduce que los ceros con $\Re z < 0$ son simples y están situados en $z = -2, -4, -6, \ldots$ mientras que el resto, llamados *ceros no triviales*, están en la *banda crítica* $0 \leq \Re s \leq 1$ y cumplen que si z es un cero, $1 - z$ también lo es. Entonces (1.15) se reescribe usando el desarrollo de Taylor $-\log\big(1 - x^{-2}\big) = \sum (x^{-2})^n/n$ como

$$\psi(x) = x - \sum_{|\Im \rho| < R} \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2} \log\big(1 - x^{-2}\big) + \texttt{Error} \tag{1.16}$$

donde ρ recorre los ceros no triviales.

Resultados condicionales e incondicionales

Para probar (1.10) y por tanto el teorema de los números primos los términos en (1.16) con $\zeta'(0)/\zeta(0)$ y el logaritmo son claramente irrelevantes. Tomando $R \asymp x$ se obtiene $\texttt{Error} = O\big(\log^2 x\big)$ que también es irrelevante.

Si hubiera aunque fuera un solo cero con $\Re\rho = 1$ entonces $\lim \psi(x)/x$ no existiría (porque $x^{1+i\alpha}/x = x^{i\alpha}$ oscila) y el teorema de número primo no sería cierto. Por otro lado, se conoce que $\sum_{|\Im\rho|<R} 1/\rho = O\big(\log^2 R\big)$, por tanto si todos los ceros cumplieran $\Re\rho \le \alpha_0$ entonces se tendría

$$\psi(x) = x + O\big(x^{\alpha_0}\log^2 x\big) \tag{1.17}$$

y de (1.9)

$$\pi(x) = \mathrm{Li}(x) + O\big(x^{\alpha_0}\log x\big). \tag{1.18}$$

Dada la simetría $\rho \leftrightarrow 1-\rho$ el mejor caso posible ocurriría cuando $\alpha_0 = 1/2$, es decir, si todos los ceros no triviales cumplieran $\Re\rho = 1/2$. Ésta es la famosa *hipótesis de Riemann*, todavía sin resolver, que daría el error más pequeño en el teorema de los números primos (en lo que se refiere al exponente) y tendría numerosas consecuencias en teoría de números.

Volviendo a la realidad, no se conoce ningún $\alpha_0 < 1$ para el que $\Re\rho \le \alpha_0$. ¿Entonces cómo concluir la prueba a partir de (1.16)? Se demuestra que si $1-\Re\rho$ fuera realmente muy pequeño entonces $|\Im\rho|$ sería tan grande que no aparecería en la suma de (1.16). Más concretamente, la forma habitual de esta afirmación, ya conocida por de la Vallée Poussin en 1899, es que la región

$$1 - \Re s < \frac{K}{\log(|\Im s| + 2)} \tag{1.19}$$

no contiene ceros de ζ donde es cierta constante (computable efectivamente). Esto implica que $\Re\rho \le 1 - K/\log(R+2)$ en (1.16) y entonces $x^\rho = O\big(xe^{-K\log x/\log(R+2)}\big)$. Esto daría un insuficiente $O(x)$ si $R \asymp x$ pero tomando un R mucho más pequeño, por ejemplo, $R \asymp e^{\sqrt{\log x}}$, todavía se tiene que `Error` no contribuye decisivamente y por tanto

$$\psi(x) = x + O\big(xe^{-K'\sqrt{\log x}}\big) \tag{1.20}$$

para cierta $K' > 0$ (que prácticamente nunca se calcula explícitamente). De aquí se concluye (1.10) y a través de (1.9) se aproxima $\pi(x)$ por $\mathrm{Li}(x)$ con el mismo tipo de término de error.

Los trabajos de Korobov y Vinogradov permiten reemplazar $\sqrt{\log x}$ por algo ligeramente menor que $(\log x)^{3/5}$. Es un poco frustrante que más de un siglo después de las primeras pruebas del teorema de los números primos, toda la mejora en el término de error sea ese aumento de un décima en la potencia de logaritmo.

1.5. Variantes de la prueba clásica

Esencialmente hay dos pruebas del teorema de los números primos, la clásica que hemos visto aquí y la elemental. Sin embargo la bibliografía induce a ver más bien dos familias de pruebas. Comentaremos aquí algunas variantes de la prueba clásica.

Las fórmulas básicas del análisis complejo tienen su contrapartida en el análisis de Fourier y es cuestión de gustos emplear uno u otro lenguaje. Por ejemplo, si $f(z) = a_0 + a_1 z + a_2 z^2 + \dots$ la fórmula de Cauchy

$$a_n = \frac{1}{2\pi i} \int_{S^1} \frac{f(z)}{z^{n+1}} dz \tag{1.21}$$

equivale al emplear la parametrización $z = e(x)$, $0 \leq x \leq 1$ de S^1 a la fórmula de los coeficientes de Fourier $a_n = \int_0^1 F(x) e(-nx)\, dx$ para $F(x) = a_0 + a_1 e(x) + a_2 e(2x) + \dots$ En esta línea, es posible reescribir la forma de extraer $\psi(x)$ a partir de $\sum \Lambda(n) n^{-s}$ en términos de transformadas de Fourier. Para conseguir el teorema sin preocuparse demasiado por el término de error todo lo que se necesita es que $\zeta'(s)/\zeta(s)$ no sea "mala" en $\Re s = 1$, en particular ζ no debe tener ceros allí. Un ejemplo de este esquema se encuentra en [16].

Se llaman *teoremas tauberianos* a los que recuperan el comportamiento de una suma (habitualmente de coeficientes en el desarrollo de una función) en términos de sumas regularizadas (habitualmente valores de una función cerca de una singularidad). El método del párrafo anterior permite deducir que $\sum_{n \leq x} a_n \sim x$ para $a_n \geq 0$ siempre que $\sum a_n n^{-s} \sim (s-1)^{-1}$ cuando $s \to 1^+$ y la diferencia de estas cantidades tenga una extensión analítica a $\Re s \geq 1$. Una forma un poco más general es lo que se llama *teorema tauberiano de Wiener–Ikehara*. Empleando teoremas tauberianos a ciegas es posible deducir muy rápido el teorema de los números primos, como reza el título de [35], sin embargo las pruebas no son autocontenidas.

D.J. Newman mostró en 1980 cómo simplificar drásticamente la prueba del teorema de los números primos con el esquema clásico (siempre a costa de perder precisión en el término de error). Su exposición es muy conveniente para cursos básicos de teoría de números por su brevedad y claridad, virtudes compartidas por el libro [41] que la contiene. Hay también una demostración breve de H. Iwaniec, incluida en [32], que tiene la sorprendente novedad de que no emplea la extensión holomorfa de ζ.

Capítulo 2

El teorema de los números primos en progresiones aritméticas

2.1. Enunciado y demostración

Consideramos una progresión aritmética $\{qn+a\}_{n=1}^{\infty}$ con $q \in \mathbb{Z}^+$ y $a \in \mathbb{Z}$ y queremos estimar el número $\pi(x;q,a)$ de primos en ella que son menores o iguales que un número grande x. Dados a_1 y a_2 congruentes módulo q, los conjuntos $\{qn+a_1\}_{n=1}^{\infty}$ y $\{qn+a_2\}_{n=1}^{\infty}$ difieren sólo en un número finito de elementos. Por tanto, si pensamos en el caso en que x sea grande en comparación con a, podemos restingirnos a $1 \leq a \leq q$.

Claramente

$$\bigcup_{a=1}^{q} \{qn+a\}_{n=1}^{\infty} = [q+1,\infty) \cap \mathbb{Z} \tag{2.1}$$

por tanto cualquier primo $p > q$ está en alguna de las progresiones aritméticas de la unión anterior. El teorema de los números primos implica entonces que si $q = o(x)$ se tiene

$$\sum_{a=1}^{q} \pi(x;q,a) = \pi(x) - \pi(q) \sim \pi(x) \sim \mathrm{Li}(x). \tag{2.2}$$

Obviamente si a y q no son coprimos no hay primos de la forma $qn+a$, entonces fijada una diferencia q sólo hay $\phi(q)$ progresiones aritméticas que pueden contribuir significativamente a la suma anterior. El teorema de los números primos en progresiones aritméticas afirma que asintóticamente los primos están igualmente distribuidos en ellas. Dicho de otro modo, no hay ninguna progresión aritmética favorita para los primos.

Teorema (de los números primos en progresiones aritméticas): *Fijados $q \in \mathbb{Z}^+$ y $a \in \mathbb{Z}$, se cumple*

$$\pi(x;q,a) \sim \frac{\mathrm{Li}(x)}{\phi(q)} \qquad \textit{cuando } x \to \infty. \tag{2.3}$$

Es posible sustituir $\mathrm{Li}(x) = \int_2^x dt/\log t$ por $x/\log x$ ya que ambas cantidades son asintóticamente equivalentes, pero es conocido que la primera función da una aproximación más precisa.

La demostración sigue las líneas de la del teorema de los números primos (ver [8]). En ella se partía de la siguiente consecuencia del producto de Euler:

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} \qquad \text{para } \Re(s) > 1, \tag{2.4}$$

y aplicando métodos de variable compleja, especialmente el teorema de los residuos y la extensión meromorfa de ζ, se conseguían despejar las sumas parciales de los coeficientes de la serie de Dirichlet del segundo miembro de (2.4) en términos de los ceros de la función ζ. El resultado concreto era que para $2 \le T \le x$

$$\psi(x) = x - \sum_{|\rho|<T} \frac{x^\rho}{\rho} + O\big(\frac{x}{T}(\log x)^2\big) \qquad \text{con } \psi(x) = \sum_{n \le x} \Lambda(n). \tag{2.5}$$

Aquí ρ recorre los ceros no triviales de ζ, los que cumplen $0 < \Re(\rho) < 1$, y T es un parámetro ajustable que se elige en función de cuánta información se tenga de los ceros. La región libre de ceros conocida hoy en día permite elegir T para deducir el teorema de los números primos en la forma $\psi(x) \sim x$ y extraer un débil término de error.

En el caso de progresiones aritméticas, para llegar a un análogo de (2.4) hay que partir de un producto de Euler, lo cual requiere una función multiplicativa. El problema es entonces detectar una progresión aritmética con funciones multiplicativas. Lo que lleva a inventar los *caracteres de Dirichlet*, funciones multiplicativas de periodo q que para cada a primo con q satisfacen la relación de ortogonalidad

$$\frac{1}{\phi(q)} \sum_{\chi} \overline{\chi}(a)\chi(n) = \begin{cases} 1 & \text{si } n \equiv a \pmod q \\ 0 & \text{si } n \not\equiv a \pmod q \end{cases} \tag{2.6}$$

y $\chi(a) = 0$ si a y q no son coprimos.

La *congruencia de Euler-Fermat* $a^{\phi(q)} \equiv 1 \pmod q$ implica que $\chi(a)$ es una raíz $\phi(q)$-ésima de la unidad (siempre bajo el supuesto de que a y q son coprimos). Esto y la propiedad

multiplicativa, es suficiente para construir todos los caracteres. Por ejemplo, para $q = 4$, $\chi(3) = \pm 1$ porque $3^2 \equiv 1 \pmod 4$ y los valores $\chi(1) = 1$, $\chi(0) = \chi(2) = 0$ vienen forzados. Entonces hay exactamente dos caracteres, uno por cada signo de $\chi(3)$. Para $q = 5$, se tiene que por ejemplo $a = 2$ es un generador de todo el grupo multiplicativo $1 \equiv 2^0$, $2 \equiv 2^1$, $3 \equiv 2^3$, $4 \equiv 2^2$. Entonces escogiendo como $\chi(2)$ una raíz cuarta de la unidad, obtenemos los cuatro caracteres posibles.

	0	1	2	3
χ_1	0	1	0	1
χ_2	0	1	0	-1

módulo 4

	0	1	2	3	4
χ_1	0	1	1	1	1
χ_2	0	1	i	$-i$	-1
χ_2	0	1	i	$-i$	-1
χ_3	0	1	-1	-1	1
χ_4	0	1	$-i$	i	-1

módulo 5

Es posble probar con técnicas elementales o con otras más abstractas y rápidas que módulo q hay exactamente $\phi(q)$ caracteres.

A cada carácter χ módulo q se le asocia un producto de Euler, su función L,

$$L(s,\chi) = \prod_p \big(1 - \chi(p)p^{-s}\big)^{-1} \qquad \text{para } \Re(s) > 1, \tag{2.7}$$

de esta forma

$$-\frac{L'(s,\chi)}{L(s,\chi)} = \sum_{n=1}^{\infty} \chi(n)\frac{\Lambda(n)}{n^s} \qquad \text{para } \Re(s) > 1. \tag{2.8}$$

Si χ_0 es el carácter que cumple $\chi_0(a) = 1$ para todo a coprimo con q, llamado *carácter principal*, entonces $L(s,\chi_0)$ es igual a $\zeta(s)$ salvo unos pocos factores correspondientes a primos que dividen a q, por ello $L(s,\chi_0)$ hereda el polo simple en $s = 1$ de ζ. Por otro lado para $\chi \neq \chi_0$ las funciones $L(s,\chi)$ son enteras utilizando el mismo argumento que da la extensión meromorfa de ζ. Por ello hay dos casos en el análogo de (2.5), válido para $2 \le T \le x$ y uniformemente en $q \le x$

$$\psi(x,\chi) = c(\chi)x - \sum_{|\rho|<T} \frac{x^\rho - 1}{\rho} + O\big(\frac{x}{T}(\log x)^2\big) \qquad \text{con} \qquad c(\chi) = \begin{cases} 1 & \text{si } \chi = \chi_0 \\ 0 & \text{si } \chi \neq \chi_0 \end{cases} \tag{2.9}$$

donde ahora ρ recorre los ceros $0 < \Re\rho < 1$ de la función L correspondiente y

$$\psi(x,\chi) = \sum_{n \le x} \chi(n)\Lambda(n). \tag{2.10}$$

El 1 que resta a x^ρ en (2.9) es conjeturalmente prescindible, se introduce para preservar la uniformidad incluso si hubiera algún ρ muy cercano a cero. Nótese que la función x^s/s es singular en $s=0$ mientras que $(x^s-1)/s$ se puede definir con continuidad en dicho punto.

Fijado q, nuestro conocimiento sobre los ceros de las funciones L y de la función ζ en la banda crítica es a grandes rasgos similar, lo que permite que la misma elección de T que en (2.5) daba en teorema de los números primos, en (2.9) produzca $\psi(x,\chi_0)\sim x$ y $\psi(x,\chi)=o(x)$ si $\chi\neq\chi_0$. Recordando (2.6) se deduce para cada q fijo

$$\psi(x;q,a)\sim\frac{x}{\phi(q)}\qquad\text{donde}\qquad\psi(x;q,a)=\sum_{\substack{n\le x\\ n\equiv a\ (q)}}\Lambda(n). \tag{2.11}$$

Sumando por partes se llega al teorema de los números primos en progresiones aritméticas.

2.2. Uniformidad y distribución de los ceros

En muchos problemas aparecen progresiones aritméticas con la diferencia q variando. Esto ocurre por ejemplo en los mayores avances hacia la conjetura de Goldbach o la conjetura de los primos gemelos. Sin entrar en detalles, digamos que para buscar primos gemelos p, $p+2$, podríamos quitar de entre los $p\in[N^2/2,N^2)$ aquellos tales que $p+2$ es divisible por algún $q\le N$, lo que conduce a $\pi(N^2,q,-2)-\pi(N^2/2,q,-2)$ con q variable. En este contexto se muestra crucial la uniformidad en q en el teorema de los números primos en progresiones aritméticas.

Las funciones L de caracteres primitivos satisfacen una ecuación funcional similar a la de la función ζ, la cual implica una simetría $\rho\leftrightarrow 1-\overline{\rho}$ en los ceros $0<\Re\rho<1$ de cada función L, entonces dado $x\ge 2$ se tiene $\sup_\rho|x^\rho|\ge x^{1/2}$. Además el caso imprimitivo (en el cual no se cumple la ecuación funcional) no añade nuevos ceros en esta región. Entonces lo mejor que podemos esperar para estimar (2.9) es la siguiente conjetura:

Hipótesis de Riemann generalizada: Si $L(\rho,\chi)=0$ con $0<\Re\rho<1$ entonces $\Re(\rho)=1/2$.

Bajo esta hipótesis, tomando $T=x^{1/2}$ en (2.9) y empleando que en cada rectángulo $k\le|\Im(s)|\le k+1$ de la banda crítica hay $O\big(\log(q(k+1))\big)$ ceros, se deduce

$$\psi(x;q,a)=\frac{x}{\phi(q)}+O(x^{1/2}(\log x)^2)\qquad\text{y}\qquad\pi(x;q,a)=\frac{\mathrm{Li}(x)}{\phi(q)}+O(x^{1/2}\log x) \tag{2.12}$$

uniformemente en q.

La segunda fórmula proviene de sumar por partes en la primera. El caso del teorema de los números primos bajo la hipótesis de Riemann corresponde formalmente a $q=1$ y $a=0$.

Nótese que las fórmulas (2.12) son triviales para $q \geq x^{1/2}$, lo cual sugiere que tendríamos que tener un control finísimo sobre los ceros, más allá de la hipótesis de Riemann generalizada, para poder explorar este rango. La conjetura es que el teorema de los números primos en progresiones aritméticas se sigue cumpliendo uniformemente para q menor que x elevado a un exponente menor que 1. Concretamente, H.L. Montgomery ha sugerido que para cualquier $\epsilon > 0$

$$\pi(x; q, a) = \frac{\mathrm{Li}(x)}{\phi(q)} + O\big(x^{\epsilon}\sqrt{\frac{x}{q}}\big) \qquad \text{uniformemente en } q \leq x. \tag{2.13}$$

Por otra parte, en [19] se prueba que el teorema de los números primos en progresiones aritméticas no puede ser cierto para todo $q < x/(\log x)^A$.

La búsqueda de resultados uniformes incondicionales choca con el insidioso fenómeno de los posibles ceros excepcionales, que describimos a continuación.

La forma clásica del término de error en el teorema de los números primos en progresiones aritméticas es

$$\pi(x) = \mathrm{Li}(x) + O\big(xe^{-C\sqrt{\log x}}\big) \tag{2.14}$$

para cierta constante positiva C. Sabiendo que hay $O\big(\log(k+1)\big)$ ceros no triviales ρ de la función ζ con $k \leq |\Im(\rho)| \leq k+1$, la prueba se reducía a utilizar en (2.5) que la región

$$\left\{\sigma + it \,:\, \sigma > 1 - \frac{C}{\log(|t|+2)}\right\} \tag{2.15}$$

está libre de ceros, donde C es una constante positiva (distinta de la del teorema anterior). Este hecho se derivaba de una desigualdad trigonométrica que conducía a

$$-3\frac{\zeta'(\sigma)}{\zeta(\sigma)} - \Re\frac{\zeta'(\sigma+2it)}{\zeta(\sigma+2it)} \geq 4\Re\frac{\zeta'(\sigma+it)}{\zeta(\sigma+it)} \qquad \text{para } \sigma > 1. \tag{2.16}$$

Un cero $\rho = \beta + it$ de ζ es un polo de ζ'/ζ y si estuviera muy cercano a $\Re(s) = 1$ permitiría elegir σ tal que $\Re\big(\zeta'(\sigma+it)/\zeta(\sigma+it)\big)$ es tan grande que contradiría la desigualdad anterior.

Para las funciones L el razonamiento es similar utilizando

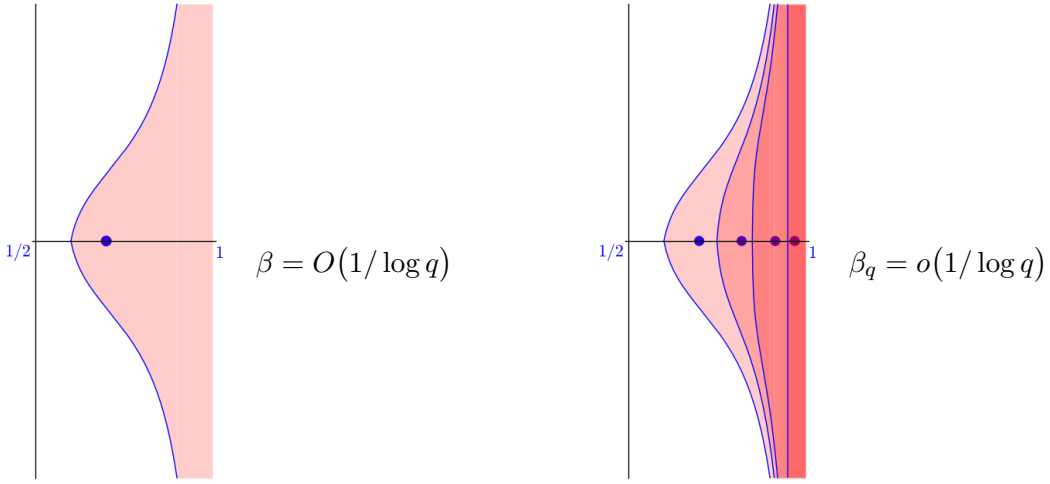
$$-3\frac{L'(\sigma,\chi_0)}{L(\sigma,\chi_0)} - \Re\frac{L'(\sigma+2it,\chi^2)}{L(\sigma+2it,\chi^2)} \geq 4\Re\frac{L'(\sigma+it,\chi)}{L(\sigma+it,\chi)} \qquad \text{para } \sigma > 1 \tag{2.17}$$

donde χ_0 es el carácter principal. Si χ es un carácter real, es decir, si $\chi^2 = \chi_0$, entonces esta fórmula no da información para $t = 0$ y σ cercano a 1, como tampoco la daba (2.16), la diferencia es que es elemental (y relativamente fácil) probar que $\zeta(\beta) \neq 0$ para $\beta > 1/2$ pero

nadie ha conseguido probarlo para todas las funciones L. El resultado conocido es entonces que existe un C tal que para cualquier función $L(s,\chi)$ la región

$$\left\{\sigma+it\,:\,\sigma>1-\frac{C}{\log(q(|t|+2))}\right\}. \tag{2.18}$$

contiene a lo más un cero, llamado *cero excepcional* (o cero de Siegel), de dicha función L. Tal posible cero, si existe, es real y simple y sólo puede aparecer cuando χ es real y no principal.



cero excepcional variación de q

Nótese que la definición de cero excepcional depende en rigor de la constante elegida. Si los ceros excepcionales para cierta constante sólo ocurriesen para un número finito de funciones L y el mayor de ellos estuviera a distancia 0.02011 de 1, entonces dejarían de existir eligiendo $C = 0.02011/2$. Por ello al referirse a ceros excepcionales se piensa más bien en sucesiones de ceros β_q con $(1-\beta_q)\log q \to 0$.

Si apareciera un cero excepcional, digamos en $s=\beta$, entonces para $q < e^{\sqrt{\log x}}$ se tendría

$$\psi(x,\chi) = \frac{x^\beta}{\beta} + O\big(xe^{-C\sqrt{\log x}}\big). \tag{2.19}$$

Si β pudiera estar muy cerca de 1, por ejemplo a distancia $o\big((\log x)^{-1/2}\big)$ entonces $\psi(x,\chi) \neq O\big(xe^{-C\sqrt{\log x}}\big)$ y no podríamos deducir el análogo de (2.14) en progresiones aritméticas.

El teorema de Siegel (ver más adelante) permite controlar la distancia de β a 1 de manera satisfactoria si q no soprepasa una potencia de logaritmo de x. Es decir, fijado $A>0$, se cumple

$$\pi(x;q,a) = \frac{\mathrm{Li}(x)}{\phi(q)} + O\big(xe^{-C\sqrt{\log x}}\big) \qquad \text{uniformemente en } q \le (\log x)^A. \tag{2.20}$$

A este resultado se le llama *teorema de Siegel-Walfisz*. La constante O depende de A de una manera desconocida.

2.3. Algunos resultados relacionados

En esta sección veremos tres teoremas relacionados con la uniformidad:

1. El teorema de Siegel (§21 [13], Th.5.28 [32]).
2. El teorema de Brun-Titchmarsh (Ch.13 [20], Th.6.6 [32]).
3. El teorema de Bombieri-Vinogradov (§28 [13], Th.17.1 [32]).

Ya hemos mencionado que el primero aparece en la prueba de (2.20), la versión uniforme del teorema de los números primos en progresiones aritméticas. Los otros dos permiten aumentar el rango de q a cambio de perder la asintótica.

Teorema (de Siegel): *Para cada* $\epsilon > 0$ *existe* $C(\epsilon) > 0$ *tal que para todo* χ *carácter primitivo real módulo* q *se cumple* $L(\sigma, \chi) > 0$ *cuando* $\sigma > 1 - C(\epsilon)q^{-\epsilon}$.

Significado: Los probablemente inexistentes ceros que no verifican la hipótesis de Riemann generalizada estropearían el término de error esperado (2.12) en el teorema de los números primos en progresiones aritméticas. Entre ellos, los ceros excepcionales son especialmente perniciosos porque limitan la dependencia en la diferencia q de la progresión aritmética, y lo son más cuanto más cerca estén de $s = 1$ en términos de q. El teorema asegura que no existe ninguno a distancia menor que una potencia negativa de q, lo cual permite salvar un poco de la uniformidad en los términos expresados en el teorema de Siegel-Walfisz.

Comentarios: Uno de los aspectos más singulares del teorema es que fijado $0 < \epsilon < 1/2$, la constante $C(\epsilon)$ no es efectiva con los conocimientos actuales. Es decir, no se conoce una manera de dar un valor válido de por ejemplo $C(0.02011)$. La razón para ello es que en la prueba del teorema hay dos casos, dependiendo de si alguna función $L(s, \chi)$ tiene ceros reales en cierto intervalo $I_\epsilon \subset (1/2, 1]$ o no. Si los tuviera, $C(\epsilon)$ estaría relacionada con el tamaño del módulo de χ.

Ideas sobre la demostración: La clave está en emplear que $\zeta(s)L(s, \chi)L(s, \chi')L(s, \chi\chi')$ tiene una serie de Dirichlet con coeficientes no negativos cualesquiera que sean χ y χ' caracteres primitivos reales. Esta serie además tiende a ∞ cuando $s \to 1$. Si $L(s, \chi')$ tuviera un cero muy cercano a $s = 1$, entonces el producto $L(s, \chi')\zeta(s)$ estaría acotado cerca de ese cero y para compensar, $L(s, \chi)L(s, \chi\chi')$ debería ser grande para todo χ. Hay una cota superior burda para $L(s, \chi\chi')$ que permite concluir algo acerca del tamaño de $L(s, \chi)$. De alguna manera lo que se prueba es que los ceros de las funciones L se repelen cerca de $s = 1$ y una que no cumpla el teorema de Siegel con cierta constante sirve para deducir el teorema para el resto de las funciones L. Finalmente la constante se reajusta para que la función L inicial no suponga una excepción.

Teorema (de Brun-Titchmarsh): *Para $x \geq 0$ y $1 \leq q < y$ se cumple*

$$\pi(x+y;q,a) - \pi(x;q,a) < \frac{2y}{\phi(q)\log(y/q)}. \tag{2.21}$$

En realidad el resultado se suele establecer en una forma más débil (con una constante mayor que 2 o con un término de error). Esta mejora de la forma inicial del teorema proviene de [39].

Significado: En el caso $q = 1$ ya nos dice que en intervalos pequeños no puede haber muchos más primos de lo que conjeturaríamos. Si todos los ceros de la función ζ estuvieran en $\Re(s) < \alpha$ entonces el teorema de los números primos tendría un error $O\big(x^\alpha \log x\big)$ y por tanto habría una fórmula asintótica para los primos en $[x, x + x^\beta]$ cuando $\beta > \alpha$. Tal fórmula no está probada para ningún $\beta < 1$ debido a nuestro escaso conocimiento de los ceros y el teorema anterior es un sustituto más débil. En el caso $x = 0$ con $q > 1$, el teorema implica que la uniformidad en q en rangos amplios no puede ir tan sumamente mal como para que la cantidad de primos sea de un orden diferente mayor que el sugerido por el teorema de los números primos.

Comentarios: La constante 2 es una barrera en los métodos empleados en la demostración y disminuirla tendría consecuencias interesantes relativas a los ceros excepcionales. Algunos avances han permitido reducirla en ciertos rangos (ver Th. 13.2 de [20]) pero no en general.

Ideas sobre la demostración: Para la formulación más débil a la que nos hemos referido tras el enunciado, la prueba depende completamente de métodos de criba. Concretamente, si aplicamos la criba de Selberg al conjunto formado por los $n \in (x, x+y]$ con $n \equiv a \pmod q$, se llega a la desigualdad del teorema salvo añadir $O\big(y/(\phi(q)\log^2(y/q))\big)$ al segundo miembro [32]. La criba de Selberg es una manera ingeniosa, basada en formas cuadráticas, de eliminar (cribar) muchos de los números compuestos de un conjunto. Hay todavía números compuestos que superan esa criba y por ello sólo ofrece una cota superior.

Teorema (de Bombieri-Vinogradov): *Para cualquier $A > 0$ existe $B > 0$ tal que*

$$\sum_{q \leq Q} \max_{y \leq x} \max_{(a,q)=1} \left| \pi(y;q,a) - \frac{\mathrm{Li}(y)}{\phi(q)} \right| \ll \frac{x}{(\log x)^A} \qquad \textit{con } Q = \frac{x^{1/2}}{(\log x)^B}. \tag{2.22}$$

Significado: Quizá en una primera lectura es mejor olvidarse de los máximos. Lo que dice el teorema es que tenemos control sobre el término de error en el teorema de los números primos en progresiones aritméticas al promediar sobre los módulos hasta algo menos de $\sqrt{x}$. Esto contrasta con que el teorema de Siegel-Walfisz (2.20) sólo permita decir algo sobre módulos individuales menores que una potencia de logaritmo. El teorema de Bombieri-Vinogradov sirve muchas veces como sustituto de la hipótesis de Riemann generalizada.

Comentarios: Si se promedia también sobre a entonces el resultado se llama *teorema de Barban-Davenport-Halberstam* [13], [32] y en ese caso se puede llevar Q hasta x. Este resultado es más fácil de probar pero menos útil en las aplicaciones.

Ideas sobre la demostración: La manera natural de proceder es utilizar (2.9) para relacionar el error con los ceros de las funciones L. Las primeras pruebas (debidas a A.I. Vinogradov y a E. Bombieri, independientemente) utilizaban este hecho y aplicaban resultados de densidad (basados en la gran criba) que afirman en cierto modo que la hipótesis de Riemann generalizada no puede violarse demasiado por demasiadas funciones L simultáneamente. Más tarde, con trabajos de P.X. Gallagher y R.C. Vaughan se comprobó que era más rápido olvidarse de los ceros y aplicar la gran criba directamente a $\psi(x,\chi)$. Sin entrar en detalles, las desigualdades llamadas de gran criba en este contexto implican que una expresión de la forma $\sum_n a_n\chi(n)$ debe tener cancelación para la mayor parte de los caracteres, sean cuales sean los coeficientes a_n.

Ilustraremos los dos últimos resultados siguiendo [11] para aplicarlos al *problema del divisor de Titchmarsh*, consistente en estimar la suma $\sum_{p\le x} d(p+a)$ para $a\neq 0$ fijado donde $d(\cdot)$ indica el número de divisores.

Utilizando que si q divide a n, entonces n/q también lo divide, uno puede restringirse a divisores menores que $\sqrt{x}$ y llegar a

$$\sum_{p\le x} d(p+a) = 2\sum_{q\le\sqrt{x}} \pi(x;q,-a) + O\big(\sqrt{x}\big). \tag{2.23}$$

Sea Q como en el teorema de Bombieri-Vinogradov, entonces

$$\sum_{q\le Q} \pi(x;q,-a) = \mathrm{Li}(x) \sum_{\substack{q\le Q\\(q,a)=1}} \frac{1}{\phi(q)} + O\Big(\frac{x}{\log^A x}\Big). \tag{2.24}$$

Por otro lado el teorema de Brun-Titchmarsh implica

$$\sum_{Q\le q\le\sqrt{x}} \pi(x;q,-a) \ll \frac{x}{\log x} \sum_{Q\le q\le\sqrt{x}} \frac{1}{\phi(q)}. \tag{2.25}$$

Sumando estas dos fórmulas y empleando que para cierta $C=C(a)>0$

$$\sum_{\substack{t\le Q\\(q,a)=1}} \frac{1}{\phi(q)} = C\log t + O(1), \tag{2.26}$$

cuya prueba es elemental (pero no inmediata, véanse los ejercicios 2, 3 y 4 de [11]), se deduce finalmente

$$\sum_{p\le x} d(p+a) = Cx + O\Big(\frac{x\log\log x}{\log x}\Big). \tag{2.27}$$

Capítulo 3

Formas cuadráticas binarias definidas positivas

3.1. Introducción

En nuestro contexto llamamos *forma cuadrática binaria* a

$$Q(x,y) = ax^2 + bxy + cy^2 \qquad \text{con} \quad a,b,c \in \mathbb{Z}. \tag{3.1}$$

Aquí únicamente consideraremos el caso definido positivo, que por el criterio de Sylvester corresponde a $d < 0 < a$ donde d es el *discriminante* $b^2 - 4ac$. También supondremos que las formas son *primitivas*, es decir, que $\text{mcd}(a,b,c) = 1$.

Se dice que Q *representa* a $n \in \mathbb{Z}^+$ si $Q(x,y) = n$ tiene solución para ciertos $x, y \in \mathbb{Z}$. Ya en los albores de la teoría de números se observó que los problemas de representación por formas cuadráticas binarias tienen una sorprendente profundidad y riqueza aritmética. Uno de los más conocidos es el resultado de Fermat de que para p primo impar

$$x^2 + y^2 = p \text{ tiene solución } x, y \in \mathbb{Z} \;\Leftrightarrow\; p \equiv 1 \pmod 4. \tag{3.2}$$

La igualdad para la norma de números complejos $|z|^2|w|^2 = |zw|^2$ permite multiplicar representaciones mediante la regla

$$(x^2 + y^2)(t^2 + u^2) = (xt - yu)^2 + (xu + yt)^2 \tag{3.3}$$

y con ello se llega a probar que $Q(x,y) = x^2 + y^2$ representa exactamente los enteros positivos de la forma kl^2 donde k no tiene factores primos $p \equiv 3 \pmod 4$. Este resultado, incluso en su form débil (3.2), aunque clásico y bien conocido no es sencillo de probar y aparentemente depende de trucos muy especiales para esta Q. Sin embargo experimentalmente parece tener réplicas para muchas otras formas cuadráticas de discriminante pequeño.

Euler atacó los casos x^2+2y^2 y x^2+3y^2 también anunciados por Fermat pero fue incapaz de probar que si $p \neq 2,5$ es primo

$$x^2+5y^2 = p \text{ tiene solución } x,y \in \mathbb{Z} \;\Leftrightarrow\; p \equiv 1,9 \pmod{20}. \tag{3.4}$$

Aquí además al tratar de pasar de primos a compuestos la multiplicatividad funciona de una forma extraña: los primos $p \equiv 3,7 \pmod{20}$ no son representables pero cualquier producto de un número par de ellos sí lo es. Así se tiene $x^2+5y^2 \neq 7$ y $x^2+5y^2 \neq 23$ pero $9^2+5\cdot 4^2 = 7\cdot 23$, lo cual tiene que ver con el aparatoso pero elemental análogo de (3.3)

$$(2x^2+2xy+3y^2)(2t^2+2tu+3u^2) = (2xt+xu+yt+3yu)^2+5(xu-yt)^2 \tag{3.5}$$

hallado por Lagrange [12, Ch.1].

Por otro lado, hay formas sencillas para las que los primos representados no admiten una caracterización por medio de congruencias simples, de hecho éste es el caso genérico. Por ejemplo $Q(x,y) = 2x^2+7y^2$ representa primos $p \neq 2,7$ que cumplen $p \equiv 1,9,15,23,25,39 \pmod{56}$ pero no a todos ellos, experimentalmente parece que se olvida de la mitad. Curiosamente resulta que éstos son justamente aquellos cuyo doble es representable.

	Primos $p \equiv 1,9,15,23,25,39 \pmod{56}$									
$2x^2+7y^2 = p$		71,79,113,		191, 193,		263,		337,		401,
$2x^2+7y^2 = 2p$	23,		127,137,151,		233,239,		281,		359,	

Es decir, los cálculos sugieren que para $p \neq 2,7$

$$2x^2+7y^2 = p \;\text{ ó }\; 2x^2+7y^2 = 2p \text{ con } x,y \in \mathbb{Z} \;\Leftrightarrow\; p \equiv 1,9,15,23,25,39 \pmod{56}. \tag{3.6}$$

Una buena parte de los misterios relativos a la aritmética de las formas cuadráticas binarias fueron desvelados por Gauss en los artículos más complicados de sus *Disquisitiones Arithmeticae* [22] publicadas en 1801. Esencialmente demostró que se las puede dotar de una estructura de grupo abeliano que permite resolver todos los problemas de representación que admitan una caracterización de los primos representados por medio de congruencias lineales. Hoy entendemos esta estructura de grupo a través de la teoría de ideales que surgió del trabajo de Kummer a mediados del XIX y fue desarrollada por Dedekind. Gauss también atisbó con sus leyes de reciprocidad cúbica y cuártica que en algunos casos se podía ir más allá utilizando congruencias polinómicas. La teoría de cuerpos de clases que nació a principios del siglo XX ha permitido dar una explicación general de este hecho [12].

3.2. Clases

Las formas x^2+y^2 y $x^2+2xy+2y^2$ representan los mismos enteros porque están ligadas por el cambio de variables invertible sobre $\mathbb{Z}$, $x \mapsto x+y$, $y \mapsto y$. Los cambios de variables lineales

invertibles sobre $\mathbb{Z}$ son en general los que tienen matrices de determinante ± 1. Gauss consideró ambos signos en relación con las formas cuadráticas pero las exposiciones modernas se olvidan del signo negativo y se centran en

$$\mathrm{SL}_2(\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \ : \ a,b,c,d \in \mathbb{Z}, \quad ad-bc=1 \right\}. \tag{3.7}$$

Diremos que dos formas cuadráticas Q_1 y Q_2 son *equivalentes*, y escribiremos $Q_1 \sim Q_2$, si $Q_1 = Q_2 \circ \gamma$ para algún $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Necesariamente en este caso Q_1 y Q_2 tienen el mismo discriminante y representan los mismos enteros (ambas cosas constituyen un ejercicio sencillo). Denotaremos con $\mathcal{C}_d$ al conjunto cociente de las formas cuadráticas de discriminante d módulo la relación $\sim$.

Nos gustaría escoger en cada clase de equivalencia en $\mathcal{C}_d$ un representante que en cierto modo tenga los coeficientes más pequeños, de igual manera que preferimos referirnos a la clase $\overline{2}$ en vez de a la clase $\overline{2012}$ cuando trabajamos en $\mathbb{Z}/5\mathbb{Z}$.

Representantes reducidos: *En cada clase de equivalencia de $\mathcal{C}_d$ existe exactamente una forma cuadrática $Q(x,y) = ax^2+bxy+cy^2$ que cumple*

$$-a < b \le a < c \qquad \text{o} \qquad 0 \le b \le a = c. \tag{3.8}$$

De aquí se deduce con un poco de esfuerzo que el cardinal de $\mathcal{C}_d$, llamado *número de clases* por antonomasia y denotado con $h(d)$, es finito (ejercicio).

Ejemplo: Para calcular $h(-4)$, de $b^2-4ac=-4$ se sigue $b=2B$ y $B^2+1=ac$. Si $B \neq 0$ entonces $a=c$ es imposible y $2|B| \le a < c$ también lleva a contradicción, por tanto $b=0$ y $a=c=1$. Es decir, todas las formas de discriminante -4 son equivalentes a x^2+y^2 y $h(-4)=1$.

Veamos una consecuencia aritmética de ello. Dado un divisor n de m^2+1 consideremos $Q(x,y) = nx^2+2mxy+\frac{m^2+1}{n}y^2$ que cumple $Q(1,0)=n$ y tiene $d=-4$, entonces $x^2+y^2 \sim Q$ y se concluye que cualquier divisor de un cuadrado más uno se escribe como suma de dos cuadrados. Esta información y un conocimiento básico sobre residuos cuadráticos es suficiente para deducir (3.2).

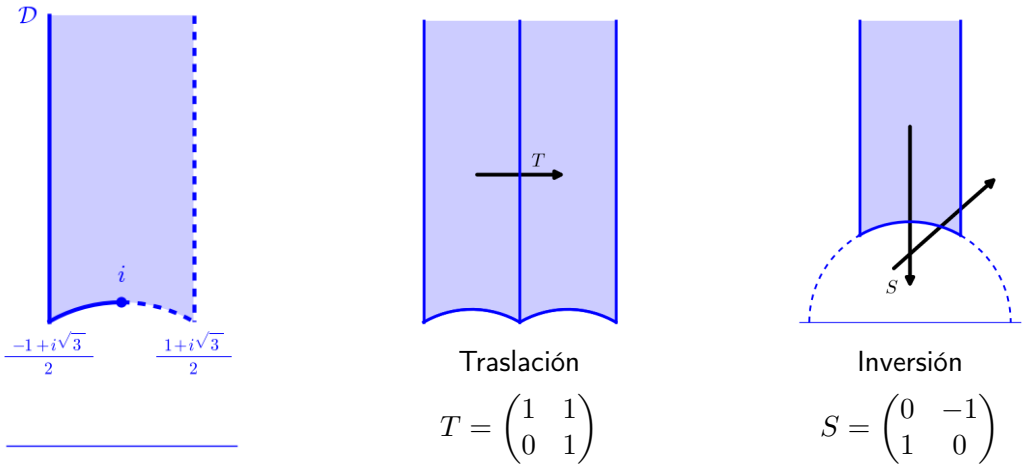
La existencia de representantes reducidos admite una prueba puramente algorítmica sin nuevos ingredientes pero hoy en día, debido a la relación con las formas modulares, es más atrayente relacionarla con la acción de $\mathrm{SL}_2(\mathbb{Z})$ sobre el semiplano superior $\mathbb{H} = \{\Im z > 0\}$ definida por

$$\gamma(z) := \frac{az+b}{cz+d} \qquad \text{para} \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \quad \text{y} \quad z \in \mathbb{H}. \tag{3.9}$$

Se cumple $\gamma\big(\tau(z)\big) = (\gamma\tau)(z)$. Es bien conocido que en cada órbita hay exactamente un elemento en el *dominio fundamental restringido*

$$\mathcal{D} := \{z \in \mathbb{H} : -\frac{1}{2} \le \Re z \le 0, \quad |z| \ge 1\} \cup \{z \in \mathbb{H} : 0 \le \Re z < \frac{1}{2}, \quad |z| > 1\}. \tag{3.10}$$

¿Cómo se prueba esto? Si $|z| < 1$ con $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ sacamos a z fuera de este (semi)círculo y si está fuera de la banda $-1/2 \le \Re z < 1/2$ con cierto $\left(\begin{smallmatrix} 1 & k \\ 0 & 1 \end{smallmatrix}\right)$ lo conseguimos meter dentro. Una iteración de estos procesos permite encontrar algorítmicamente [5, Ch.1] una colección $\{\gamma_j\}_{j=1}^{N}$ tal que $(\gamma_1\gamma_2\cdots\gamma_N)(z) \in \mathcal{D}$.



A cada forma Q le asignamos el $z \in \mathbb{H}$ tal que $Q(z, 1) = 0$. Recíprocamente cada $z \in \mathbb{H}$ en una extensión cuadrática proviene de una forma Q_z obtenida al homogeneizar el polinomio mínimo de z sobre $\mathbb{Z}$. Es fácil ver que $Q_z \circ \gamma = Q_{\gamma^{-1}(z)}$ (ejercicio), en particular $Q_z \sim Q_w$ si y sólo si $z = \gamma(w)$ con $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ y entonces, según lo dicho anteriormente, en cada clase hay exactamente una forma Q_z con $z \in \mathcal{D}$. Para $Q(x, y) = ax^2 + bxy + cy^2$ se tiene $z = (-b + i\sqrt{-d})/2a$ y la condición $z \in \mathcal{D}$ da justamente las condiciones (3.8).

Ejemplo: La forma $Q(x, y) = 27x^2 + 124xy + 143y^2$ debe ser equivalente a una que satisface (3.8). Para hallarlar, vemos que el z correspondiente a Q es $z = (-62+i\sqrt{17})/27$ y trasladado dos unidades con $\left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)$ da lugar a $z_1 = (-8 + i\sqrt{17})/27$ que cumple $-1/2 \le \Re z_1 < 0$ pero $|z_1| < 1$, por tanto pasamos a $z_2 = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)(z_1) = (8+i\sqrt{17})/3$. Finalmente la traslación $\left(\begin{smallmatrix} 1 & -3 \\ 0 & 1 \end{smallmatrix}\right)$ lleva a $z_3 = (-1 + i\sqrt{17})/3$ que ya pertenece a $\mathcal{D}$ y corresponde a la forma $3x^2 + 2xy + 6y^2$.

Una pregunta natural es si dos formas son equivalentes de manera única, es decir si $Q_2 = Q_1 \circ \gamma$ puede ser válido para más de un γ. Esto es lo mismo que decidir cuándo se cumple $Q \circ \gamma = Q$ o estudiar $\gamma(z) = z$ para $z \in \mathcal{D}$. En general esta última ecuación tiene sólo las dos

soluciones obvias $\gamma = \pm I$ pero cuando $z = i$ y $z = (-1+i\sqrt{3})/2$, que corresponden a las únicas formas (salvo equivalencias) de discriminantes $d = -4$ y $d = -3$, hay respectivamente otras dos y otras cuatro soluciones [49], [5, Ch.1]. En conclusión, si Q_1 y Q_2 son formas equivalentes de discriminante d entonces $Q_2 = Q_1 \circ \gamma$ para w_d matrices $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ donde

$$w_d = \begin{cases} 6 & \text{si } d = -3 \\ 4 & \text{si } d = -4 \\ 2 & \text{en otro caso.} \end{cases} \tag{3.11}$$

3.3. Número de representaciones por las clases

El número de representaciones de $m \in \mathbb{Z}^+$ por Q es el número de soluciones de $Q(x, y) = m$. Llamemos $r_d(m)$ a la suma de estos números de representaciones cuando Q recorre un sistema completo de representantes de las clases en $\mathcal{C}_d$. Curiosamente $r_d(m)$ admite una fórmula sencilla.

Restringiremos los posibles valores de m en aras de la claridad pero con un poco más de esfuerzo y razonamientos similares se consigue una fórmula muy general [28, §12.4].

Fórmula del número de representaciones: *Para m libre de cuadrados y primo con $2d$ se tiene*

$$r_d(m) = w_d \prod_{p|m} \left(1 + (\frac{d}{p})\right). \tag{3.12}$$

La prueba elabora el esquema que hemos apuntado para deducir el resultado de Fermat.

Consideremos (si existe) una solución $x = b$ de

$$x^2 \equiv d \pmod{4m} \qquad \text{con} \qquad 0 \le x < 2m. \tag{3.13}$$

Entonces

$$Q(x, y) = mx^2 + bxy + \frac{b^2 - d}{4m} y^2 \tag{3.14}$$

tiene discriminante d y existen w_d matrices $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ tales que $Q \circ \gamma = \widetilde{Q}$ es una de las formas del sistema completo de representantes. Por tanto $Q = \widetilde{Q} \circ \gamma^{-1}$ que aplicado a $(1, 0)$ prueba $m = \widetilde{Q}\big(\gamma^{-1}(1, 0)\big)$. No puede ser que otra solución $x = b_*$ de (3.13) dé lugar a la misma representación porque si $(1, 0)$ tiene la misma imagen por γ^{-1} y por γ_*^{-1} entonces $\gamma^{-1}\left(\begin{smallmatrix} 1 & k \\ 0 & 1 \end{smallmatrix}\right) = \gamma_*^{-1}$ para algún $k \in \mathbb{Z}$ y aplicando $\widetilde{Q}$ se tendría $\big(Q \circ \left(\begin{smallmatrix} 1 & k \\ 0 & 1 \end{smallmatrix}\right)\big)(x, y) = mx^2 + b_* xy + \ldots$ que lleva a $b_* = b + 2mk$ y contradice $0 \le b, b_* < 2m$ si $b_* \neq b$.

En definitiva, por cada solución de (3.13) hay w_d representaciones. Además el proceso se puede invertir asignando a cada representación de m una forma (3.14). Para $p|m$ la ecuación $x^2 \equiv d \pmod{p}$ tiene $1 + \left(\frac{d}{p}\right)$ soluciones y aplicando el teorema chino del resto con algún detalle

menor para incluir el factor 4 se tiene que el número de soluciones de (3.13) es justamente el producto del enunciado.

Si $h(d) = 1$ entonces la fórmula está contando el número de representaciones de una sola forma (cualquiera con determinante d). Esto ocurre para

$$d = -3,\ -4,\ -7,\ -8,\ -11,\ -12,\ -16,\ -19,\ -27,\ -28,\ -43,\ -67,\ -163. \tag{3.15}$$

Hay un teorema con una historia azarosa que prueba que no hay ningún caso más pero es difícil de probar.

Ejemplo: Determinemos los primos que son representados por $Q(x, y) = 9x^2 + 13xy + 5y^2$. Lo podemos hacer gracias a que Q tiene discriminante $d = -11$ el cual está en la lista (3.15). Es fácil ver que 2 no está representado y $11 = Q(-1, 2)$. En el resto de los casos la proposición implica que la condición necesaria y suficiente es $\left(\frac{-11}{p}\right) = 1$, que por la ley de reciprocidad cuadrática lleva a concluir que para cualquier primo p

$$p = 9x^2 + 13xy + 5y^2 \text{ con } x, y \in \mathbb{Z} \ \Leftrightarrow\ p \equiv 0, 1, 3, 4, 5, 9 \pmod{11}. \tag{3.16}$$

Así $23 = Q(-3, 2)$, $47 = Q(-3, 1)$, $37 = Q(-1, 4)$, $71 = Q(-6, 11)$ y $31 = Q(-2, 5)$.

3.4. Géneros

Tomando representaciones reducidas como en 3.8, tenemos que $\mathcal{C}_{-20}$ está compuesto por las clases de equivalencia de

$$x^2 + 5y^2 \qquad \text{y} \qquad 2x^2 + 2xy + 3y^2. \tag{3.17}$$

Los argumentos de la sección anterior sugieren que en lo que se refiere a representaciones ambas formas son inseparables, sin embargo (3.4) sugiere lo contrario.

La solución a esta paradoja es muy sencilla trabajando módulo 20. Unos cálculos implican que

$$\{x^2 + 5y^2 \ :\ x, y \in \mathbb{Z}/20\mathbb{Z}\} \cap \left(\mathbb{Z}/20\mathbb{Z}\right)^* = \{1, 9\}, \tag{3.18}$$

$$\{2x^2 + 2xy + 3y^2 \ :\ x, y \in \mathbb{Z}/20\mathbb{Z}\} \cap \left(\mathbb{Z}/20\mathbb{Z}\right)^* = \{3, 7\}. \tag{3.19}$$

Por consiguiente si $p \neq 2, 5$ es representable por $x^2 + 5y^2$ no puede serlo por $2x^2 + 2xy + 3y^2$ y viceversa, porque ambas formas representan diferentes clases de congruencias. La fórmula (3.12) asegura que $r_{-20}(p) = 4$ para $p \equiv 1, 3, 7, 9 \pmod{20}$ por consiguiente obtenemos el resultado (3.4) que no supo demostrar Euler y además lo complementamos con

$$2x^2 + 2xy + 3y^2 = p \text{ tiene solución } x, y \in \mathbb{Z} \ \Leftrightarrow\ p \equiv 3, 7 \pmod{20}. \tag{3.20}$$

Alentados por este éxito decimos que dos formas Q_1 y Q_2 de discriminante d están en el mismo *género* si representan los mismos valores de $\big(\mathbb{Z}/|d|\mathbb{Z}\big)^*$. Recordemos que $Q_1 \sim Q_2$ representan los mismos enteros, por tanto estar en el mismo género es una relación de equivalencia más débil que la equivalencia de formas en el sentido de que los géneros se dividen en clases.

Para los discriminantes de interés por su relación con la teoría de ideales (véase la siguiente sección) se puede probar que Q_1 y Q_2 están en el mismo género si y sólo si $Q_1 = Q_2 \circ \gamma$ para $\gamma \in \mathrm{GL}_2(\mathbb{Q})$, dando un paralelismo algebraicamente atractivo entre los géneros y la equivalencia de clases. Otras definiciones alternativas emplean los números p-ádicos [12, Th 3.21]. Originariamente Gauss determinó los géneros por medio de caracteres [6].

Recapitulando, siempre que cada género contenga una sola clase podemos separar las representaciones en clases de congruencia y decidir la representabilidad de los primos. Para $4 \mid d$ desde tiempos de Euler se sabe que esto ocurre para 65 discriminantes (Euler llamó a los valores de $d/4$ *números convenientes* [12, p.61]) y se ha probado que a lo más hay otro ejemplo.

Si hay más de una clase por género no queda más remedio que agrupar varias formas. Por ejemplo $\mathcal{C}_{-56}$ está compuesto por las clases de equivalencia de

$$x^2+14y^2, \qquad 2x^2+7y^2, \qquad 3x^2+2xy+5y^2 \quad \text{y} \quad 3x^2-2xy+5y^2. \tag{3.21}$$

Tras algunos cálculos se deduce que las dos primeras representan ambas $\overline{1}$, $\overline{9}$, $\overline{15}$, $\overline{23}$, $\overline{25}$ y $\overline{39}$ en $\big(\mathbb{Z}/56\mathbb{Z}\big)^*$ por tanto están el mismo género. Las otras dos formas constituyen otro género pues representan $\overline{3}$, $\overline{5}$, $\overline{13}$, $\overline{19}$, $\overline{27}$ y $\overline{45}$. La conclusión es entonces que para $p \neq 2, 7$

$$x^2+14y^2 = p \;\text{ ó }\; 2x^2+7y^2 = p \text{ tiene solución } x, y \in \mathbb{Z} \;\Leftrightarrow\; p \equiv 1, 9, 15, 23, 25, 39 \pmod{56}. \tag{3.22}$$

Escribiendo la primera ecuación como $2x^2 + 7(2x)^2 = 2p$ se obtiene (3.6).

La teoría de cuerpos de clases [12] permite "caracterizar" los primos que son representables por x^2+14y^2 pero la condición resultante es mucho más complicada que una simple congruencia. Concretamente, para $p \neq 2, 7$ es que $p \equiv 1, 9, 15, 23, 25, 39 \pmod{56}$ y que además la ecuación $x^4+2x^2-7 \equiv 0$ tenga sólución módulo p. Por ejemplo para $p = 23$, se tiene que $x = 3$ es solución y por tanto 23 es representable. Computacionalmente no está claro que sea más fácil resolver una ecuación de grado superior que comprobar directamente si un primo es representable dando valores.

3.5. Formas cuadráticas y teoría de ideales

Para motivar cuáles son los discriminantes "de interés" que hemos mencionado antes, pensemos en qué ocurriría si hubiéramos admitido formas cuadráticas que no tienen sus coeficientes coprimos (no primitivas) como $6x^2+3xy+9y^2$. Su discriminante es $-207 = -23 \cdot 3^2$ y el $z \in \mathbb{H}$

que le corresponde es $z = \big(-1 + i\sqrt{23}\big)/4$, el mismo que corresponde a $2x^2 + xy + 3y^2$ de discriminante -23. Es decir, admitiendo formas no primitivas perderíamos la biyectividad con los $z \in \mathbb{H}$ en extensiones cuadráticas y por otro lado, las formas no primitivas no añaden nada nuevo al problema de las representaciones.

En la teoría de ideales se estudian las formas cuadráticas a través de extensiones cuadráticas y queremos evitar cualquier discriminante que pudiera introducir ambigüedad con formas no primitivas. Así eliminando los discriminantes $d = -23k^2$ con $k > 1$ siempre que veamos $z = \big(\cdots + i\sqrt{23}\big)/\ldots$ estaremos seguros de que la forma correspondiente es de discriminante -23. No podemos decretar que el discriminante $-207 = -23 \cdot 3^2$ es realmente el importante y eliminar los otros que difieren en un factor cuadrático, porque por ejemplo $z = \big(1 + i\sqrt{207}\big)/4$ correspondería a $2x^2 - xy + 26y^2$ que tiene $d = -207$ pero $z = \big(-9 + i\sqrt{207}\big)/36$ llevaría a $2x^2 + xy + 3y^2$ con $d = -23$. Visto de otro modo, $z = (-9 + i\sqrt{207})/36$ corresponde a la forma $6x^2 + 3xy + 9y^2$ de discriminante $d = -207$, que no es primitiva y "simplificar" modifica el discriminante.

Se dice que d es un *discriminante fundamental* si todas las formas que se pueden construir de discriminante d son necesariamente primitivas. El párrafo anterior sugiere que los discriminantes fundamentales d son los libres de cuadrados pero si d es par la cosa es un poco más complicada. Exactamente,

$$(3.23)\quad d \text{ es discriminante fundamental } \Leftrightarrow \begin{cases} 2 \nmid d \text{ y } d \text{ es libre de cuadrados} \\ \text{o} \\ 4 \mid d,\ \dfrac{d}{4} \text{ es libre de cuadrados y } 4 \nmid (\dfrac{d}{4} - 1). \end{cases}$$

Nótese que si un discriminante es par necesariamente es divisible por 4.

Cualquier extensión cuadrática imaginaria $K/\mathbb{Q}$ se escribe de forma única como $K = \mathbb{Q}(\sqrt{d})$ con d un discriminante fundamental. El anillo de enteros de esta extensión es [6, Prop.6.6]

$$(3.24)\qquad \mathcal{O} = \mathbb{Z}\big[\frac{d - \sqrt{d}}{2}\big] = \big[1, \frac{d - \sqrt{d}}{2}\big]$$

donde $[\alpha, \beta]$ significa $\{n\alpha + m\beta \,:\, n, m \in \mathbb{Z}\}$. No hay que dejarse asustar por esta notación sintética. Por ejemplo para $d = -4$ tenemos $K = \mathbb{Q}(i)$ y $\mathcal{O} = \mathbb{Z}[i]$ mientras que para $d = -3$, $K = \mathbb{Q}(i\sqrt{3})$ y $\mathcal{O} = \mathbb{Z}[\big(1+i\sqrt{3}\big)/2]$. En general [10, III.7] para $d = 4k$, $K = \mathbb{Q}(\sqrt{k})$ y $\mathcal{O} = \mathbb{Z}[\sqrt{k}]$ mientras que para d impar $K = \mathbb{Q}(\sqrt{d})$ y $\mathcal{O} = \mathbb{Z}[\big(1 + \sqrt{d}\big)/2]$.

A cada par de enteros algebraicos $\alpha, \beta \in \mathcal{O} - \{0\}$ con $\alpha/\beta \notin \mathbb{Z}$ le asignamos una forma cuadrática de discriminante d mediante la regla

$$(3.25)\qquad \phi(\alpha, \beta) = \frac{(\alpha x + \beta y)(\bar{\alpha} x + \bar{\beta} y)}{n(\alpha, \beta)}$$

donde $n(\alpha,\beta)$ es el entero necesario para que la forma sea primitiva. Por ejemplo, para $\alpha = 2$ y $\beta = 1-\sqrt{-5}$, que corresponde a $d=-20$,

$$(\alpha x+\beta y)(\bar\alpha x+\bar\beta y) = 4x^2+4xy+6y^2 \implies \phi(2,1-\sqrt{-5}) = 2x^2+2xy+3y^2. \tag{3.26}$$

Un *ideal* $I\subset\mathcal{O}$ es un subanillo $I\neq\{0\}$ tal que $\alpha I\subset I$ para todo $\alpha\in\mathcal{O}$. Esta definición es muy poco intuitiva pero a la postre se prueba que para cualquier ideal I existen $\alpha,\beta\in\mathcal{O}-\{0\}$ y $\alpha/\beta\notin\mathbb{Z}$ tales que $I=[\alpha,\beta]$, por tanto geométricamente todo ideal en $\mathcal{O}$ es un retículo en $\mathbb{C}$ (pero no al revés). Evidentemente $[\alpha,\beta]=[\beta,\alpha]$ pero por razones que veremos en seguida ordenaremos α y β de manera que $\Im(\alpha/\beta)>0$. Ni con este convenio α y β están determinados unívocamente por I, por ejemplo $[\alpha,\beta]=[\alpha+\beta,\beta]=[\alpha+\beta,\alpha+2\beta]$. En general $[\alpha,\beta]=[a\alpha+b\beta,c\alpha+d\beta]$ para cualquier matriz entera $\gamma=\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)$ invertible sobre $\mathbb{Z}$. El convenio que hemos adoptado fuerza el signo positivo del determinante (se conserva la orientación) y por tanto $\gamma\in\mathrm{SL}_2(\mathbb{Z})$.

Las formas $Q_1=\phi(\alpha,\beta)$ y $Q_2=\phi(a\alpha+b\beta,c\alpha+d\beta)$ cumplen $Q_2=Q_1\circ\gamma^t$, entonces

$$\begin{aligned}\Phi:\{\text{Ideales en }\mathcal{O}\} &\longrightarrow \mathcal{C}_d\\ I=[\alpha,\beta] &\longmapsto \text{clase de } \phi(\alpha,\beta)\end{aligned} \tag{3.27}$$

es una aplicación bien definida entre los ideales y las clases de formas cuadráticas. Dicho sea de paso, $n(\alpha,\beta)$ es $|\mathcal{O}/I|$, lo que se llama *norma* del ideal I y se escribe $N(I)$, en particular es independiente de la elección de α y β para un ideal fijado. La aplicación Φ es sobreyectiva porque para cada $Q(x,y)=ax^2+bxy+cy^2$ de discriminante d el conjunto $I=\left[a,\frac{b-\sqrt{d}}{2}\right]$ es un ideal (ejercicio) y $\phi\big(a,(b-\sqrt{d})/2\big)=Q$.

Claramente Φ no es inyectiva ya que $\phi(\delta\alpha,\delta\beta)=\phi(\alpha,\beta)$ para cualquier $\delta\in\mathcal{O}-\{0\}$ y por consiguiente $\Phi(I)=\Phi(\delta I)$. Para eliminar esta ambigüedad definimos la relación de equivalencia entre ideales $I\sim J\Leftrightarrow\delta_1 I=\delta_2 J$ para ciertos $\delta_1,\delta_2\in\mathcal{O}-\{0\}$. Llamemos $\mathcal{I}_d$ al conjunto cociente resultante, entonces a partir de (3.27) conseguimos una biyeccción

$$\begin{aligned}\mathcal{I}_d &\longrightarrow \mathcal{C}_d\\ \text{clase de } \left[a,\frac{b-\sqrt{d}}{2}\right] &\longmapsto \text{clase de } ax^2+bxy+cy^2\end{aligned} \tag{3.28}$$

3.6. Grupo de clases y representaciones

Acabamos de ver que es lo mismo estudiar clases de ideales que clases de formas. ¿Y qué ganamos con ello? A fin de cuentas las formas cuadráticas son objetos naturales mientras que los ideales son poco intuitivos. Buscamos lo que con la mayoría de los "es lo mismo" en Matemáticas, conectar dos áreas y compartir sus técnicas.

En nuestro caso los ideales tienen una estructura algebraica más rica que las formas. Por ejemplo, al ser subanillos cerrados por multiplicaciones en $\mathcal{O}$ los ideales se pueden multiplicar entre sí definiendo $I \cdot J$ como el menor subanillo de $\mathcal{O}$ que contiene a todos los productos de elementos de I y J. Una propiedad importante es que esta multiplicación de ideales se comporta bien con respecto a la norma cumpliéndose [28, §16.9.4]

$$N(I \cdot J) = N(I)N(J). \tag{3.29}$$

No entraremos en cómo calcular explícitamente el producto de ideales en general. Nótese, no obstante, que no es difícil comprobar que el producto del ideal $\left[a, \frac{b-\sqrt{d}}{2}\right]$ por el ideal formado por los conjugados de sus elementos $\left[a, \frac{-b+\sqrt{d}}{2}\right]$ es $a\mathcal{O}$. Como $a\mathcal{O} \sim \mathcal{O}$ y $\mathcal{O}$ funciona como elemento neutro del producto, en $\mathcal{I}_d$ podemos invertir clases de ideales y $(\mathcal{I}_d, \cdot)$ adquiere estructura de grupo abeliano. Gracias a (3.28) también $\mathcal{C}_d$ adquiere estructura de grupo bajo la cual se conoce como *grupo de clases*.

Gauss definió directamente en $\mathcal{C}_d$ el producto de clases de formas, al que llamó *composición*, con unas fórmulas muy complicadas que en cierto modo incluyen todas las posibles variantes de (3.3) y (3.5). Entonces no existía todavía la teoría de ideales y tampoco existía la teoría de grupos como tal. De hecho la primera demostración del teorema de estructura de los grupos abelianos finitos está implícita en su estudio de las formas cuadráticas.

Saber multiplicar clases de formas en $\mathcal{C}_d$ sirve para algo más que satisfacer nuestro orgullo algebraico. La estructura de grupo de $\mathcal{C}_d$ contiene información acerca de cómo se construyen las representaciones de un número m a partir de sus factores. También los géneros admiten una explicación como cogrupos en $\mathcal{C}_d$ (en particular, siempre agrupan al mismo número de clases) con propiedades muy especiales, pero aquí no entraremos en ello.

Sea $\{Q_j\}_{j=1}^{h(d)}$ es un sistema completo de representaciones de $\mathcal{C}_d$ y sean los ideales correspondientes $I_j = \left[a_j, \frac{b_j-\sqrt{d}}{2}\right]$. Tomemos un m libre de cuadrados y primo con $2d$, como en la fórumla del número de representaciones (3.12). Si p es un factor primo de m con $\left(\frac{d}{p}\right) = -1$ sabemos que m no es representable por ningún Q_j y en otro caso tenemos $Q_j(x_p, y_p) = p$ para cierto $j = j(p)$. Entonces

$$m = \prod_{p|m} Q_{j(p)}(x_p, y_p) = \prod \frac{z_p \bar{z}_p}{N\big(I_{j(p)}\big)} \qquad \text{con} \qquad z_p = a_{j(p)}x_p + \frac{b_{j(p)} - \sqrt{d}}{2} y_p. \tag{3.30}$$

Consideremos $I = \prod I_{j(p)} = [\alpha, \beta]$ y $w = \prod z_p$. Como $w \in I$ se tiene $w = \alpha x + \beta y$ para ciertos $x, y \in \mathbb{Z}$ y de ello $w\bar{w}/N(I) = Q(x, y)$. Hemos probado entonces que Q representa a m cuando la clase de Q se puede escribir como producto de clases de formas cuadráticas que

representan a cada uno de los factores. Se puede probar que es posible invertir el proceso y módulo w_d simetrías las soluciones de $Q(x,y)=m$ están en biyección con las representaciones de los factores primos de m que sean coherentes con la condición de producto.

Sabemos que $h(-20)=2$ entonces $\mathcal{C}_{-20}$ es un grupo de dos elementos. Ya habíamos señalado los representantes $Q_1(x,y)=x^2+5y^2$ y $Q_2(x,y)=2x^2+2xy+3y^2$. El primero corresponde a $[1,-\sqrt{-5}]=\mathcal{O}$ y por tanto su clase es el elemento neutro. Sin apelar a los géneros deducimos de lo anterior que $m=p_1p_2\cdots p_r$ con $\left(\frac{-20}{p_j}\right)=1$ es representable por Q_1 si y sólo si el número de los p_j representables por Q_2 es par (así el producto de clases será el elemento neutro) y es representable por Q_2 en caso contrario.

La estructura de $\mathcal{C}_d$ nos ayuda en ocasiones a tratar problemas que parecen sin esperanza pensando sólo en términos de géneros.

Ejemplo: Veamos que para p primo

$$x^2+14y^2=3p \text{ tiene solución } x,y\in\mathbb{Z} \;\Leftrightarrow\; p\equiv 3,5,13,19,27,45 \pmod{56}. \tag{3.31}$$

Los casos $p=2,7$ los excluimos porque se pueden tratar directamente. De un ejemplo anterior sabemos que $Q_1(x,y)=x^2+14y^2$ y $Q_2(x,y)=2x^2+7y^2$ están en el mismo género, entonces aparentemente no se pueden separar las representaciones y las clases de congruencia allí indicadas sólo dan la implicación directa. Por ese mismo ejemplo sabemos que alguna de las formas $3x^2\pm 2xy+5y^2$ representa a p bajo las condiciones de congruencia del enunciado. De hecho ambas lo representan por la simetría $x\leftrightarrow -x$. Las clases de ideales que corresponden a estas dos formas tiene como producto la clase de $\mathcal{O}=[1,\sqrt{-14}]$ que corresponde a x^2+14y^2, por tanto $3p$ es representable por x^2+14y^2.

En realidad la condición de que m sea libre de cuadrados que hemos ido arrastrando hasta ahora no es muy relevante. Además los métodos anteriores permiten contar las soluciones.

Ejemplo: El grupo $\mathcal{C}_{-47}$ es un grupo cíclico de 5 elementos (véase por ejemplo la tabla de $[\overline{10}]$), $\mathcal{C}_{-47}=\langle g\rangle$ donde g^j es la clase de Q_j con

$$\begin{aligned} &Q_0(x,y)=x^2+xy+12y^2, \quad Q_1(x,y)=2x^2+xy+6y^2, \quad Q_2(x,y)=3x^2-xy+4y^2,\\ &Q_3(x,y)=3x^2+xy+4y^2, \quad Q_4(x,y)=2x^2-xy+6y^2. \end{aligned} \tag{3.32}$$

Con esta información, hallemos para cada $n,m\in\mathbb{Z}^+$ el número de soluciones de $Q_1(x,y)=17^n\cdot 83^m$.

Los primos 17 y 83 satisfacen $\left(\frac{-47}{p}\right)=1$ y, salvo las w_d simetrías, cada uno admite dos representaciones por la fórmula (3.12), las cuales son

$$17=Q_2(1,2)=Q_3(1,-2) \qquad \text{y} \qquad 83=Q_0(5,2)=Q_0(7,-2). \tag{3.33}$$

Con la notación anterior tenemos dos elecciones para z_{17}, digamos $z_{17}^{\pm}$ correspondientes a $g^{\pm 2}$ (nótese que $g^{-2} = g^3$). De la misma forma tenemos $z_{83}^{\pm}$ ambos correspondientes al elemento neutro. Deseamos que el resultado de multiplicar n de estos $z_{17}^{\pm}$ por m de los $z_{83}^{\pm}$ esté en el ideal que corresponde a g^1, la clase de Q_1. Es decir, debemos hacer una elección de signos tales que

$$\left(g^{\pm 2} \cdot g^{\pm 2} \cdot \overset{n \text{ veces}}{\cdots\cdots} \cdot g^{\pm 2}\right) \cdot \left(g^{\pm 0} \cdot g^{\pm 0} \cdot \overset{m \text{ veces}}{\cdots\cdots} \cdot g^{\pm 0}\right) = g^1. \tag{3.34}$$

Sea k y j el número de signos $+$ en el primer y segundo paréntesis, respectivamente. Debemos entonces contar el número de soluciones de

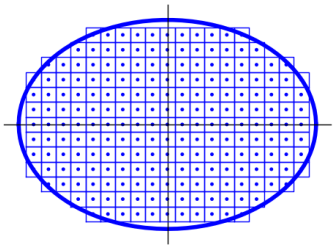
$$\begin{cases} 2k - 2(n-k) + j \cdot 0 - 0 \cdot (m-j) \equiv 1 \pmod 5 \\ 0 \le k \le n, \qquad 0 \le j \le m. \end{cases} \tag{3.35}$$

Hay $m+1$ posibilidades para j y escribiendo la congruencia como $-k - 2n \equiv 1 \pmod 5$ hay $\big[(3n+1)/5\big] - \big[2n/5\big]$ posibilidades para k, con $[\,\cdot\,]$ denotando la parte entera. Incorporando las $w_d = 2$ simetrías de cada solución se tienen en total

$$2(m+1)\left(\big[\frac{3n+1}{5}\big] - \big[\frac{2n}{5}\big]\right) \text{ soluciones.} \tag{3.36}$$

3.7. Indicaciones sobre la fórmula del número de clases

A cada forma Q y cada $N \in \mathbb{Z}^+$ le asignamos la elipse $Q(x,y) \le N$, la cual tiene área $2\pi N/\sqrt{|d|}$ con d el discriminante (ejercicio). Esta cantidad aproxima el número de puntos de coordenadas enteras dentro de ella cuando N es grande.



Elipse $Q(x,y) \le N$

Área $= 2\pi N/\sqrt{|d|}$
Número de puntos $\sim$ Área

Si tomamos representantes de todas las clases obtendremos

$$\lim_{N\to\infty} \frac{1}{N} \sum_{m=1}^{N} r_d(m) = \frac{2\pi h(d)}{\sqrt{|d|}}. \tag{3.37}$$

Por otra parte la multiplicatividad del símbolo de Jacobi permite escribir (3.12) como

$$r_d(m) = w_d \sum_{n|m} \chi_d(n) \qquad \text{donde} \qquad \chi_d(n) = \big(\frac{d}{n}\big). \tag{3.38}$$

Si esta fórmula fuera válida para todos los valores de m, no sólo para los de las hipótesis de (3.12), entonces intercambiando el orden de sumación se tendría

$$\frac{1}{N}\sum_{m=1}^{N} r_d(m) = \frac{w_d}{N}\sum_{n=1}^{N} \big[\frac{N}{n}\big]\chi_d(n). \tag{3.39}$$

Cuando N crece es de esperar que podamos quitar la parte entera en el límite y que el resultado converja a $L(1,\chi_d)$. Comparando con (3.37) obtendríamos la *fórmula del número de clases*

$$h(d) = \frac{w_d\sqrt{|d|}}{2\pi}L(1,\chi_d) \qquad \text{con} \qquad L(1,\chi_d) = \sum_{n=1}^{\infty}\frac{\chi_d(n)}{n}. \tag{3.40}$$

Este resultado es cierto para discriminantes fundamentales pero la prueba anterior no es correcta literalmente, sólo en su esquema, porque de hecho (3.38) no es válida para todos los valores de m.

Una de las consecuencias inmediatas de (3.40) es que $L(1,\chi_d) > 0$. Recuérdese que ello era esencial en el teorema de los números primos en progresiones aritméticas, por ello la prueba originaria empleaba la teoría de formas cuadráticas.

Sorprendentemente $L(1,\chi_d)$ se puede calcular explícitamente usando las sumas de Gauss y sustituyendo en (3.40) se sigue

$$h(d) = \frac{w_d}{2d}\sum_{n=1}^{-d} n\big(\frac{d}{n}\big). \tag{3.41}$$

Por ejemplo, para $d=-4$

$$h(-4) = \frac{4}{-8}\big(1\cdot 1 + 2\cdot 0 + 3\cdot(-1) + 4\cdot 0\big) = 1. \tag{3.42}$$

Los detalles de la prueba de (3.40) se pueden consultar en [28] o en el capítulo 6 de [13]. En [10] hay una extensión cuando d no es discriminante fundamental.

Capítulo 4

El método de van der Corput

4.1. Integrales oscilatorias

Comencemos con un poco de notación física. Representaremos una onda mediante una función $y(x) = g(x)e(f(x))$ donde $e(t) = e^{2\pi it}$ y f y g son funciones reales[1] que supondremos C^∞.

La función $g(x)$, o más precisamente $|g(x)|$, es la *amplitud* que controla el tamaño de la onda mientras que $f(x)$ es la *fase*, cuya derivada es la *frecuencia*.

Para un tono puro $y(x) = A\,e(mx)$, $m \in \mathbb{Z}^+$, los senos y los cosenos de siempre, la frecuencia m indica el número de oscilaciones (también llamadas *ciclos*) cuando x varía en una unidad. Claramente en $\Re y$ y en $\Im y$ cada parte positiva se compensa con una negativa al integrar y si la amplitud varía suavemente tal cancelación se conservará en gran medida, sobre todo para frecuencias altas. Por ejemplo, si $g \in C_0^\infty$ (a veces se dice que en este caso la onda es un *pulso* o un *paquete de ondas*) se tiene

$$\left| \int_{-\infty}^{\infty} g(x)e(mx)\,dx \right| \leq \frac{C_{g,k}}{m^k} \qquad \forall m, k \in \mathbb{Z}^+ \tag{4.1}$$

donde $C_{g,k}$ significa que la constante que hay que poner depende de g y del k que deseemos elegir. La prueba es muy sencilla, simplemente integrar por partes k veces. La condición de que m sea entero es irrelevante, basta $m \in \mathbb{R}^+$ y el caso $m \in \mathbb{R}^-$ es similar conjugando y escribiendo $|m|^k$ en el segundo miembro.

¿Qué ocurriría si integramos en un intervalo finito $[a, b]$ tal que g no se anula en los extremos? Entonces ya en la primera integración por partes aparecen términos de frontera que contribuyen $O(m^{-1})$ y por ello en general no es posible mejorar la cota (4.1) para $k = 1$. Demos a esta

[1] En principio podría parecer que los valores complejos de $y = y(x)$ están reñidos con cualquier significado físico y que esta función es un artificio matemático para considerar pares de ondas "de verdad" dadas por sus partes real e imaginaria, sin embargo la ecuación de Schrödinger conduce naturalmente a considerar ondas complejas.

observación un aspecto más general tomando f con $f' \geq m > 0$ y g/f' monótona, entonces

$$\left| \int_a^b g(x) e\big(f(x)\big)\, dx \right| \leq \frac{C_g}{m}. \tag{4.2}$$

La prueba se reduce a integrar por partes con $dv = 2\pi i f'(x) e\big(f(x)\big)$ y emplear $\big|e\big(f(x)\big)\big| = 1$ o, si se prefiere, a hacer primero el cambio de variable $f(x) = mt$ y proceder como antes.

En el caso de amplitud constante se tiene que si f' es monótona en $[a, b]$ y cumple $f' \geq m > 0$ entonces

$$\left| \int_a^b e\big(f(x)\big)\, dx \right| \leq \frac{C}{m} \tag{4.3}$$

donde C es una constante absoluta, por ejemplo $C = 3/2\pi$ es válida. Se suele llamar a este resultado *primer lema de van der Corput*. De nuevo el caso $f' \leq m < 0$ es similar escribiendo $|m|$ en el segundo miembro.

Las acotaciones (4.2) y (4.3) degeneran cuando $m \to 0$ lo cual es natural porque si una onda no oscila no se espera cancelación ($f = \text{cte}$, $b = -a \to \infty \Rightarrow \int_a^b e(f) \to \infty$). Sin embargo si la frecuencia es cero sólo en un punto, o dicho de otra forma, si la fase es estacionaria para dicho punto, entonces todavía cabe esperar cancelación. El caso prototípico es $f(x) = \lambda x^2$, $\lambda > 0$, con $x = 0$ como única solución de $f'(x) = 0$. Si $|x| > K/\lambda$ entonces obtenemos una acotación $O\big(K^{-1}\big)$ por (4.3) mientras que si $|x| \leq K/\lambda$ la cota trivial es $O(K/\lambda)$. Para minimizar la sumas de estas acotaciones elegimos $K = \sqrt{\lambda}$. El resultado es que si $g(x)/x$ es monótona en cada uno de los intervalos que componen $|x| \geq \sqrt{\lambda}$ y $0 \neq |x| \leq \sqrt{\lambda}$ entonces

$$\left| \int_a^b g(x) e\big(\lambda x^2\big)\, dx \right| \leq \frac{C_g}{\sqrt{\lambda}}. \tag{4.4}$$

El *principio de fase estacionaria* dice algo más preciso cuando g es una función C_0^∞ sin hipótesis adicionales:

$$\int_{-\infty}^{\infty} g(x) e(\lambda x^2)\, dx = \frac{1+i}{2\sqrt{\lambda}} g(0) + O\big(\lambda^{-3/2}\big). \tag{4.5}$$

La demostración consiste en utilizar la fórmula de Parseval $\int gf = \int \widehat{g}\widehat{f}$ con $f(x) = e(\lambda x^2)$ y sustituir la transformada $\widehat{f}(\xi)$ por una aproximación suya de Taylor de orden cero. Empleando Taylor hasta orden n, se obtienen resultados mejores con precisión $O\big(\lambda^{-n-3/2}\big)$.

Si repasamos el argumento que ha llevado a (4.4) en el caso $g = 1$ nos percataremos de que realmente la función λx^2 no tiene nada de particular, todo lo que necesitamos es una cota para la derivada en las cercanías del punto crítico. Si reemplazamos λx^2 por $f(x)$ con $x = 0$ como valor crítico, de $f'(x) = f'(x) - f'(0) = x f''(\xi)$ se sigue que basta acotar la derivada segunda inferiormente. Evidentemente que el punto crítico se alcance en 0 o en otro lugar es

indiferente tras una traslación. Todavía más, si no existiera un punto crítico, más a nuestro favor. Pensando con cuidado estas afirmaciones se prueba el *segundo lema de van der Corput*: Si $|f''| > \lambda$ en el intervalo $[a, b]$ entonces

$$\left| \int_a^b e\big(f(x)\big)\, dx \right| \leq \frac{C}{\sqrt{\lambda}} \tag{4.6}$$

para cierta constante C. El valor absoluto en f'' es sólo para enunciar simultáneamente el caso positivo y el negativo.

La motivación de J. van der Corput al desarrollar sus técnicas eran problemas de teoría analítica de números, pero la acotación y estimación de integrales oscilatorias tiene un ámbito mayor y los primeros ejemplos vinieron de la física matemática. Por ejemplo, F. Bessel (famoso por sus trabajos y por su correspondencia con Gauss) introdujo unas importantes integrales oscilatorias, hoy llamadas *funciones de Bessel*, para estudiar el movimiento de los planetas. Una de las más básicas es

$$J_1(\lambda) = \int_0^1 \cos(\pi x) \operatorname{sen}\big(\lambda \cos(\pi x)\big) dx. \tag{4.7}$$

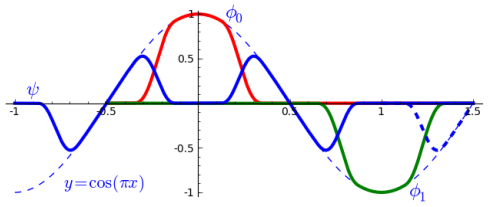
Para ilustrar las ideas anteriores, vamos a estudiar la asintótica de $J_1(2\pi\lambda)$ cuando $\lambda \to +\infty$, lo cual es importante en algunas situaciones.

Por la periodicidad,

$$J_1(2\pi\lambda) = \frac{1}{2} \int_0^1 \cos(\pi x) \operatorname{sen}\big(\lambda \cos(\pi x)\big) dx = \frac{1}{2} \Im \int_{-1/2}^{3/2} \cos(\pi x) e\big(\lambda \cos(\pi x)\big) dx. \tag{4.8}$$

La fase es $f(x) = \lambda \cos(\pi x)$ que tiene puntos estacionarios en $x = 0$ y en $x = 1$. Primero los aislamos para poder usar (4.5) cerca de ellos. Con esta fin introducimos $\phi_0 \in C_0^\infty$ tal que $\phi_0\big|_{[-0.1,0.1]} = \cos(\pi x)$, con soporte incluido en $[-0.5, 0.5]$ y $\phi_1(x) = -\phi_0(1-x)$ (que coincide con $\cos(\pi x)$ cerca de $x = 1$). La diferencia $d(x) = \cos(\pi x) - \phi_0(x) - \phi_1(x)$ se transforma en una función C_0^∞ "moviendo" la parte con $x \in [1, 1.5]$ a $[-1, -0.5]$. Es decir, considerando

$$\psi(x) = \begin{cases} d(x) & \text{si } x \in [-1/2, 1] \\ d(x+2) & \text{si } x \in [-1, -1/2] \\ 0 & \text{en el resto} \end{cases} \tag{4.9}$$



Van der Corput llamó a funciones de este tipo "neutralizadores", pues neutralizan el efecto de los puntos conflicitivos, en nuestro caso los estacionarios $x = 0, 1$, que requieren un estudio aparte.

De esta forma se obtiene:

$$J_1(2\pi\lambda) = \frac{1}{2}\Im \int_{-\infty}^{\infty} \big(\phi_0(x) + \phi_1(x) + \psi(x)\big) e\big(f(x)\big) dx. \tag{4.10}$$

Como f' no se anula en el soporte de ψ, podemos escribir sin peligro $\int \psi(x) e\big(f(x)\big) dx = \int \big(\psi(x)/f'(x)\big) f'(x) e\big(f(x)\big) dx$ e integrar por partes como en el primer lema de van der Corput. Así se obtiene $O\big(\lambda^{-1}\big)$ o incluso $O\big(\lambda^{-N}\big)$ para cualquier N integrando N veces por partes. Por otro lado, con el cambio $\cos(\pi x) = 1 - t^2/2$ o equivalentemente $t = 2\,\mathrm{sen}(\pi x/2)$ y aplicando (4.5), se deduce

$$\int \phi_0(x) e\big(f(x)\big) dx = \frac{2e(\lambda)}{\pi} \int \frac{\phi_0(\frac{2}{\pi}\operatorname{arc\,sen}\frac{t}{2})}{\sqrt{4-t^2}} e\big(-\frac{\lambda}{2}t^2\big) dt = \frac{e(\lambda - 1/8)}{\pi\sqrt{\lambda}} + O\big(\lambda^{-3/2}\big). \tag{4.11}$$

Por la simetría o repitiendo el cálculo, $\int \phi_1(x) e\big(f(x)\big) dx$ da el mismo resultado y se concluye

$$J_1(2\pi\lambda) = \frac{1}{\pi\sqrt{\lambda}} \operatorname{sen}\big(2\pi\lambda - \frac{\pi}{4}\big) + O\big(\lambda^{-3/2}\big). \tag{4.12}$$

Por ejemplo, para $\lambda = 2011$, el término principal difiere de $J_1(2\pi\lambda)$ en menos de $1.5 \cdot 10^{-7}$.

4.2. La estimación básica de van der Corput

Nuestro objetivo es conseguir acotaciones no triviales para *sumas trigonométricas*

$$S = \sum_{a \le m \le b} e\big(f(m)\big). \tag{4.13}$$

Este tipo de problemas aparecen en teoría analítica de números en numerosas ocasiones tras utilizar el análisis de Fourier para escribir una función como superposición de ondas. Después de separar una amplitud no oscilatoria mediante sumación por partes, el estudio del promedio de la función en $[a, b] \cap \mathbb{Z}$ lleva a sumas como la anterior.

Hasta ahora hemos acotado integrales oscilatorias, ¿qué relación guardan con las sumas trigonométricas? La conexión entre ambos temas es la *fórmula de sumación de Poisson*, un arma fundamental que transforma cualquier suma suficientemente regular en una suma de integrales oscilatorias:

$$\sum_{m=-\infty}^{\infty} h(m) = \sum_{m=-\infty}^{\infty} \widehat{h}(m) \qquad \text{con} \quad \widehat{h}(\xi) = \int_{-\infty}^{\infty} h(x) e(-x\xi). \tag{4.14}$$

La prueba habitual de esta fórmula emplea la función 1-periódica $H(x) = \sum_{m=-\infty}^{\infty} h(x+m)$. El primer miembro de (4.14) coincide con $H(0)$ y el segundo con la evaluación en cero de su

serie de Fourier. Formalmente considerando $h(x)$ como $e\big(f(x)\big)$ multiplicada por la función característica de $[a,b]$ se llegaría a

$$\sum_{a\le m\le b} e\big(f(m)\big) = \sum_{m=-\infty}^{\infty} \int_a^b e\big(f(x)-mx\big)\,dx. \tag{4.15}$$

Sin embargo tanto el enunciado (4.14) como su prueba requieren cierta regularidad para asegurar la convergencia. La discontinuidad de nuestra elección de h lleva a complicaciones técnicas y de hecho (4.15) requiere una pequeña modificación si a o b son enteros. Nótese que tampoco estamos tan lejos de una situación tratable si f' es monótona y $M_1 < f' < M_2$ porque la integral del término m de (4.15) está acotada por $O\big(1/(m-M_2)\big)$ si $m > M_2$ y por $O\big(1/(M_1-m)\big)$ si $m < M_1$, gracias al primer lema de van der Corput (4.3). Entonces estamos al borde de la convergencia ($\sum (m+\delta)^{-\alpha}$ diverge para $\alpha = 1$ y converge para $\alpha > 1$) y algunas técnicas analíticas permiten ganar un poco en los términos $m \notin [M_1, M_2]$ consiguiendo la convergencia y una cota pequeña para su contribución.

El resultado preciso que se prueba es que si $f'' \neq 0$ en $[a,b]$ entonces

$$\sum_{a\le m\le b} e\big(f(m)\big) = \sum_{M_1\le m\le M_2} \int_a^b e\big(f(x)-mx\big)\,dx + O\big(\log(2+M_2-M_1)\big) \tag{4.16}$$

donde M_1 y M_2 son enteros tales que $M_1 < f' < M_2$ en $[a,b]$.

La condición $f'' \neq 0$ asegura la monotonía de f'. Respecto a las técnicas analíticas a las que nos hemos referido, se pueden emplear versiones truncadas de Poisson, especialmente la fórmula de Euler-Maclaurin, o se puede introducir una función meseta $\phi \in C_0^\infty$ que pase de 0 a 1 y viceversa entre dos enteros para que S coincida con $\sum_{m=-\infty}^{\infty} \phi(m) e\big(f(m)\big)$.

Gracias a (4.16) estimar sumas trigonométricas se vuelve tan fácil o difícil como estimar sumas de integrales oscilatorias. Empleando el segundo lema de van der Corput (4.6), la suma del segundo miembro es $O\big((M_2-M_1)\lambda^{-1/2}\big)$ cuando $|f''| > \lambda$. Para dar a esta acotación un aspecto más manejable, usamos el teorema del valor medio garantizando que se pueden escoger M_2 y M_1 de forma que $M_2 - M_1 \le |f''(\xi)|(b-a)+2$.

En definitiva, obtenemos que si f satisface $\lambda \le |f''| \le C\lambda$ entonces

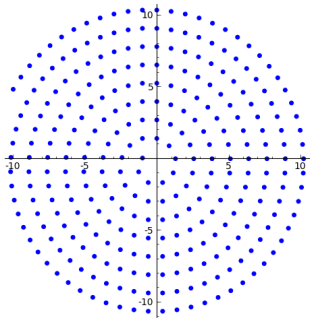
$$\sum_{a\le m\le b} e\big(f(m)\big) \ll (b-a+1)\lambda^{1/2} + \lambda^{-1/2} \tag{4.17}$$

donde la constante $\ll$ depende de C.

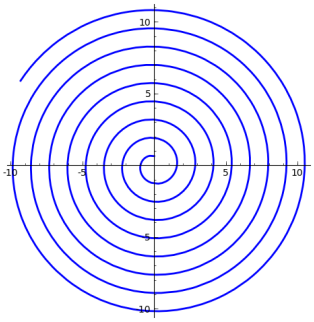
Ésta es la acotación más básica del método introducido por van der Corput y permite obtener acotaciones no triviales en rangos en los que la derivada segunda de la fase es moderadamente pequeña.

Por ejemplo, para la suma $\sum_{N\le m<2N} e(R/m)$ con $N \le R^{1/2}$, que aparece en el llamado problema del divisor [31], (4.17) da la acotación $O\big(R^{1/2}N^{-1/2}\big)$ que es no trivial si $N > R^{1/3}$.

Para funciones con derivada pequeña, (4.16) puede dar más información que (4.17) aproximando las integrales en vez de acotándolas, d hecho veremos más adelante que se puede hacer una teoría de esta idea. Aquí nos conformamos con mencionar un ejemplo curioso: si dibujamos los números complejos $z_N = \sum_{n=1}^{N} e\big(\frac{1}{2}\sqrt{n}\big)$, el resultado revela una estructura espiral inesperada.



z_N para $1 \le N \le 300$



$\frac{2}{\pi}\sqrt{x}\big(\sin(\pi\sqrt{x}), -\cos(\pi\sqrt{x})\big)$ para $1 \le x \le 300$

La explicación es que si aplicamos (4.16) con $f(x) = \frac{1}{2}\sqrt{x}$ y $M_1 = -1$, $M_2 = 1$, por los lemas de van der Corput $\int_1^N e\big(f(x) \pm x\big)dx = O(1)$ y por otro lado, $\int_1^N e\big(f(x)\big)dx$ se puede calcular explícitamente con el cambio de variable $x \mapsto x^2$. El resultado es

$$z_N = \frac{2\sqrt{N}}{\pi i} e\big(\frac{1}{2}\sqrt{n}\big) + O(1) = \frac{2\sqrt{N}}{\pi}\big(\sin(\pi\sqrt{N}) - i\cos(\pi\sqrt{N})\big) + O(1). \tag{4.18}$$

Por tanto los z_N recorren circunferencias con radio que crecen como $2\sqrt{N}/\pi$, lo que representa una espiral.

Se deja como reto para el lector explicar por qué el caso $z_N = \sum_{n=1}^{N} e(\sqrt{n})$ tiene un dibujo aparentemente tan distinto del anterior. En [38] hay otros ejemplos y gráficos.

4.3. La iteración del proceso de Weyl y van der Corput

Al aplicar (4.17) con $\lambda \gg 1$ no mejoramos la cota trivial. Teniendo en cuenta el ejemplo $\sum_{m=1}^{N} e\big(m^2\big) = N$ no hay nada raro en ello. Sin embargo este caso es muy especial y no parece totalmente ligado al crecimiento de f, por ejemplo, cabría esperar cancelación en $\sum_{m=1}^{N} e\big(m^{5/2}\big)$ a pesar de que, de nuevo, (4.17) no da ninguna información y $m^{5/2}$ crece más rápido que m^2.

En un famoso artículo sobre equidistribución, H. Weyl trató algunas sumas trigonométricas elevando su módulo al cuadrado varias veces para obtener incrementos de las fases a través de

$$\Big|\sum_{a\le m\le b} e\big(f(m)\big)\Big|^2 = \sum_{a\le m\le b}\sum_{a\le n\le b} e\big(f(n)-f(m)\big). \tag{4.19}$$

El problema es que cuando n y m son muy diferentes no hay mucho control sobre el incremento. La solución es agrupar los sumandos en bloques cortos y elevar al cuadrado cada bloque (empleando la desigualdad de Cauchy), con ello sólo aparecen incrementos de valores cercanos y podremos de alguna forma reducir el crecimiento de la fase. En la práctica, siguiendo a van der Corput, es importante no hacer una división en bloques disjuntos para evitar los efectos de los extremos. Lo que se hace es superponer los bloques unos sobre otros de forma que cada sumando aparezca cierto número H de veces, partiendo de la identidad

$$H\sum_{a\le m\le b} e\big(f(m)\big) = \sum_{m}\sum_{n\in I_m} e\big(f(n+m)\big) \qquad \text{con } I_m = [1,H]\cap[a-m,b-m]. \tag{4.20}$$

Tras aplicar la desigualdad de Cauchy, al desarrollar el cuadrado de la suma interior se obtienen fases $f(n+m)-f(n'+m)$ con $|n-n'|<H$. Separando los términos $n=n'$ y renombrando las variables con este argumento se prueba la desigualdad básica del *proceso de Weyl y van der Corput*

$$\Big|\sum_{a\le m\le b} e\big(f(m)\big)\Big|^2 \le \frac{4(b-a)^2}{H} + \frac{4(b-a)}{H}\sum_{1\le r<H}\Big|\sum_{a\le m\le b-r} e\big(f(m+r)-f(m)\big)\Big|. \tag{4.21}$$

Imaginemos que tenemos una fase tal que $f''\gg 1$ y por tanto (4.17) es trivial, la idea es que con (4.21) podemos pasar a considerar $f(x+r)-f(x)$ cuya derivada segunda es como $rf'''(\xi)$ por el teorema del valor medio. Si f''' fuera pequeña escogiendo un H adecuado podremos estar en buenas condiciones para aplicar (4.17). Si f''' no fuera pequeña aplicaríamos de nuevo (4.21) y tendríamos que considerar un incremento de derivadas terceras, esto es, una derivada cuarta.

En general con el control de alguna derivada de orden superior tendremos una acotación. El resultado concreto que se obtiene es que si f satisface $\lambda\le|f^{(k)}|\le C\lambda$ en $[a,b]$ para cierto $k\ge 2$, entonces

$$\sum_{a\le m\le b} e\big(f(m)\big) \ll (b-a+1)\lambda^{1/(2K-2)} + (b-a+1)^{1-2/K}\lambda^{-1/(2K-2)} \tag{4.22}$$

donde $K=2^{k-1}$ y la constante $\ll$ depende de C.

Escribir la prueba con detalle es un poco farragoso aunque se simplifica bastante utilizando inducción en k porque así sólo hay que aplicar (4.21) una vez. El paso inicial $k=2$ corresponde a (4.17).

Si consideramos de nuevo el ejemplo $\sum_{N\le m<2N} e(R/m)$ con $N \le R^{1/2}$, tomando $k = 3$ en (4.22) se llega a una cota no trivial para $N > R^{1/4}$, lo cual extiende el rango obtenido anteriormente pero la cota es peor para N cerca de $R^{1/2}$, que es lo que se necesita para el problema del divisor. Para el ejemplo $\sum_{m=1}^{N} e\big(m^{5/2}\big)$, intratable con (4.17), una división previa en intervalos diádicos $[N, N/2)$, $[N/2, N/4)$, ... del rango de sumación y la aplicación de (4.22) en cada uno de ellos con $k = 4$ demuestra que la suma es $O\big(N^{25/28}\big)$.

4.4. El método de pares de exponentes

Comencemos con unas consideraciones generales más intuitivas que rigurosas.

Si $|f'| < K$ con K no muy grande sabemos por (4.16) es esencialmente como unas pocas integrales que podríamos estimar o aproximar con métodos numéricos. La dificultad de la teoría está por tanto en el caso de las frecuencias altas. Supongamos entonces $|f'| > 1$. Bajo esta condición, $e\big(f(x)\big)$ oscila y si los valores de $e\big(f(m)\big)$ son aleatorios entonces el teorema central del límite sugiere que el tamaño típico de la suma debería ser como la raíz cuadrada del número de términos[2]. Por otro lado las oscilaciones grandes no siempre causan cancelación porque $f(m)$ podría caer por casualidad a menudo muy cerca de valores enteros.

En definitiva, las cotas para sumas trigonométricas dependen del tamaño de la frecuencia $D \asymp f'$ y del número de términos N. Digamos que buscamos acotaciones del tipo

$$S \ll D^p N^q \qquad \text{para} \quad S = \sum_{N\le m<2N} e\big(f(m)\big). \tag{4.23}$$

Siempre podemos dividir un rango de sumación en intervalos diádicos por tanto la forma de S no establece una gran restricción. En primera instancia diremos que (p, q) es un *par de exponentes* para la fase f si se cumple (4.23).

El objetivo es obtener unos pares de exponentes a partir de otros utilizando las ideas de las secciones anteriores y llegar a resultados no triviales a partir del par trivial $(0, 1)$. Al intentar poner en práctica esta idea se aparecen unas cuantas dificultades técnicas insidiosas cuando se trabaja con funciones generales, por ello en las exposiciones teóricas se restringe enormemente la clase de funciones que pueden aparecer en la fase. No definiremos esta clase aquí (véase [24] p. 30) pero sí indicaremos que tienen la propiedad de que $D \asymp N^{k-1} f^{k)}(x)$, lo cual es lógico porque el teorema del valor medio $f^{k-1)}(2N) - f^{k-1)}(N) = N f^{k)}(\xi)$ sugiere que en cada

[2]Lo que afirma el teorema es que al sumar variables aleatorias independientes de media cero y desviación típica σ entonces el resultado tiende a tener la distribución de una normal centrada de desviación típica $\sigma\sqrt{n}$. Por ejemplo si representamos salir cara con $+1$ y salir cruz con -1 ($\sigma = 1$) entonces al tirar una moneda n veces la probabilidad de que haya al menos m caras más que cruces está bien aproximada cuando n es grande por $\text{Prob}\{\xi > m,\ \xi \sim N(0, \sqrt{n})\} = (2\pi)^{-1/2} \int_{m/\sqrt{n}}^{\infty} e^{-x^2/2}\, dx$ que es exponencialmente pequeño si m es mucho mayor que $\sqrt{n}$.

derivada perdemos un N o, más informalmente, porque derivar es como bajar uno el grado. En la práctica sólo se habla de *pares de exponentes* para este conjunto restringido de funciones. Además se suele exigir $0 \le p \le 1/2 \le q \le 1$. Nótese que para $q > 1$ tendríamos acotaciones peores que la trivial y que las consideraciones probabilistas anteriores sugieren que $q < 1/2$ es imposible.

Consideremos el efecto de aplica el proceso de Weyl y van der Corput (4.21) a S como antes. Esquemáticamente se tiene

$$|S|^2 \ll N^2 H^{-1} + N H^{-1} |\widetilde{S}| \tag{4.24}$$

donde $\widetilde{S}$ son H sumas como S pero con frecuencias $f'(n+h) - f'(n) \ll H f''(\xi) \ll HDN^{-1}$. Si conocemos el par de exponentes (p, q) entonces $|\widetilde{S}| \ll H(HDN^{-1})^p N^q$. Sustituyendo en (4.24) y eligiendo $H = D^{-p/(p+1)} N^{1-q/(p+1)}$ para minimizar el resultado, se sigue $|S| \ll D^{p'} N^{q'}$ con $p' = p/(2p+2)$ y $q' = (p+q+1)/(2p+2)$. Es decir, hemos encontrado un proceso que aplica pares de exponentes en pares de exponentes.

Proceso A: *Si (p, q) es un par de exponentes entonces también es par de exponentes*

$$A(p, q) = \big(\frac{p}{2p+2}, \frac{p+q+1}{2p+2}\big). \tag{4.25}$$

Interesante pero todavía inútil si sólo contamos con el par de exponentes trivial $(0, 1)$ porque A lo fija.

Partamos ahora de la aplicación de (4.16) a S que da lugar a una suma de $O(D)$ integrales $\int_N^{2N} e\big(h(x)\big)\, dx$ con $h(x) = f(x) - mx$. Estas integrales una vez acotadas con el segundo lema de van der Corput (4.6) conducían a (4.17). Con vistas a mejorar el resultado tratamos de aproximarlas en lugar de acotarlas. Típicamente h alcanzará un punto crítico en $x = c \in (N, 2N)$. Digamos que sustituimos h por su aproximación de Taylor de orden dos alrededor de este punto, $\widetilde{h}(x) = h(c) + \frac{1}{2} h''(c)(x-c)^2$, entonces

$$\int_N^{2N} e\big(\widetilde{h}(x)\big)\, dx = e\big(f(c) - mc\big) \int_{N-c}^{2N-c} e\big(\lambda x^2\big)\, dx \qquad \text{con} \quad \lambda = \frac{1}{2} f''(c). \tag{4.26}$$

La última integral se puede aproximar por $\int_{-\infty}^{\infty} e\big(\lambda x^2\big)\, dx = (1 \pm i)/2\sqrt{|\lambda|}$, donde el $\pm$ es el signo de λ, y si todas estas aproximaciones funcionan bien[3] se tiene

$$\int_N^{2N} e\big(h(x)\big)\, dx = \frac{1 \pm i}{\sqrt{2|f''(c)|}} e\big(f(c) - mc\big) + \texttt{términos de error}. \tag{4.27}$$

[3]La prueba no es muy difícil ([38] §3 Th. 9) pero sí lo suficientemente larga y técnica como para evitar dar más detalles aquí.

Nótese que $f'(c) - m = 0$ y por tanto $f(c) - mc = F(m)$ con $F(x) = f\big(g(x)\big) - mg(x)$ y g la función inversa de f'. Además $f''(c) \asymp DN^{-1}$. En definitiva, sumando por partes para separar los coeficientes tenemos algo de la forma

$$|S| \ll (DN^{-1})^{-1/2}|\widetilde{S}| + \texttt{términos de error} \tag{4.28}$$

donde la fase de $\widetilde{S}$ es $F(m)$ y el intervalo de sumación tiene longitud $O(D)$. Un poco de Cálculo I prueba que $F' \asymp N$. Si conociéramos el par de exponentes (p, q) y nos olvidamos de los términos de error, entonces $|S| \ll (DN^{-1})^{-1/2}N^pD^q = D^{q-1/2}N^{p+1/2}$ y tenemos un nuevo proceso que preserva la propiedad de ser par de exponentes.

Proceso B: *Si (p, q) es un par de exponentes entonces también es par de exponentes*

$$B(p, q) = \big(q - \frac{1}{2}, p + \frac{1}{2}\big). \tag{4.29}$$

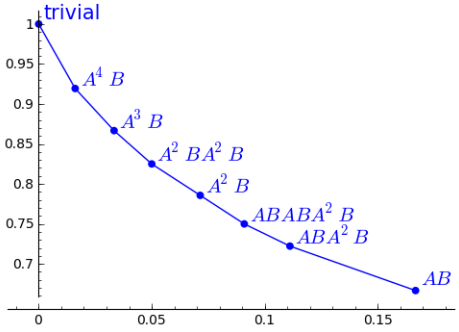
Aplicando B al par de exponentes trivial $(0, 1)$ llegamos a $(1/2, 1/2)$ que de hecho corresponde a (4.17) porque $\lambda_2 \asymp D/N$. Iteraciones sucesivas de A y B dan lugar a infinidad de pares de exponentes a partir de $(0, 1)$, llamados *pares de exponentes de van der Corput*.

Nótese que el proceso B es involutivo (esencialmente porque corresponde a la fórmula de sumación de Poisson) y no tiene sentido aplicarlo dos veces. S.W. Graham dio un algoritmo [24] para hallar una sucesión de composiciones de A y B que minimicen las acotaciones que aparecen típicamente en teoría analítica de números. Ya R.A. Rankin había probado anteriormente que el ínfimo de $p + q$ es $0{,}8290213568\ldots$ cuando se consideran pares de exponentes de van der Corput.

En la práctica se gana muy poco al tomar cadenas muy largas de A y B. Por ejemplo, la suma $\sum_{N\le m<2N} e(N^2 \log m)$, relacionada con la función ζ, se acota por $O\big(N^{5/6}\big)$ usando el par $AB(0, 1)$, la acotación se mejora a $O\big(N^{5/6-1/234}\big)$ empleando $ABA^3BA^2B(0, 1)$ y el algoritmo de Graham implica que la mejor elección sólo reduce el último exponente en menos de $4 \cdot 10^{-5}$.

Se muestran a continuación algunos pares con $0 \le p \le 1/2$. Aplicando el proceso B se extiende la lista a $1/6 \le p \le 1/2$.

$\mathsf{trivial} = (0, 1)$
$A^4B = (1/62, 57/62)$
$A^3B = (1/30, 13/15)$
$A^2BA^2B = (1/20, 33/40)$
$A^2B = (1/14, 11/14)$
$ABABA^2B = (1/11, 3/4)$
$ABA^2B = (1/9, 13/18)$
$AB = (1/6, 2/3)$



4.5. Algunos comentarios bibliográficos

Un libro breve y claro dedicado por entero al método de van der Corput en sentido amplio es [24]. Desafortunadamente la edición no está muy cuidada y hay bastantes erratas menores. En el tercer capítulo de [38] y en el segundo capítulo de [31][4] se incluyen buenas introducciones más breves y menos completas.

Hay algunas sumas que escapan de los métodos de van der Corput y que requieren otras técnicas, empleadas por ejemplo en el mejor término de error conocido para el teorema de los números primos. Algunas de las más notorias se debe a I.M. Vinogradov quien escribió un interesante e influyente libro [54]. Una de las mejores exposiciones del método de Vinogradov está en el capítulo 8 de [32] que también incluye todo el material básico del método de van der Corput.

Más centrado en las últimas novedades de sumas trigonométricas aplicables a problemas de puntos del retículo está [30] pero también incluye un gran número de ideas básicas y estimaciones simples en los capítulos iniciales. Si uno quiere evitar a toda costa el análisis, en [23] hay pruebas elementales de algunas de las aplicaciones típicas.

Hay también algunas sumas trigonométricas que requieren métodos de geometría algebraica. Gracias al desarrollo de las ideas de S.A. Stepanov es posible obtener demostraciones bastante simplificadas con técnicas que recuerdan a las de aproximación diofántica. Una exposición autocontenida que da todos los detalles para el caso de sumas de Kloosterman está en el capítulo 11 de [32]. Otros casos se tratan en [46].

[4] Comparando con otros textos, no parece claro que la teoría de pares de exponentes se aplique a una clase de funciones tan grande como se afirma en esta monografía.

Capítulo 5

Rudimentos sobre métodos de criba

5.1. Notación y argumentos probabilísticos

Supongamos que queremos decidir si un número es primo a partir de una tabla de primos bastantes más pequeños. Claramente si $p \nmid n$ para todo $p \leq \sqrt{n}$ entonces n es primo porque a lo más un factor suyo puede ser mayor que $\sqrt{n}$. Si sólo disponemos de una tabla de primos que no llega hasta $\sqrt{n}$, con este procedimiento no podemos decidir infaliblemente la primalidad de n pero si la tabla no es demasiado pequeña, estaremos bastante seguros de ella.

Los métodos de criba utilizan técnicas combinatorias y analíticas para estimar el número de elementos de un conjunto finito $\mathcal{A} \subset \mathbb{N}$ que no son divisible por primos p pequeños. La notación habitual es

$$S(\mathcal{A}, z) = \big|\{a \in \mathcal{A} \,:\, p \nmid a,\ \forall p < z\}\big| \tag{5.1}$$

con $|\cdot|$ representando el cardinal.

Para simplificar algunas fórmulas se suele introducir

$$P(z) = \prod_{p<z} p \tag{5.2}$$

y de esta forma

$$S(\mathcal{A}, z) = \big|\{a \in \mathcal{A} \,:\, (a, P(z)) = 1\}\big| \tag{5.3}$$

donde $(\cdot, \cdot)$ indica el máximo común divisor.

El objetivo básico es obtener $S(\mathcal{A}, z)$ eliminando de $\mathcal{A}$ los múltiplos de muchos números como en el procedimiento de Eratóstenes que aparece en los libros de aritmética elemental. Estos conjuntos a eliminar son

$$\mathcal{A}_d = \{a \in \mathcal{A} \,:\, d \mid a\} \tag{5.4}$$

donde d tiene factores primos menores que z.

Por último introducimos una función multiplicativa $0 \le g(d) < 1$ que aproxime $|\mathcal{A}_d|/|\mathcal{A}|$, la probabilidad de ser divisible por d. La multiplicatividad está reñida con que represente exactamente la probabilidad (enseguida volveremos sobre ello) y por ello admitimos un error r_d al aproximar $|\mathcal{A}_d|$ con esta probabilidad imperfecta. En símbolos

$$|\mathcal{A}_d| = |\mathcal{A}|g(d) + r_d. \tag{5.5}$$

Basta dar un vistazo a cualquier monografía sobre el tema (por ejemplo el clásico [25]) para darse cuenta de la complejidad de la notación que se suma a las dificultades intrínsecas de los métodos de criba. Aquí se ha introducido una notación esencial mínima y simplificada suficiente para entender las siguientes páginas pero que se queda corta si uno está interesado en otros temas o en una visión más general.

Para practicar con la notación y ahondar más en la interpretación probabilística consideremos $\mathcal{A} = [1, N]$ y $z = \sqrt{N}$. Según el argumento con el comenzamos, $S(\mathcal{A}, z)$ cuenta los primos entre $\sqrt{N}$ y N, el 1 y N si fuera un cuadrado de primo. Por tanto

$$S(\mathcal{A}, \sqrt{N}) = \pi(N) - \pi\big(\sqrt{N}\big) + O(1) \sim \frac{N}{\log N}. \tag{5.6}$$

Por otro lado, ¿cuál es la probabilidad de que un número no sea divisible por un primo p? Claramente $1 - 1/p$. ¿Y de que no sea divisible ni por p_1 ni por p_2, primos distintos? Sería $(1 - 1/p_1)(1 - 1/p_2)$. Considerando la probabilidad de que un elemento de $\mathcal{A}$ no sea divisible por ningún primo $p < \sqrt{N}$ uno estaría tentado a escribir

$$\frac{S(\mathcal{A}, \sqrt{N})}{N} \sim \prod_{p<\sqrt{N}} \big(1 - \frac{1}{p}\big) \qquad \text{que es FALSO.} \tag{5.7}$$

De hecho la *fórmula de Mertens* (cuya demostración es elemental pero no sencilla [11] §5.2) implica

$$\lim_{n\to\infty} \log n \prod_{p<\sqrt{n}} \big(1 - \frac{1}{p}\big) = 2e^{-\gamma} = 1.1229\ldots \tag{5.8}$$

donde $\gamma = 0.5772\ldots$ es la constante de Euler. El orden de magnitud cuadra con (5.6) pero no la constante. Por otro lado, cuando z es suficientemente pequeño entonces la idea probabilista funciona. Por ejemplo, más adelante veremos que sí es cierto

$$\frac{S(\mathcal{A}, z_0)}{N} \sim \prod_{p<z_0} \big(1 - \frac{1}{p}\big) \qquad \text{para } z_0 = \frac{\log N - 2\log\log N}{\log 2} \tag{5.9}$$

que, según (5.8), es asintóticamente $e^{-\gamma}/\log\log N$.

El fallo en (5.7) radica en la falta de independencia entre diferentes propiedades de divisibilidad[1]. Por ejemplo, en los números del 1 al 10 la probabilidad de ser par es 5/10 y la de ser múltiplo de 3 es 3/10, sin embargo la de ser múltiplo de 6 es 1/10 en lugar de $5/10 \cdot 3/10 = 15/100$. Si se remplaza 10 por un número N muy grande el error disminuye, por ejemplo, en los $n \leq 1000$ la probabilidad de ser par por la probabilidad de ser múltiplo de 3 es 0.1665 mientras que la probabilidad de ser múltiplo de 6 es 0.166. Esperamos entonces en este y en otros problemas

$$\frac{|\mathcal{A}_{p_1}|}{|\mathcal{A}|} \cdot \frac{|\mathcal{A}_{p_2}|}{|\mathcal{A}|} \approx \frac{|\mathcal{A}_{p_1 p_2}|}{|\mathcal{A}|} \qquad \text{para } p_1 \neq p_2 \tag{5.10}$$

con cierto grado de aproximación y algo similar con más factores primos, por ello suponemos que la probabilidad de ser divisible por d, la cantidad $|\mathcal{A}_d|/|\mathcal{A}|$, es aproximadamente una función multiplicativa $g(d)$.

Los métodos de criba consiguen a menudo en rangos adecuados estimaciones del tipo

$$S(\mathcal{A}, z) \asymp |\mathcal{A}| \prod_{p<z} \big(1 - g(p)\big), \tag{5.11}$$

probando así que la idea probabilista tiene al menos el orden correcto.

De la igualdad en (5.6) tenemos que el teorema de los números primos es equivalente a una fórmula asintótica para $S(\mathcal{A}, \sqrt{N})$ con $\mathcal{A} = [1, N]$. Es posible dar interpretaciones similares en problemas abiertos clásicos.

Por ejemplo, consideremos $\mathcal{A} = \{n(n+2) \,:\, n \in \mathbb{Z}^+\} \cap [1, N]$ entonces $S(\mathcal{A}, N^{1/4})$ cuenta n y $n+2$ cada uno de ellos menor que $\sqrt{N}$ que tienen factores primos mayores que $N^{1/4}$, por tanto n y $n+2$ deben ser primos y cualquier cota inferior $S(\mathcal{A}, N^{1/4}) > 0$ válida para todo N suficientemente grande probaría la conjetura de los primos gemelos.

Definimos la función que cuenta primos gemelos

$$\pi_2(x) = \{p \leq x \,:\, p \text{ y } p+2 \text{ son primos}\}. \tag{5.12}$$

Según la conjetura, todavía abierta, $\pi_2(x) \to \infty$. Veamos cómo la interpretación probabilística permite formular una conjetura más precisa. Definimos el suceso A consistente en que un número $n \leq x$ elegido al azar sea primo y B lo mismo con $n+2$. Con este lenguaje

$$\pi_2(x) \sim x\mathrm{Prob}(A \cap B) \qquad \text{y} \qquad \mathrm{Prob}(A) \sim \mathrm{Prob}(B) \sim \frac{1}{\log x}. \tag{5.13}$$

[1]Recuérdese que la independencia de dos sucesos A y B se traduce en que $\mathrm{Prob}(A \cap B) = \mathrm{Prob}(A)\mathrm{Prob}(B)$.

¿Se cumplirá $\text{Prob}(A \cap B) = \text{Prob}(A)\text{Prob}(B)$? No es lógico porque si sabemos que $n+2$ es primo, en particular no es divisible por ejemplo por 3 y por tanto es imposible que n sea un primo de la forma $3k-2$. Hay por tanto un condicionamiento, una falta de independencia, que hace más difícil que n sea primo si $n+2$ ya lo es. Escribimos entonces una fórmula con probabilidades condicionadas

$$\pi_2(x) \sim x\text{Prob}(A|B)\text{Prob}(B) \sim \frac{\text{Prob}(A|B)}{\text{Prob}(A)} \frac{x}{\log^2 x}. \tag{5.14}$$

Supongamos que podemos aproximar $\text{Prob}(A|B)/\text{Prob}(A)$ reemplazando A y B por la no divisibilidad por muchos primos y multiplicando los resultados como hicimos sin éxito en (5.7) y con éxito en (5.9). Si A_p y B_p son las probabilidades de que n y $n+2$ no sean divisibles por p, salvo un error que tiende a cero con x, se tiene $\text{Prob}(A_p) = 1 - 1/p$ mientras que $\text{Prob}(A_p|B_p) = (p-2)/(p-1)$ para $p > 2$ ya que $p \nmid n+2$ limita n a $p-1$ posibles clases de congruencias y una de ellas, $\overline{0}$, es imposible si además $p \nmid n$. Finalmente $\text{Prob}(A_2|B_2) = 1$ porque si $n+2$ es par, n también lo es. Conjeturamos entonces que se verifica

$$\pi_2(x) \sim \frac{2x}{\log^2 x} \prod_{p>2} \frac{p(p-2)}{(p-1)^2}. \tag{5.15}$$

El producto converge ya que sus factores son $1 - O\big(p^{-2}\big)$. ¿Por qué esperamos que las cosas no vayan mal como en (5.7)? Justamente por la convergencia. Por ella da igual asintóticamente limitar el producto hasta un logaritmo, lo cual funcionó en (5.9), o extenderlo hasta el infinito como hemos hecho.

Un argumento similar para la conjetura de Goldbach exigiendo que n y $N-n$ sean primos lleva a conjeturar que el número de representaciones $r_2(N)$ de un número par N como suma de dos primos verifica

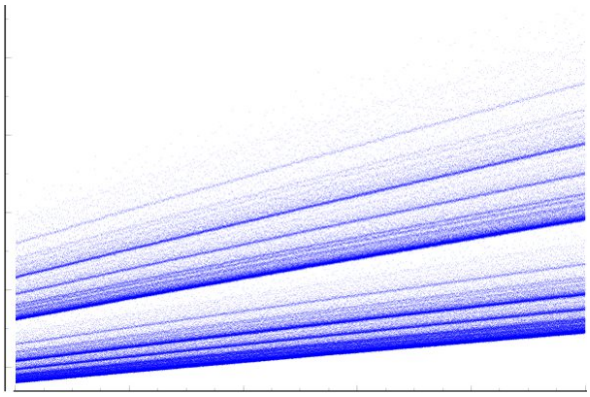
$$r_2(N) \sim \frac{x}{\log^2 x} \prod_{p|N} \frac{p}{p-1} \prod_{p\nmid N} \frac{p(p-2)}{(p-1)^2} = \frac{2x}{\log^2 x} \prod_{p>2} \Big(1 - \frac{1}{(p-1)^2}\Big) \prod_{\substack{p|N \\ p>2}} \Big(1 + \frac{1}{p-2}\Big). \tag{5.16}$$

En cada intervalo $[N, 2N]$ la función $\log^2 x$ no varía sustancialmente y por tanto la gráfica del segundo miembro de (5.16) será un conjunto de puntos que se agrupan aproximadamente en rectas, cuyas pendientes corresponden a los divisores primos.

Este comportamiento se refleja experimentalmente en la gráfica real de $r_2(N)$. En particular, los números pares N que sólo tienen factores primos grandes distintos de 2 tienen pocas representaciones en comparación con los que tienen muchos factores primos pequeños distintos. Por ejemplo, en el intervalo $[10^6, 2\cdot 10^6]$ el mínimo valor de $r_2(N)$ sobre los pares se alcanza en $N = 1002002 = 2 \cdot 501001$ y el máximo en $N = 1981980 = 2^2 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11^2 \cdot 13$.

Tanto (5.15) como (5.16) están respaldadas por extensos cálculos computacionales.

Con razonamientos análogos se llega a una fórmula conjetural para primos representados por polinomios, cuantificando así una hipótesis sobre su existencia atribuida a A. Schinzel [42].



$r_2(N)$ para $10^6 \leq N < 2 \cdot 10^6$

5.2. La criba de Eratóstenes-Legendre

Todos los métodos de criba parten de la identidad

$$S(\mathcal{A}, z) = \sum_{d|P(z)} \mu(d)|\mathcal{A}_d|. \tag{5.17}$$

La prueba más intuitiva consiste en utilizar simplemente el principio de inclusión-exclusión. Para quitar de $\mathcal{A}$ los múltiplos de primos menores que z nos deshacemos primero de los pares y quedan $|\mathcal{A}| - |\mathcal{A}_2|$. Al restar los múltiplos de 3, $|\mathcal{A}_3|$, habrá que compensar los múltiplos de 6, $|\mathcal{A}_2 \cap \mathcal{A}_3| = |\mathcal{A}_6|$, eliminados dos veces, entonces quedan $|\mathcal{A}| - |\mathcal{A}_2| - |\mathcal{A}_3| + |\mathcal{A}_6|$. El procedimiento continúa con todos los primos $p < z$ y según el principio de inclusión-exclusión el signo que corresponde a $|\mathcal{A}_{p_1} \cap \mathcal{A}_{p_2} \cap \ldots \cap \mathcal{A}_{p_k}| = |\mathcal{A}_{p_1 p_2 \ldots p_k}|$ es $(-1)^k$, o lo que es lo mismo $\mu(p_1 p_2 \ldots p_k)$. También es fácil dar una prueba directa con las propiedades básicas de la función de Möbius.

Empleando (5.5) y usando que la función g es multiplicativa tenemos la desigualdad a veces conocida como *criba de Eratóstenes-Legendre*

$$\left| S(\mathcal{A}, z) - |\mathcal{A}| \prod_{p<z} \big(1 - g(p)\big) \right| \leq \sum_{d|P(z)} |r_d|. \tag{5.18}$$

Estudiemos la fuerza de esta fórmula en el ejemplo $\mathcal{A} = [1, N]$. La probabilidad de ser divisible por d es como $1/d$ y por tanto elegimos $g(d) = 1/d$. El valor exacto de $|\mathcal{A}_d|$ es $[N/d]$ (donde $[\,\cdot\,]$ denota la parte entera) y por (5.5) tenemos que r_d es, salvo el signo, la parte fraccionaria de N/d. Así pues $|r_d| \leq 1$. El número de sumandos que hay en el segundo miembro

de (5.18) está acotado por 2^z ya que obviamente las posibles elecciones de los d corresponden a subconjuntos de primos y en particular a subconjuntos de $[1,z]\cap\mathbb{Z}$ (un conjunto de n elementos tiene 2^n subconjuntos). Se tiene entonces con estas acotaciones

$$S(\mathcal{A},z) = N\prod_{p<z}\big(1-\frac{1}{p}\big)+O\big(2^z\big). \tag{5.19}$$

Si sustituimos $z=\sqrt{N}$ obtenemos un resultado ridículo exponencialmente peor que la desigualdad trivial $S(\mathcal{A},z)\leq N$. Tenemos que bajar hasta $z=(\log N)/\log 2$ para tener la estimación trivial. Teniendo en cuenta que el producto es comparable a $(\log z)^{-1}$ según (5.8), con un z algo menor ya tenemos una fórmula asintótica. Por ejemplo para $z\leq(\log N-2\log\log N)/\log 2$

$$S(\mathcal{A},z)\sim\frac{e^{-\gamma}N}{\log z} \tag{5.20}$$

cuando z y N crecen indefinidamente, lo que implica (5.9).

5.3. Limitaciones

Si queremos usar los métodos de criba para detectar primos, es desalentador que sólo hayamos podido tomar en el apartado anterior un z tan pequeño cuando $\mathcal{A}=[1,N]$. Según vimos en (5.6) con $z=\sqrt{N}$ conseguimos primos y este valor es crítico. Con $S\big(\mathcal{A},N^{0.49}\big)$ ya estamos incluyendo en la cuenta elementos de $\mathcal{A}$ con dos factores primos mayores que $N^{0.49}$.

Selberg [47] probó mediante dos famosos ejemplos que hay un obstáculo teórico que impide llegar a la barrera $\sqrt{N}$ empleando sólo el tipo de información que manejan los métodos de criba. Este obstáculo radica en la imposibilidad de distinguir entre números con una cantidad par o impar de factores y se le conoce con el nombre de *fenómeno de la paridad*. Impide por ejemplo que se pueda probar el teorema de los números primos[2] o incluso $\pi(x)\gg x/\log x$ utilizando métodos de criba [27] §4. Por otro lado, en un revolucionario artículo J.B. Friedlander y H. Iwaniec mostraron en 1998 cómo superar el fenómeno de la paridad con una hipótesis adicional [21] pero los años transcurridos han dado escasas aplicaciones de este resultado.

Si nos resignamos a vivir con el fenómeno de la paridad los métodos de criba nunca llegarán a probar por sí solos la conjetura de los primos gemelos ni la conjetura de Goldbach sin embargo el ingenio de muchos autores los han conseguido llevar al borde de sus límites teóricos en estos y otros problemas. Por ejemplo, J.-R. Chen probó (véase [25] y §25.6 [20]) que (5.15) y (5.16) son ciertas reemplazando $\sim$ por $\asymp$ y $\pi_2(x)$ y $r_2(N)$ respectivamente por las funciones

$$\pi_2^*(x)=\{p\leq x\,:\,p \text{ es primo y } p+2 \text{ tiene a lo más dos factores primos}\} \tag{5.21}$$

[2]Curiosamente, Selberg obtuvo su demostración elemental del teorema de los números primos mientras trabajaba sobre métodos de criba.

y

$$r_2^*(N) = \{p < N : p \text{ es primo y } N - p \text{ tiene a lo más dos factores primos}\}. \tag{5.22}$$

Estos resultados requieren llevar z hasta cierta potencia de $|\mathcal{A}|$ (que es comparable a x y N) mientras que con (5.18) parece que estamos limitados a un logaritmo de $|\mathcal{A}|$. La razón última es que hay demasiados sumandos en el segundo miembro de (5.17), una cantidad exponencial en z.

Los métodos de *criba combinatoria* (como el de Brun que veremos después) buscan tachar casi todos los sumandos de (5.17) conservando desigualdades. Es decir, se buscan conjuntos no muy grandes $\mathcal{D}^-$ y $\mathcal{D}^+$ tales que

$$\sum_{\substack{d|P(z)\\ d\in\mathcal{D}^-}} \mu(d)|\mathcal{A}_d| \leq S(\mathcal{A}, z) \leq \sum_{\substack{d|P(z)\\ d\in\mathcal{D}^+}} \mu(d)|\mathcal{A}_d|. \tag{5.23}$$

A primera vista parece extraño que exista una manera eficiente de tachar términos y lo es más todavía que haya una profunda teoría acerca de ello.

Hay otros métodos no combinatorios (como el de Selberg) en los que se buscan en general funciones f^- y f^+ que sean pequeñas muchas veces y tales que

$$\sum_{d|P(z)} f^-(d)|\mathcal{A}_d| \leq S(\mathcal{A}, z) \leq \sum_{d|P(z)} f^+(d)|\mathcal{A}_d|. \tag{5.24}$$

De nuevo esta idea resulta chocante porque intuitivamente las funciones f^- y f^+ deberían parecerse a $\mu(d)$ si buscamos buenos resultados y por otra parte queremos que sean pequeñas en valor absoluto casi todo el tiempo.

En (5.23) y (5.24) no sólo hay que buscar la menor contribución de los términos de error sino también que sea posible interpretar el término principal al sustituir $|\mathcal{A}_d|$ por $|\mathcal{A}|g(d)$. En general el aspecto de estas fórmulas con desigualdades impide obtener fórmulas asintóticas excepto en casos límite. Además (5.6), (5.7) y (5.8) ya sugieren que hay una pérdida inherente de una constante.

5.4. La criba de Brun

V. Brun originó los métodos de criba modernos con diferentes trabajos a partir de 1917. Lo que hoy en día se suele llamar criba de Brun (o criba pura de Brun) es la versión más simple y menos poderosa de sus ideas con una notación actualizada.

Es una criba combinatoria en la que se elige $\mathcal{D}^- = \{d : \nu(d) < 2l\}$ y $\mathcal{D}^+ = \{d : \nu(d) < 2l+1\}$ donde $\nu(d)$ es el número de factores primos distintos y l un número natural arbitrario. No

es difícil dar una prueba directa de (5.23) con esta elección (véase [9] Lema 10.12) sin embargo aquí utilizaremos la llamada *identidad de Buchstab* que desempeña un papel destacado en algunas demostraciones de criba:

$$S(\mathcal{A},z)=|\mathcal{A}|-\sum_{p<z}S(\mathcal{A}_p,p). \tag{5.25}$$

La prueba es muy sencilla: cualquier elemento de $\mathcal{A}$ o bien está contado en $S(\mathcal{A},z)$ o bien es de la forma $p^\alpha m$ con $p<z$ y todos los factores primos de m mayores que p. Fijado p estos elementos $p^\alpha m$ son justamente los contados por $S(\mathcal{A}_p,p)$.

Iterando un par de veces (5.25) se tiene

$$\begin{aligned}S(\mathcal{A},z)&=|\mathcal{A}|-\sum_{p<z}|\mathcal{A}_p|+\sum_{q<p<z}S(\mathcal{A}_{pq},q)\\&=|\mathcal{A}|-\sum_{p<z}|\mathcal{A}_p|+\sum_{q<p<z}|\mathcal{A}_{pq}|-\sum_{r<q<p<z}S(\mathcal{A}_{pqr},r)\end{aligned}$$

donde p, q y r recorren los primos. Iterando $k-1$ veces se llega a

$$S(\mathcal{A},z)=\sum_{\substack{d|P(z)\\ \nu(d)<k}}\mu(d)|\mathcal{A}_d|+(-1)^k\sum_{\substack{d|P(z)\\ \nu(d)=k}}S(\mathcal{A}_d,p_d) \tag{5.26}$$

donde p_d indica el menor factor primo de d.

Con esta especie de principio de inclusión-exclusión truncado es evidente que se cumple (5.23) con la elección antes indicada de $\mathcal{D}^-$ y $\mathcal{D}^+$. El parámetro k en (5.26) permite controlar el número de términos y por tanto la acumulación de errores.

En la criba de Eratóstenes-Legendre el término principal aparecía directamente pero ahora requiere la identidad

$$\sum_{\substack{d|P(z)\\ \nu(d)<k}}\mu(d)g(d)=\prod_{p<z}\big(1-g(p)\big)-\sum_{\substack{d|P(z)\\ \nu(d)=k}}\mu(d)g(d)\prod_{q<p_d}\big(1-g(q)\big) \tag{5.27}$$

donde también aquí q recorre los primos. Lo único que hay que comprobar para obtener (5.27) es que

$$\sum_{\substack{d|P(z)\\ \nu(d)\geq k}}\mu(d)g(d)=\sum_{\substack{d|P(z)\\ \nu(d)=k}}\mu(d)g(d)\prod_{q<p_d}\big(1-g(q)\big). \tag{5.28}$$

El producto es $\sum_{d|P(p_d)}\mu(l)g(l)$ y al ser g multiplicativa, lo que indica (5.28) es que todo d en el primer miembro se escribe de forma única como sus k mayores factores primos por un número l con factores menores.

Si en (5.26) sustituimos $|\mathcal{A}_d| = |\mathcal{A}|g(d) + r_d$ y utilizamos (5.27) tenemos que la diferencia $S(\mathcal{A}, z) - |\mathcal{A}| \prod_{p<z} \big(1 - g(p)\big)$ es idénticamente igual a

$$\sum_{\substack{d|P(z)\\ \nu(d)=k}} \mu(d) \Big(S(\mathcal{A}_d, p_d) - |\mathcal{A}|g(d) \prod_{q<p_d} \big(1 - g(q)\big) \Big) + \sum_{\substack{d|P(z)\\ \nu(d)<k}} \mu(d) r_d \tag{5.29}$$

cualquiera que sea k. Con $0 \le S(\mathcal{A}_d, p_d) \le |\mathcal{A}_d| = |\mathcal{A}|g(d) + r_d$ y $0 \le g(q) < 1$ limpiamos el primer paréntesis y obtenemos la *criba de Brun* en la versión

$$\Big| S(\mathcal{A}, z) - |\mathcal{A}| \prod_{p<z} \big(1 - g(p)\big) \Big| \le |\mathcal{A}| \sum_{\substack{d|P(z)\\ \nu(d)=k}} g(d) + \sum_{\substack{d|P(z)\\ \nu(d)\le k}} |r_d|. \tag{5.30}$$

Para k grande esperamos que $g(d)$ en el primer sumatorio sea muy pequeño porque es el producto de k valores $0 \le g(p) < 1$ pero en el segundo sumatorio se produce una acumulación de términos de error. En la práctica se busca un balance entre ambos sumatorios.

Para comparar con la criba de Eratóstenes-Legendre elijamos $\mathcal{A} = [1, N]$ de modo que $|\mathcal{A}_d| = |\mathcal{A}|g(d) + r_d$ con $g(d) = 1/d$ y $|r_d| \le 1$. El primer sumatorio de (5.30) está acotado por $\big(\sum_{p<z} p^{-1}\big)^k/k!$ mientras que el segundo sumatorio está acotado por z^k. Sin entrar en detalles, empleando $a^k/k! \ll (ae/k)^k$ es posible probar (utilícese $\sum_{p<z} p^{-1} = \log\log z + O(1)$, [9] §2.8) que tomando $k = [4 \log\log N]$ ambos términos son $o(N/\log z)$ cuando $z \le N^{1/5 \log\log N}$ y z tendiendo a infinito con N. Esto permite ampliar el rango de validez de (5.20) desde un tamaño logarítmico de z hasta algo que casi tiene un crecimiento potencial.

Analicemos ahora las consecuencias de (5.30) cuando se aplica al conjunto $\mathcal{A} = \{n(n+2) : n \in \mathbb{Z}^+\} \cap [1, N]$ que aparecía en relación con los primos gemelos. Claramente $|\mathcal{A}_2| = |\mathcal{A}|/2 + O(1)$ y $|\mathcal{A}_p| = 2|\mathcal{A}|/p + O(1)$ si $p > 2$. Entonces definamos $g(d) = 2^{\nu(d)}/d$ si d es impar y $g(d) = 2^{\nu(d)}/2d$ si d es par. Es fácil comprobar que $|\mathcal{A}_d| = |\mathcal{A}|g(d) + r_d$ con $r_d = O\big(2^{\nu(d)}\big)$. El término principal es en este caso

$$\frac{|\mathcal{A}|}{2} \prod_{2<p<z} \big(1 - \frac{2}{p}\big) \asymp \frac{\sqrt{N}}{\log^2 z} \tag{5.31}$$

por (5.8), teniendo en cuenta que $\prod (1 - 2/p)(1 - 1/p)^{-2}$ converge. Un razonamiento similar al del ejemplo anterior acotando por $\big(\sum_{p<z} 2p^{-1}\big)^k/k!$ prueba que el primer sumatorio en (5.30) es de orden inferior que (5.31) tomando $k = [8 \log\log N]$ para cualquier $z \le N^{1/17 \log\log N}$ que

crezca con N. El segundo sumatorio requiere el siguiente artificio:

$$\sum_{\substack{d|P(z)\\ \nu(d)\leq k}} 2^{\nu(d)} \leq z^k \sum_{d|P(z)} \frac{2^{\nu(d)}}{d} = z^k \prod_{p<z} \big(1+\frac{2}{p}\big) \asymp z^k \prod_{2<p<z} \big(1-\frac{2}{p}\big)^{-1} \ll z^k \log^2 z. \tag{5.32}$$

Esta manera de estimar una suma de funciones multiplicativas creando artificialmente una función que se puede desarrollar ventajosamente como un producto, se denomina *truco de Rankin* (en inglés, *Rankin's trick*). En el rango de z indicado, también este segundo sumatorio tiene orden menor que (5.31). En particular, concluimos

$$S\big(\mathcal{A}, N^{1/17\log\log N}\big) \asymp \sqrt{N}\bigg(\frac{\log\log N}{\log N}\bigg)^2. \tag{5.33}$$

¿Es posible deducir algo sobre los primos gemelos a partir de la estimación (5.33)? Como $z = N^{1/17\log\log N}$ es a la larga mucho menor que $N^{1/4}$ en $S(\mathcal{A}, z)$ estarán contados los primos gemelos $z \leq p, p+2 < \sqrt{N}$ y muchos otros números. Por consiguiente podemos asegurar que

$$\pi_2(x) \ll x\bigg(\frac{\log\log x}{\log x}\bigg)^2 \tag{5.34}$$

donde se ha hecho el cambio $x = \sqrt{N}$. Éste es un resultado no trivial y que además está a sólo $(\log\log x)^2$ del orden de magnitud esperado si creemos en (5.15). Por supuesto no nos acerca ni un ápice a la conjetura de los primos gemelos porque para ello necesitaríamos una cota inferior. Sin embargo es suficientemente fuerte para enunciar una consecuencia impactante llamada *teorema de Brun*: La suma de los inversos de los primos gemelos converge. En símbolos

$$\sum_{p,p+2 \text{ primos}} \bigg(\frac{1}{p}+\frac{1}{p+2}\bigg) < \infty. \tag{5.35}$$

El número al que converge la suma es[3] $1.902\ldots$ y se le llama *constante de Brun*. La prueba de (5.35) es un ejercicio empleando sumación por partes y (5.34).

5.5. La criba de Selberg

En los ejemplos que hemos visto de la criba de Eratóstenes-Legendre y de Brun el parámetro z estaba restringido a valores que no alcanzaban una potencia de $|\mathcal{A}|$. Selberg creó en 1947 un método de criba que no tenía esa limitación y que además proporciona cotas superiores

[3] Según [42] esta aproximación requiere algún tipo de de hipótesis al hacer los cálculos.

de orden de magnitud correcto en problemas como la conjetura de los primos gemelos o la conjetura de Goldbach.

La idea básica es muy sencilla y es sorprendente que algo tan simple funcione y dé tan buenos resultados. Para fijar ideas pensemos que $z > 2$ es entero y tomemos $z-1$ números reales $\lambda_1, \lambda_2, \ldots, \lambda_{z-1}$ con $\lambda_1 = 1$ entonces

$$S(\mathcal{A}, z) \leq \sum_{a \in \mathcal{A}} \Big(\sum_{\substack{d<z \\ d|a}} \lambda_d \Big)^2 \tag{5.36}$$

simplemente porque si a no tiene factores primos menores que z el paréntesis es $\lambda_1^2 = 1$. Desarrollando el cuadrado e invirtiendo el orden de sumación se obtiene

$$S(\mathcal{A}, z) \leq \sum_{n,m<z} \lambda_n \lambda_m \big|\mathcal{A}_{[n,m]}\big| \tag{5.37}$$

donde $[n, m]$ indica el mínimo común múltiplo. Sabemos que salvo un término de error $\big|\mathcal{A}_{[n,m]}\big|$ es como $|\mathcal{A}|g([n, m])$. Ahora Selberg considera la forma cuadrática

$$Q = \sum_{n,m<z} \lambda_n \lambda_m g([n, m]) \tag{5.38}$$

y elige los λ_i, $1 < i < z$, de forma que Q alcance un mínimo. En principio calcular este mínimo es tan fácil como derivar, igualar a cero y resolver el sistema, aunque la presencia de $g([n, m])$ hace suponer complicaciones aritméticas. Realmente todo es razonablemente rápido con las definiciones adecuadas. Supongamos que n y m recorren siempre valores $n, m < z$ libres de cuadrados y tales que $g(n), g(m) \neq 0$, lo cual obviamente no constituye una restricción en (5.38). En estas condiciones $g([n, m]) = g(n)g(m)/g\big((n, m)\big)$. Definiendo $h(n) = \big(\sum_{d|n} \mu(d)/g(n/d)\big)^{-1}$ se tiene $h(p) = g(p)/(1 - g(p))$ y $1/g(n) = \sum_{d|n} 1/h(d)$, por la fórmula de inversión de Möbius. Por tanto

$$Q = \sum_{n,m} \lambda_n \lambda_m \frac{g(n)g(m)}{g\big((n, m)\big)} = \sum_{n,m} \lambda_n \lambda_m g(n)g(m) \sum_{\substack{d|n \\ d|m}} \frac{1}{h(d)}. \tag{5.39}$$

Con ello hemos separado las variables. Intercambiando el orden de sumación se obtiene

$$Q = \sum_d \frac{x_d^2}{h(d)} \qquad \text{con} \qquad x_d = \sum_{n \equiv 0 \ (d)} \lambda_n g(n). \tag{5.40}$$

La condición λ_1 se puede escribir como $\sum_\mu(d)x_d = 1$. Utilizando multiplicadores de Lagrange o técnicas más elementales (la forma cuadrática es ahora diagonal), se sigue que el mínimo de Q es exactamente $\big(\sum_{n<z} h(n)\big)^{-1}$ y nos podemos olvidar de las restricciones sobre n escribiendo

$$h(n) = \mu^2(n) \prod_{p|n} \frac{g(p)}{1-g(p)}. \tag{5.41}$$

Es posible (aunque un poco más complicado) probar que los λ_n para los que se alcanza el mínimo cumplen $|\lambda_n| \le 1$ y de ello no es difícil deducir que los términos de error contribuyen en (5.37) menos que $\sum_{d<z^2} 3^{\nu(d)}|r_d|$ y se obtiene la siguiente forma de la *criba de Selberg*:

$$S(\mathcal{A}, z) \le |\mathcal{A}| \bigg(\sum_{n<z} h(n) \bigg)^{-1} + \sum_{d<z^2} 3^{\nu(d)} |r_d|. \tag{5.42}$$

Consideremos por ejemplo el caso $\mathcal{A} = [1, N]$ con $g(d) = 1/d$ y $|r_d| \le 1$, como antes, entonces $h(p) = 1/(p-1) > 1/p$ y de aquí $\sum_{n<z} h(n) \gg \log z$. Por otro lado, el término de error se acota por $O\big(z^2 \log^3 z\big)$ con el truco de Rankin como en (5.32). En definitiva, (5.42) prueba

$$S(\mathcal{A}, z) \ll \frac{N}{\log z} + z^2 \log^3 z. \tag{5.43}$$

Cualquier $z = N^\alpha$, $0 < \alpha < 1/2$, es válido para concluir $\pi(N) \ll N/\log N$.

Esta conclusión es débil a la luz del teorema de los números primos y se obtendría de manera menos enrevesada, y todavía elemental, con los razonamientos de Chebyshev. La fuerza de (5.42) radica en su versatilidad. Por ejemplo, el razonamiento que lleva a (5.43) funciona sin cambios en $\mathcal{A} = [M+1, M+N]$ y con la misma elección de z se llega a

$$\pi(M+N) - \pi(M) \ll \frac{N}{\log N}. \tag{5.44}$$

Cuando $N = O\big(M^\delta\big)$ con $\delta < 1$, tal resultado está fuera del alcance del teorema de los números primos con nuestro pobre conocimiento actual sobre el término de error.

5.6. Comentarios sobre la criba lineal

Un problema esencial de la criba de Selberg es que, en la forma aquí presentada, sólo proporciona cotas superiores. En principio la identidad de Buchstab (5.25) permite crear cotas inferiores a partir de cotas superiores y viceversa gracias al signo negativo del sumatorio. Sin embargo esta identidad no muestra su utilidad hasta que se itera, como hemos visto en la criba de Brun.

La búsqueda de la invariancia por el límite de este proceso iterativo da lugar a una criba combinatoria, llamada *criba de Rosser-Iwaniec* o más recientemente *criba-β*, correspondiente a una elección extraña de $\mathcal{D}^-$ y $\mathcal{D}^+$ en (5.23). Cuando se tiene una condición sobre $g(p)$ algo más fuerte que

$$\prod_{p<z}\big(1-g(p)\big)\asymp\frac{1}{\log z} \tag{5.45}$$

se dice que esta criba es la *criba lineal* y lo curioso es que da resultados óptimos a la luz de los contraejemplos de Selberg. En particular permite obtener

$$S(\mathcal{A},z)\asymp\frac{|\mathcal{A}|}{\log z}\qquad\text{suponiendo}\sum_{\substack{d|P(z)\\ d\le z^s}}|r_d|\ll\frac{|\mathcal{A}|}{(\log|\mathcal{A}|)^\delta} \tag{5.46}$$

para algunos números reales $s>2$ y $\delta>1$. En el caso $\mathcal{A}=[1,N]$ considerado anteriormente conseguimos entonces $S(\mathcal{A},z)\gg N/\log N$ para $z=N^\alpha$, $0<\alpha<1/2$, complementando lo obtenido con la criba de Selberg. Es decir, nos quedamos a las puertas de $\alpha=1/2$ que permitiría detectar primos como en (5.6). Cualquier otro valor $1/(k+1)<\alpha<1/k$ contaría en $S(\mathcal{A},z)$ números con a lo más k factores primos. Es por ello que muchos resultados de criba hablan de casiprimos (números con un número máximo fijado de factores primos).

Para ilustrar la situación en un caso más interesante, partimos del conjunto $\mathcal{A}=\{n^2+1:n\le N\}$. Usando resultados bien conocidos sobre residuos cuadráticos, $|\mathcal{A}_p|=g(p)N+O(1)$ donde $g(2)=1/2$, $g(p)=2/p$ si $p\equiv1\pmod4$ y $g(p)=0$ si $p\equiv3\pmod4$. La condición a la que nos referimos con (5.45) se cumple porque $pg(p)$ es 1 en promedio sobre los primos p. Extendiendo g de manera multiplicativa se verifica en general $|\mathcal{A}_d|=g(d)N+O\big(2^{\nu(d)}\big)$ por el teorema chino del resto, lo que asegura que para $z=N^\alpha$ con $0<\alpha<1/2$ se satisface la hipótesis sobre el error en (5.46), gracias al truco de Rankin. La conclusión $S(\mathcal{A},N^\alpha)\gg N/\log N$ para cualquier $\alpha<1/2$ implica que *hay infinitos primos de la forma n^2+1 con a lo más cuatro factores primos*. Lo mejor que se conoce en la actualidad en este sentido es un resultado de Iwaniec que reduce los cuatro factores a dos. Para ello hay que introducir nuevas ideas respecto al tratamiento del término de error.

La reciente monografía [20] es una referencia general para los métodos de criba en la que se explican de manera asequible las ideas principales y se dan las demostraciones completas de muchos resultados avanzados.

Capítulo 6

La desigualdad de Erdős-Turán

6.1. La discrepancia

Una sucesión $\{x_n\}_{n=1}^{\infty}$ está *uniformemente distribuida* o *equidistribuida* si al tomar muchos elementos la proporción de ellos que está en un subintervalo $[a,b]$ de $[0,1]$ se acerca a $b-a$. En términos matemáticos

$$\lim_{N\to\infty} \frac{\#\{n \le N \,:\, a \le x_n \le b\}}{N} = b-a \qquad \text{para cualesquiera } \ 0 \le a < b \le 1. \tag{6.1}$$

La *discrepancia* es una medida de lo bien uniformemente distribuidos que están los N primeros términos de una sucesión en $[0,1]$. Su definición es

$$D(N) = \sup_{0\le a<b\le 1} \left| \frac{\#\{n \le N \,:\, a \le x_n \le b\}}{N} - (b-a) \right|. \tag{6.2}$$

Algunos autores consideran $a = 0$. Ambas definiciones dan lugar a cantidades comparables [38].

Está claro que si $D(N) \to 0$ cuando $N \to \infty$ entonces $\{x_n\}_{n=1}^{\infty}$ está uniformemente distribuida. El recíproco también es cierto gracias a un sencillo argumento que se puede encontrar en [40, 11.4.1].

Obviamente dados N números en $[0,1]$ siempre hay algún subintervalo de longitud $(N+1)^{-1}$ que no contiene ninguno en su interior entonces $D(N) \neq o(N^{-1})$. Schmidt [46] probó en 1972 que de hecho

$$\limsup_{N\to\infty} \frac{ND(N)}{\log N} > 0{,}01 \tag{6.3}$$

y esto es lo mejor posible salvo el valor de la constante en el sentido de que se conocen ejemplos con este límite acotado superiormente. Por otra parte, en cierto sentido, para "sucesiones típicas"

la discrepancia admite cotas que difieren de $O\big(N^{-1/2}\big)$ en un factor logarítmico [2] (véase también [34]).

La rapidez con que $D(N)$ tiende a cero para una sucesión uniformemente distribuida es habitualmente una medida de su calidad en las aplicaciones, como la generación de números pseudoaletorios o métodos de integración de Montecarlo. A este respecto no es difícil probar integrando por partes [38, §1.2] que

$$\left| \int_0^1 f(x)\,dx - \frac{1}{N}\sum_{n=1}^{N} f(x_n) \right| \le \frac{D(N)}{2} \int_0^1 |f'(x)|\,dx. \tag{6.4}$$

6.2. Utilizando el análisis armónico

Reescribamos la expresión en el interior del supremo en (6.2) como

$$\frac{1}{N}\sum_{n=1}^{N} \chi(x_n) - (b-a) \qquad \text{con } \chi \text{ la función característica de } [a,b]. \tag{6.5}$$

Su desarrollo de Fourier es $b-a+\sum_{m\neq 0} f_m e(mx)$ para ciertos f_m que decaen como $1/m$, lo cual impide la convergencia absoluta. El remedio habitual es la regularización. Considerando una función φ que es igual a χ con unos pequeños añadidos C^∞ en los extremos de anchura δ se puede conseguir $\chi \le \varphi \in C^\infty$ con $\varphi' \ll \delta$ (hay una variación de 1 en una longitud de δ) y en general $\varphi^{(k)} \ll \delta^{-k}$. Integrando por partes los nuevos coeficientes de Fourier serían $\widetilde{f}_m \ll m^{-1}$ para $m < \delta^{-1}$ y algo mejor para $m \ge \delta^{-1}$ integrando más veces. Por otro lado $\widetilde{f}_0 = b-a+O(\delta)$, entonces

$$\frac{1}{N}\sum_{n=1}^{N} \chi(x_n) - (b-a) \ll \delta + \frac{1}{N}\sum_{m<\delta^{-1}} \frac{1}{m}\left|\sum_{n=1}^{N} e(mx_n)\right| + \frac{1}{N}\sum_{|m|\ge\delta^{-1}} |\widetilde{f}_m|\left|\sum_{n=1}^{N} e(mx_n)\right|. \tag{6.6}$$

Tomando φ de forma análoga pero por debajo de χ se obtiene la desigualdad $\gg$. Los $|\widetilde{f}_m|$ decaen muy rápido y por ello el último término debería ser poco importante. La desigualdad de Erdős-Turán dice que nos podemos olvidar completamente de él. Lo habitual es escribir en ella $\delta^{-1} = M+1$ con M entero.

Teorema 6.2.1 (Desigualdad de Erdős-Turán) *Para cualquier $M \in \mathbb{N}$*

$$D(N) \le \frac{1}{M+1} + \frac{3}{N}\sum_{m=1}^{M} \frac{1}{m}\left|\sum_{n=1}^{N} e(mx_n)\right|. \tag{6.7}$$

Se puede interpretar esta desigualdad como una versión cuantitativa del criterio de Weyl [9], [34] pues asegura que si $\sum_{n=1}^{N} e(mx_n)$ son pequeñas entonces la sucesión $\{x_n\}_{n=1}^{\infty}$ tiene discrepancia pequeña. En el artículo original de P. Erdős y P. Turán [18] se establecía la desigualdad sin constantes explícitas, es decir, con $\ll$ en lugar de $\le$.

6.3. Idea de la demostración

Lo que se necesitaría en el argumento de la sección anterior es conseguir que los coeficientes de Fourier $\widetilde{f}_m$ de φ se anulen para $|m| \geq \delta^{-1}$ conservando el resto de las acotaciones empleadas.

Es más cómodo pensar el problema en $\mathbb{R}$ en vez de en $[0,1]$ y después "enrollar" $\mathbb{R}$ sobre $[0,1]$ con los extremos identificados, es decir pasar de φ a $\sum_{n\in\mathbb{Z}} \varphi(\cdot + n)$ cuyo coeficiente de Fourier m-ésimo es la transformada de Fourier $\widehat{\varphi}(m)$.

Queremos saber entonces si para cada $0 < \delta < 1$ existen $\varphi_-, \varphi_+ \in L^1(\mathbb{R}) \cap C^\infty(\mathbb{R})$ tales que

$$\begin{cases} \text{a)} & \varphi_-(x) \leq \chi(x) \leq \varphi_+(x) \\ \text{b)} & \widehat{\varphi}_\pm(0) = b - a \pm \delta \\ \text{c)} & \widehat{\varphi}_\pm(t) = 0 \quad \text{si } |t| \geq \delta^{-1}. \end{cases} \tag{6.8}$$

Aparentemente falta la condición que da $\widehat{\varphi}_\pm(m) \ll m^{-1}$ para $0 \neq |m| < \delta^{-1}$ pero se deduce de a) y b) fácilmente:

$$\begin{aligned} \big|\widehat{\varphi}_\pm(t)\big| &= \left| \int_{-\infty}^{\infty} \big(\varphi_\pm(x) - \chi(x) + \chi(x)\big) e(-tx) dx \right| \leq \int_{-\infty}^{\infty} \big|\varphi_\pm(x) - \chi(x)\big| dx + \big|\widehat{\chi}(t)\big| \\ &= \big|\widehat{\varphi}_\pm(0) - \widehat{\chi}(0)\big| + \big|\widehat{\chi}(t)\big| \leq \delta + \left| \frac{\operatorname{sen}\big(\pi(b-a)t\big)}{\pi t} \right|. \end{aligned}$$

El problema (6.8) es, en principio, de análisis armónico pero lo resolvió Selberg [48, p.213] en relación con ciertas aplicaciones aritméticas. La prueba original de Erdős y Turán utilizaba algo menos preciso. Selberg redescubrió la siguiente función entera introducida originalmente por Beurling en un trabajo no publicado de 1938:

$$B(z) = \frac{\operatorname{sen}^2(\pi z)}{\pi^2} \left(\sum_{n=0}^{\infty} \frac{1}{(z-n)^2} - \sum_{n=1}^{\infty} \frac{1}{(z+n)^2} + \frac{2}{z} \right), \tag{6.9}$$

llamada *función de Beurling-Selberg*, que verifica

$$\begin{cases} 1) & B(x) \geq \operatorname{sgn} x \quad \forall x \in \mathbb{R} \\ 2) & \int_{-\infty}^{\infty} \big(B(x) - \operatorname{sgn} x\big)\, dx = 1 \\ 3) & B(z) = O\big(e^{2\pi|\Im z|}\big) \end{cases} \tag{6.10}$$

De hecho Beurling demostró que era la única con estas propiedades [52]. La prueba de ellas, en contra de lo que pueda parecer, no es complicada empleando la identidad

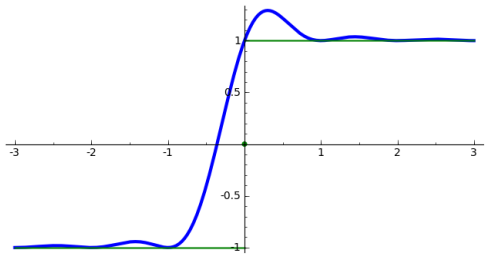
$$\frac{\operatorname{sen}^2(\pi z)}{\pi^2} \sum_{n=-\infty}^{\infty} \frac{1}{(z-n)^2} = 1 \tag{6.11}$$

que se puede deducir a partir de la fórmula clásica $\operatorname{sen}(\pi z) = \pi z \prod_{n=1}^{\infty} \left(1 - z^2/n^2\right)$ tomando logaritmos y derivando dos veces. La prueba detallada de (6.10) está en [40, Th.11.4.3].

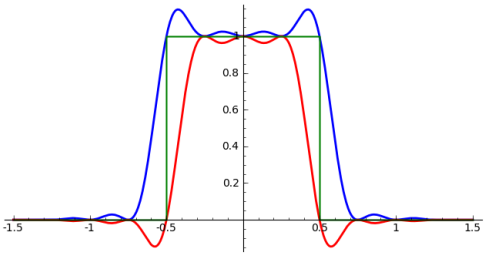
Consideremos ahora

$$\varphi_+(x) = \frac{B\left(\frac{x-a}{\delta}\right) + B\left(\frac{b-x}{\delta}\right)}{2} \qquad \text{y} \qquad \varphi_-(x) = -\frac{B\left(\frac{a-x}{\delta}\right) + B\left(\frac{x-b}{\delta}\right)}{2}. \tag{6.12}$$

De 1) en (6.10) se deduce a) en (6.8) y de 2) se deduce b). Escribiendo $\varphi_\pm(t) = \int \widehat{\varphi}_\pm(x) e(xt)\, dx$ parece claro que si $\widehat{\varphi}_\pm$ tuviera soporte mayor que $[-\delta^{-1}, \delta^{-1}]$ entonces se tendría $\varphi_\pm(iu) \neq O\left(e^{2\pi|u|/\delta}\right)$ contradiciendo 3). El teorema de Paley-Wiener [45] caracteriza todas las transformadas de Fourier enteras con crecimiento exponencial y en particular tiene esta consecuencia. En [40, Th.11.4.4] hay una prueba directa.



La función de Beurling-Selberg



φ_+ y φ_- para $b = -a = 1/2$, $\delta = 1/4$

Es posible dar también una demostración del Teorema 6.2.1 sin pasar por $\mathbb{R}$ construyendo explícitamente las series de Fourier mayorantes y minorantes de χ, a las cuales se denomina a veces polinomios de Vaaler. El procedimiento está descrito con detalle en [38].

6.4. Algunos ejemplos

1. Múltiplos de irracionales cuadráticos. Consideremos la sucesión $x_n = \operatorname{Frac}(n\alpha)$ con $\alpha \notin \mathbb{Q}$. Es bien conocido que está uniformemente distribuida [9]. Si $\|\cdot\|$ denota la distancia al entero más cercano, usando $2\|t\| \leq |\operatorname{sen}(\pi t)|$ se tiene

$$\left|\sum_{n=1}^{N} e(nt)\right| = \frac{|\operatorname{sen}(\pi N t)|}{|\operatorname{sen}(\pi t)|} \leq \frac{1}{2\|t\|} \tag{6.13}$$

y entonces, según el Teorema 6.2.1,

$$D(N) \leq \frac{1}{M+1} + \frac{3}{2N} \sum_{m=1}^{M} \frac{1}{m\|m\alpha\|}. \tag{6.14}$$

Si α es un irracional cuadrático las convergentes a_n/q_n de su fracción continua verifican que q_n es comparable a una progresión geométrica y $|\alpha - a_n/q_n| \asymp q_n^{-2}$. De aquí $\|m\alpha\| = \|ma_n/q_n\| + O\big(mq_n^{-2}\big)$ y siempre eligiendo un q_n adecuado (comparable a m) se tiene $\|m\alpha\| \gg 1/m$, entonces

$$\|m\alpha\| \gg \|ma_n/q_n\| + 1/m \qquad \text{para cualquier} \quad m \ll q_n. \tag{6.15}$$

Tomemos $M = q_0 + q_1 + \cdots + q_K$ que por el crecimiento geométrico satisface $M \asymp q_K$ y $K \asymp \log M$. Sustituyendo en (6.14)

$$\begin{aligned} D(N) &\ll \frac{1}{M} + \frac{1}{N}\sum_{k=1}^{K} \sum_{m=q_0+\cdots+q_{k-1}}^{q_0+\cdots+q_k} \frac{1}{m\|ma_k/q_k\| + 1} \\ &\ll \frac{1}{M} + \frac{1}{N}\sum_{k=1}^{K}\left(1 + \frac{1}{2} + \cdots + \frac{1}{q_k}\right) \ll \frac{1}{M} + \frac{\log^2 M}{N}. \end{aligned}$$

Finalmente la elección $q_K \asymp N/\lg^2 N$ lleva a

$$D(N) \ll \frac{\log^2 N}{N} \qquad \text{para } N > 1 \text{ con } \alpha \text{ irracional cuadrático.} \tag{6.16}$$

Recordando (6.3) vemos que la cota para la discrepancia sólo difiere en un factor logarítmico de lo mejor posible.

2. Espaciamiento entre residuos cuadráticos. En lugar de una sucesión propiamente dicha, consideremos el conjunto $\{x_1, x_2, \ldots, x_N\}$ de residuos cuadráticos módulo p normalizados en $[0,1]$ (divididos entre p) con $p > 3$ primo y $N = (p-1)/2$. Si $1 \le m < p$, empleando las sumas de Gauss

$$\left|\sum_{n=1}^{N} e(mx_n)\right| = \frac{1}{2}\left|\sum_{n=1}^{p} e(mn^2/p) - 1\right| \le \frac{\sqrt{p+1}}{2}. \tag{6.17}$$

La desigualdad de Erdős-Turán asegura

$$D(N) \le \frac{1}{M+1} + \frac{3\sqrt{p+1}}{p-1}\sum_{m=1}^{M}\frac{1}{m} \le \frac{1}{M+1} + \frac{3}{\sqrt{p-3}}\big(\log(M+1) + \gamma\big) \tag{6.18}$$

con $\gamma = 0{,}5772\ldots$ la constante de Euler.

Escogiendo como M la parte entera de $\sqrt{p-3}/\big(3\log\frac{\sqrt{p-3}}{3}\big)$ se tiene $D(N) \ll (\log p)/\sqrt{p}$ y, de hecho, con ayuda de un ordenador se pueden apurar las constantes para conseguir

$$D(N) \le 3\frac{\log p}{\sqrt{p}} \qquad \text{para } p > 13. \tag{6.19}$$

Revisando la definición de la discrepancia (6.2) se deduce que un intervalo entero de longitud mayor que $3\sqrt{p}\log p$ siempre contiene residuos cuadráticos módulo p. Resultados de este tipo se obtienen habitualmente utilizando estimaciones para sumas de caracteres. Con ellas se conoce que para cualquier $\sigma > 1/4$ existe un $P = P(\sigma)$ tal que si $p > P$, entonces un intervalo de longitud p^σ contiene siempre residuos cuadráticos (y también no residuos) módulo p.

3. Puntos del retículo bajo gráficas. Queremos contar el número $\mathcal{N}$ de puntos de coordenadas enteras bajo la gráfica de una función $f \geq 0$ limitada por el eje X en el intervalo entero $[A, B]$. Esto es,

$$\mathcal{N} = \big\{(n,m) \in \mathbb{Z}^2 \,:\, A \leq n \leq B,\ 0 \leq m \leq f(n)\big\}. \tag{6.20}$$

Definiendo $\psi(x) = [x] - x + 1/2$ se tiene

$$\mathcal{N} = \mathcal{P} + \mathcal{E} \qquad \text{con} \quad \mathcal{P} = \frac{B-A+1}{2} + \sum_{n=A}^{B} f(n) \quad \text{y} \quad \mathcal{E} = \sum_{n=A}^{B} \psi\big(f(n)\big). \tag{6.21}$$

El *término principal* $\mathcal{P}$ es fácil de aproximar si f tiene alguna suavidad (que suponemos). Por ejemplo, utilizando la fórmula de sumación de Euler-Maclaurin se obtiene

$$\mathcal{P} = \frac{B-A+f(A)+f(B)}{2} + \int_A^B f(x)\,dx + O\bigg(1 + \int_A^B |f''(x)|\,dx\bigg). \tag{6.22}$$

La dificultad está en el *término de error* $\mathcal{E}$. Lo que vamos a ver es cómo la desigualdad de Erdős-Turán traslada este problema a uno de sumas trigonométricas.

Escribamos $N = B - A + 1$ y consideremos

$$x_n = \mathrm{Frac}\big(f(n+A-1)\big) \qquad \text{con} \quad n = 1, 2, \ldots, N. \tag{6.23}$$

En $[0,1)$, $\psi(x) = 1/2 - x$, por tanto $\int_0^1 \psi(x)\,dx = 0$ y ψ es 1-periódica. Gracias a (6.4)

$$|\mathcal{E}| = \bigg|\sum_{n=1}^{N} \psi(x_n)\bigg| \ll ND(N). \tag{6.24}$$

Una pequeña tecnicalidad es que $\psi'(x)$ estrictamente no existe en 1 porque $\psi(1^-) = -1/2$ y $\psi(1) = 1/2$ y realmente en (6.4) hay que añadir a $\int_0^1 |f'(x)|\,dx = 0$ el valor del salto en 1 (véase el enunciado general preciso en [38]) pero ello sólo afecta al valor de la constante $\ll$.

Uno de los casos típico es en el que $f'' \asymp N^{-1}$, el cual aparece de forma natural al dilatar la gráfica de una función g como $y = Ng(x/N)$. La acotación más básica de van der Corput de sumas trigonométricas afirma que

$$\sum_{n=A}^{B} e\big(mf(n)\big) \ll N^{1/2} m^{1/2}, \tag{6.25}$$

que tras el Teorema 6.2.1 con $M = N^{1/3}$ lleva a

$$\mathcal{N} = \frac{B - A + f(A) + f(B)}{2} + \int_A^B f(x)\,dx + O\big((B - A + 1)^{2/3}\big). \tag{6.26}$$

En general, los métodos de sumas trigonométricas normalmente dan lugar a acotaciones del tipo

$$\sum_{n=A}^{B} e\big(mf(n)\big) \ll (m\Delta)^p N^q. \tag{6.27}$$

para ciertos $0 < p \le 1/2 \le q < 1$ donde $f' \asymp \Delta \gg 1/m$. El Teorema 6.2.1 implica

$$\mathcal{E} \ll \frac{N}{M} + \Delta^p N^q \sum_{m=1}^{M} m^{p-1} \ll \frac{N}{M} + \Delta^p M^p N^q. \tag{6.28}$$

Si $\Delta^p \ll N^{1-q}$ entonces se puede escoger $M^{p+1} \asymp N^{1-q}\Delta^{-p}$ y conseguir

$$\mathcal{E} \ll \Delta^{p/(p+1)} N^{(p+q)/(p+1)}. \tag{6.29}$$

Si $\Delta^p \gg N^{1-q}$ sólo se obtiene la cota trivial.

Capítulo 7

La gran criba

7.1. Introducción

En diversos problemas de teoría analítica de números surge el problema de obtener cancelación en una forma bilineal

(7.1)
$$\mathcal{B}(\vec{x},\vec{y}) = \sum_{m=1}^{M}\sum_{n=1}^{N} x_m b_{mn} y_n = \vec{x}^t B\vec{y} \quad \text{con } \vec{x} = (x_m)_{m=1}^{M}, \quad \vec{y} = (y_n)_{n=1}^{N} \quad \text{y} \quad B = \big(b_{mn}\big)_{m,n=1}^{M,N}$$

donde las coordenadas de $\vec{x}$ e $\vec{y}$ tienen significado demasiado aritmético como para tratar de atacar directamente las sumas en m o en n con métodos analíticos. La cota trivial aplicando dos veces la desigualdad de Cauchy-Schwarz es

$$\mathcal{B}(\vec{x},\vec{y}) \le \|\vec{x}\|\|\vec{y}\|\|B\|_2 \qquad \text{con} \qquad \|B\|_2 = \Big(\sum_{m=1}^{M}\sum_{n=1}^{N} |b_{mn}|^2\Big)^{1/2}. \tag{7.2}$$

En términos generales se llama *desigualdad de gran criba* a una mejora de esta acotación para cierta B con $\vec{x}$ e $\vec{y}$ arbitrarios. Se busca extraer la cancelación inducida por la estructura de la forma bilineal con la idea de que la debida a una elección particular de $\vec{x}$ e $\vec{y}$ es intratable.

El nombre apareció en el trabajo fundacional de Yu.V. Linnik en 1941 y posiblemente quedó definitivamente asentado con la publicación en 1974 del libro de E. Bombieri "La gran criba en la teoría analítica de números" [3]. Proviene de que algunas de estas desigualdades fueron fundamentales para construir métodos de criba que permitían eliminar muchas clases de congruencia. El nombre se debe entonces a razones históricas y es un poco desafortunado y confuso, pues se aplica por igual a las desigualdades y a los métodos de criba (véase la introducción de [44]). Además ya I.M. Vinogradov había explotado la estructura bilineal en 1937 para su famoso teorema sobre sumas de tres primos sin construir ningún nuevo método de criba.

Está claro que

$$\sup_{\vec{x}\neq\vec{0}} \frac{|\mathcal{B}(\vec{x},\vec{y})|^2}{\|\vec{x}\|^2} = \|B\vec{y}\|^2 = \sum_{m=1}^{M} \Big|\sum_{n=1}^{N} b_{mn}y_n\Big|^2. \tag{7.3}$$

Por tanto, si lo preferimos, podemos pensar en formas cuadráticas en vez de en formas bilineales.

La acotación trivial correspondiente a (7.2) es

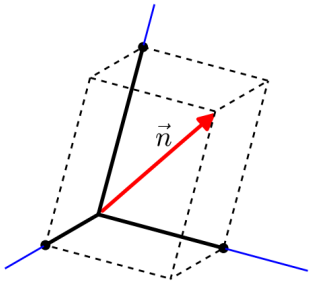
$$Q(\vec{y}) \le \|B\|_2^2\|\vec{y}\|^2 \qquad \text{con} \quad Q(\vec{y}) = \|B\vec{y}\|^2. \tag{7.4}$$

Imaginemos que las filas $\vec{b}_1, \vec{b}_2, \dots, \vec{b}_M$ de B son ortonormales. Esto obliga a $M \le N$ y se puede extender B a una matriz cuadrada ortogonal (o unitaria) $B^* \in \mathcal{M}_{N\times N}$ con el proceso de Gram-Schmidt. Se tiene $\|B\|_2^2 = M$ (cada fila es de norma 1) y (7.4) afirma $Q(\vec{y}) \le M\|\vec{y}\|^2$, sin embargo

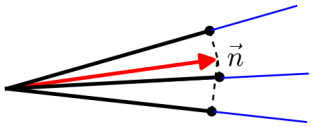
$$Q(\vec{y}) = \|B\vec{y}\|^2 = |\vec{b}_1\cdot\vec{y}|^2 + |\vec{b}_2\cdot\vec{y}|^2 + \cdots + |\vec{b}_M\cdot\vec{y}|^2 \le \|B^*\vec{y}\|^2 = \|\vec{y}\|^2. \tag{7.5}$$

De alguna forma la "independencia" (ortogonalidad) entre las filas es la razón de que (7.4) se mejore M veces.

Las desigualdades de gran criba sustituyen la ortogonalidad, que es demasiado fuerte como para poder emplearla en la práctica, por la hipótesis de que los productos escalares sean pequeños. La idea geométrica es que un vector de norma fijada no puede tener proyección grande con respecto a muchos vectores que tengan direcciones bien distintas porque no puede acercarse a ser paralelo a todos ellos.



$\big(\sum |\mathsf{proyecc.}|^2\big)^{1/2} \approx \|\vec{n}\|$



$\sum |\mathsf{proyecc.}| \approx 3\|\vec{n}\|$

La cadena de desigualdades

$$\begin{aligned} |\vec{x}^t B\vec{y}|^2 \le \|\vec{x}^t B\|^2\|\vec{y}\|^2 &= \|\vec{y}\|^2 \sum_n \Big|\sum_{m=1}^{M} x_m b_{mn}\Big|^2 = \|\vec{y}\|^2 \sum_{k,l=1}^{M} x_k\overline{x}_l \sum_{n=1}^{N} b_{kn}\overline{b}_{ln} \\ &\le \|\vec{y}\|^2 \sum_{k,l=1}^{M} \frac{|x_k|^2+|x_l|^2}{2} \Big|\sum_{n=1}^{N} b_{kn}\overline{b}_{ln}\Big| \le \|\vec{y}\|^2\|\vec{x}\|^2 \max_k \sum_{l=1}^{M} \Big|\sum_{n=1}^{N} b_{kn}\overline{b}_{ln}\Big|, \end{aligned} \tag{7.6}$$

junto con (7.3), implica

$$\|B\vec{y}\|^2 \le \Delta(B)\|\vec{y}\|^2 \qquad \text{con} \qquad \Delta(B) = \max_k \sum_{l=1}^{M} \Big| \sum_{n=1}^{N} b_{kn}\overline{b}_{ln} \Big|. \tag{7.7}$$

Entonces $\Delta(B)$ es el máximo valor que puede tomar la suma de los valores absolutos del producto escalar de una fila por las otras, y esta cantidad cuantifica desde el punto de vista del álgebra lineal la posible mejora de (7.4).

7.2. La desigualdad clásica

Sea $x_1, x_2, \ldots$ números reales y $0 < \delta < 1$ tales que $\delta \le \|x_\nu - x_\mu\|$ para $\nu \ne \mu$ donde $\|\cdot\|$ indica la distancia al entero más cercano. Una de las desigualdades de gran criba más conocida es

$$\sum_{\nu} \Big| \sum_{n=1}^{N} a_n e(nx_\nu) \Big|^2 \le (N + \delta^{-1} - 1) \sum_{n=1}^{N} |a_n|^2 \tag{7.8}$$

para cualesquiera $a_1, a_2, \ldots, a_n \in \mathbb{C}$.

Esta desigualdad es óptima en cierto modo. Si hay un solo x_ν, tomando $a_n = \lambda e(-nx_1)$ ambos miembros de la desigualdad coinciden cuando $\delta \to 1^-$. Por otro lado, si hay muchos x_ν y están equidistribuidos $x_\nu = \nu/K$, $\nu = 1, 2, \ldots, K$; multiplicando ambos miembros por K^{-1} y eligiendo $\delta = 1/K$ se tiene cuando $K \to \infty$ que ambos convergen a $\sum |a_n|^2$.

Si aplicamos (7.7) con $b_{mn} = e(nx_m)$ e $y_n = a_n$, tendríamos

$$\Delta(B) = \max_k \sum_{l=1}^{M} \big| D_N(x_k - x_l) \big| \qquad \text{con } D_N(x) = \sum_{n=1}^{N} e(nx). \tag{7.9}$$

Esencialmente el *núcleo de Dirichlet* $D_N(x)$ restringido (por la periodicidad) a $[-1/2, 1/2]$ es grande, de tamaño a lo más N (la cota trivial), si $|x| \ll N^{-1}$ y se acota por una función que decae si nos alejamos de esta zona [16]. Si despreciamos esta última contribución, $\Delta(B)$ debería estar acotado por

$$\max_k \sum_{l} N \cdot \#\big\{ l \,:\, \|x_k - x_l\| \ll N^{-1} \big\} \ll N\big(\delta^{-1}N^{-1} + 1\big) \ll N + \delta^{-1}, \tag{7.10}$$

ya que $\|x_k - x_l\| \ge \delta$ excepto si $l = k$.

La dificultad para transformar este argumento en una prueba es que al contabilizar la contribución de $|x| \gg N^{-1}$ sale un logaritmo de más debido a que $D_N(x)$ está acotado por

una función que decae demasiado lentamente, como $1/x$, al alejarse del origen. La solución es regularizar el problema con una función ϕ y añadir artificialmente nuevos a_n cubriendo todo el soporte de ϕ, de forma que

$$\sum_\nu \big|\sum_n \widetilde{a}_n \phi(n) e(nx_\nu)\big|^2 = \sum_\nu \big|\sum_{n=1}^N a_n e(nx_\nu)\big|^2, \tag{7.11}$$

para lo cual basta elegir $\widetilde{a}_n = a_n/\phi(n)$ si n está en el soporte de ϕ y $\widetilde{a}_n = 0$ en el resto. Con ello todo funciona como antes reemplazando el núcleo de Dirichlet por $\sum_n \phi(n)e(nx)$ cuyo decaimiento podemos ajustar (véase un ejemplo en [13, §27]). Sin gran esfuerzo se consigue de esta forma probar (7.8) con $\ll$ en lugar de $\leq$. Conseguir la constante óptima 1 es bastante más difícil y las pruebas conocidas emplean la función de Beurling-Selberg o desigualdades de Hilbert generalizadas dentro del esquema anterior [37].

Renunciando a la constante óptima es posible dar una demostración muy breve debida a P.X. Gallagher [13] que no apela a (7.7) y sólo emplea rudimentos de análisis.

Por el teorema fundamental del cálculo, $|f(x)| \leq |f(y)| + \int_I |f'|$ para todo x e y en un intervalo I. Eligiendo un y para el que se alcance el mínimo de $|f|$, se obtiene una hermana menor de las desigualdades de Sobolev:

$$|f(x)| \leq \frac{1}{|I|}\int_I |f| + \int_I |f'|. \tag{7.12}$$

Si subdividimos el intervalo $[0,1)$, donde podemos suponer que están los x_ν, en intervalos semiabiertos I_k de longitud $\delta/2$ entonces cada uno de ellos contiene a lo más un x_ν, y tomando $f(x) = \big(g(x)\big)^2$

$$\sum_\nu |g(x_\nu)|^2 \ll \sum_k \left(\delta^{-1}\int_{I_k}|g|^2 + \int_{I_k}|g'g|\right) \ll \delta^{-1}\int_0^1 |g|^2 + \big(\int_0^1 |g|^2\big)^{1/2}\big(\int_0^1 |g'|^2\big)^{1/2}. \tag{7.13}$$

Eligiendo $g(x) = \sum_{n=1}^N a_n e(nx)$ y aplicando la identidad de Parseval se concluye

$$\sum_\nu \big|\sum_{n=1}^N a_n e(nx_\nu)\big|^2 \ll (N+\delta^{-1})\sum_{n=1}^N |a_n|^2. \tag{7.14}$$

Haciendo los cálculos con un poco de cuidado [13] es posible refinar esta conclusión reemplazando $\ll$ por $\leq$ y $N+\delta^{-1}$ por $\pi N + \delta^{-1}$, lo cual no difiere mucho de (7.8).

Si por ejemplo $a_n = \mu(n)$, los conocimientos actuales permiten ganar a la trivial sólo una potencia de logaritmo en cada una de las sumas interiores (Th.13.10 [32]), mientras que si tenemos N puntos x_ν con $\delta \gg N^{-1}$, entonces (7.14) prueba

$$\frac{1}{N}\sum_{\nu=1}^N \big|\sum_{n=1}^N \mu(n)e(nx_\nu)\big|^2 \ll N. \tag{7.15}$$

La hipótesis de Riemann equivale a que la suma interior para $x_\nu = 0$ sea $O\big(N^{1/2+\epsilon}\big)$ y otros valores de x_ν son conjeturalmente más difíciles de estudiar, lo que da una idea del poder de la gran criba.

En teoría analítica de números el protagonismo de las sumas oscilatorias lo comparten las sumas trigonométricas y las de caracteres. Curiosamente el análogo más natural de (7.8)

$$\sum_{\chi\in\mathcal{C}_q} \big|\sum_{n=1}^{N} a_n\chi(n)\big|^2 \le (q+N)\sum_{n=1}^{N}|a_n|^2 \tag{7.16}$$

donde $\mathcal{C}_q$ es el conjunto de caracteres módulo q, es sencillo a partir de las relaciones de ortogonalidad.

En las aplicaciones, la verdadera desigualdad hermana de (7.8) es

$$\sum_{q\le Q}\frac{q}{\phi(q)}\sum_{\chi\in\mathcal{P}_q} \big|\sum_{n=1}^{N} a_n\chi(n)\big|^2 \le (Q^2+N-1)\sum_{n=1}^{N}|a_n|^2 \tag{7.17}$$

donde $\mathcal{P}_q$ es el conjunto de caracteres primitivos módulo q.

Para probarla se comienza eligiendo en (7.8) como x_ν las fracciones irreducibles en $(0,1]$ con denominador a lo más Q (las fracciones de Farey), obteniéndose

$$\sum_{q\le Q}\sum_{\substack{a=1\\(a,q)=1}}^{q} \big|\sum_{n=1}^{N} a_n e\big(\frac{an}{q}\big)\big|^2 \le (Q^2+N-1)\sum_{n=1}^{N}|a_n|^2. \tag{7.18}$$

El resto de la deducción de (7.17) pasa por hacer algunas manipulaciones (§7.5 [32]) con la serie de Fourier discreta de los caracteres [13]:

$$\chi(a) = \frac{1}{\tau(\overline{\chi})}\sum_{n=1}^{q}\overline{\chi}(n)e\big(\frac{an}{q}\big) \qquad \text{donde } \tau(\overline{\chi}) = \sum_{n=1}^{q}\overline{\chi}(n)e\big(\frac{n}{q}\big). \tag{7.19}$$

Existen muchas otras desigualdades de gran criba útiles en diferentes contextos. Hay una buena colección de ellas en el capítulo 7 de [32].

7.3. La gran criba como método de criba

En líneas generales los métodos de criba investigan cómo se modifica el cardinal de un conjunto al eliminar ciertas clases de congruencia.

Nosotros trabajaremos en $[1,N]$ queriendo acotar superiormente el cardinal Z de un conjunto $\mathcal{Z}\subset[1,N]$ que para cada primo $p\le Q$ omite $\omega(p)$ clases de congruencia. Si $b_1,b_2,\dots,b_{p-\omega(p)}$ son las únicas clases permitidas módulo p, entonces

$$Z^2\le\Big(\sum_{j=1}^{p-\omega(p)}\#\big\{n\in\mathcal{Z}\,;\,n\equiv b_j\ (p)\big\}\Big)^2\le\big(p-\omega(p)\big)\sum_{k=1}^{p}\big(\#\big\{n\in\mathcal{Z}\,;\,n\equiv k\ (p)\big\}\big)^2. \tag{7.20}$$

Ahora escribimos la desigualdad con sumas trigonométricas

$$Z^2\le\big(p-\omega(p)\big)\sum_{n\in\mathcal{Z}}\sum_{m\in\mathcal{Z}}\frac{1}{p}\sum_{a=1}^{p}e\big(\frac{a(n-m)}{p}\big)=\frac{p-\omega(p)}{p}\sum_{a=1}^{p}\big|\sum_{n\in\mathcal{Z}}e\big(\frac{an}{p}\big)\big|^2. \tag{7.21}$$

Para $a=p$ la suma interior es Z, agrupando esta contribución con el primer miembro y despejando,

$$h(p)Z^2\le\sum_{\substack{a=1\\(a,p)=1}}^{p}\big|\sum_{n\in\mathcal{Z}}e\big(\frac{an}{p}\big)\big|^2\qquad\text{con}\qquad h(p)=\frac{\omega(p)}{p-\omega(p)}. \tag{7.22}$$

Dado q no divisible por p, se tiene

$$\sum_{\substack{a=1\\(a,pq)=1}}^{pq}\big|\sum_{n\in\mathcal{Z}}e\big(\frac{an}{pq}\big)\big|^2=\sum_{\substack{a=1\\(a,p)=1}}^{p}\sum_{\substack{b=1\\(b,q)=1}}^{q}\big|\sum_{n\in\mathcal{Z}}e\big(\frac{an}{p}\big)e\big(\frac{bn}{q}\big)\big|^2\ge h(p)\sum_{\substack{b=1\\(b,q)=1}}^{q}\big|\sum_{n\in\mathcal{Z}}e\big(\frac{bn}{q}\big)\big|^2 \tag{7.23}$$

donde la igualdad se debe al teorema chino del resto y la desigualdad se deduce de un cálculo similar al que permitió llegar (7.22) pero ahora con una suma trigonométrica en lugar de Z^2.

Por consiguiente un argumento inductivo prueba que para cualquier q libre de cuadrados

$$h(q)Z^2\le\sum_{\substack{a=1\\(a,q)=1}}^{q}\big|\sum_{n\in\mathcal{Z}}e\big(\frac{an}{q}\big)\big|^2\qquad\text{con}\qquad h(q)=\prod_{p|q}\frac{\omega(p)}{p-\omega(p)}. \tag{7.24}$$

Sumando (7.24) en $q\le Q$ y aplicando (7.18) con $a_n=1$ si $n\in\mathcal{Z}$ y $a_n=0$ en otro caso, se tiene

$$Z\le\frac{N+Q^2}{\sum_{q\le Q}h(q)}\qquad\text{con}\qquad h(q)=\mu^2(q)\prod_{p|q}\frac{\omega(p)}{p-\omega(p)}. \tag{7.25}$$

Aquí el $\mu^2(q)$ sólo sirve para excluir de la suma a los no libres de cuadrados que no hemos considerado (es posible incluirlos complicando notablemente el enunciado §2 [44]).

Esta fórmula se revela como un poderoso método de criba, llamado *gran criba* que, como hemos visto, proviene indirectamente de (7.8).

La expresión aparatosa para $h(q)$ requiere a menudo apelar al *teorema de Wirsing*. Este resultado dice esencialmente que, bajo ciertas hipótesis, si una función multiplicativa no negativa tiene promedio κ sobre los primos entonces sobre los enteros crece como $(\log N)^{\kappa-1}$.

Por ejemplo, para la función idénticamente uno se tiene:

$$\frac{1}{\pi(N)}\sum_{p\le N} 1 \sim 1 \qquad \text{y} \qquad \frac{1}{N}\sum_{n\le N} 1 \sim 1, \tag{7.26}$$

y para la función divisor restringida a los libres de cuadrados:

$$\frac{1}{\pi(N)}\sum_{p\le N} 2 \sim 2 \qquad \text{y} \qquad \frac{1}{N}\sum_{n\le N} \mu^2(n)d(n) \sim C\log N. \tag{7.27}$$

En §6.6 de [32] hay estimaciones más concretas del denominador de (7.25) para situaciones más o menos genéricas.

Un ejemplo típico que se suele poner para ilustrar la fuerza de (7.25) es el conjunto $\mathcal{Z}$ obtenido al quitar de $[1,N]$ todos los elementos que no sean residuos cuadráticos módulo los primos $p \le \sqrt{N}$. En esta situación $\omega(p) = (p+1)/2$ y $h(p) = (p+1)/(p-1)$ que tiene promedio 1 sobre los primos, por tanto el denominador en (7.25) es $\gg Q$ y eligiendo $Q = \sqrt{N}$ se llega a $Z \ll \sqrt{N}$. Obviamente $\mathcal{Z} \supset \{1^2, 2^2, 3^2, \dots, [\sqrt{N}]^2\}$ y entonces la cota obtenida es la mejor posible salvo una constante multiplicativa.

La forma de (7.25) induce a pensar que hay alguna relación intrínseca entre la gran criba y la criba de Selberg. Este punto se investiga y desarrolla en [20] y [44]. En la práctica (7.25) es menos versátil, pues practicamente obliga a tomar $Q = \sqrt{N}$ mientras que la criba de Selberg permite incorporar la información sobre términos de error. Por otro lado, sobre todo cuando se omiten muchas clases, un control aceptable del error es difícil de conseguir y la simplicidad de (7.25) se muestra más conveniente.

7.4. Algunas aplicaciones

Primos gemelos. Buscamos estimar $\pi_2(N)$, el cardinal de los $n \le N$ tales que n y $n+2$ sean primos (gemelos). Para ello consideramos el conjunto

$$\mathcal{Z} = \big\{n \in [1,N] \,:\, n \not\equiv 0,\ n \not\equiv -2 \ (p) \quad \forall p \le \sqrt{N}\big\}. \tag{7.28}$$

Entonces para cada $2 < p \le \sqrt{N}$ se tiene $h(p) = 2/(p-2)$ y $h(2) = 1/2$ y se cumple para $q \le \sqrt{N}$

$$h(q) \ge \frac{1}{2}\mu^2(q)\prod_{p\mid q}\frac{2}{p} = \frac{\mu^2(q)d(q)}{2q}. \tag{7.29}$$

Sumando por partes usando (7.27), se tiene

$$\sum_{q\le\sqrt{N}} h(q) \gg (\log N)^2. \tag{7.30}$$

Evidentemente $\pi_2(N) \le Z + \sqrt{N}$ y (7.25) con $Q = [\sqrt{N}]$ prueba

$$\pi_2(N) \ll \frac{N}{(\log N)^2}. \tag{7.31}$$

El resultado es espectacular teniendo en cuenta que la fórmula asintótica de la conjetura de los primos gemelos es $\pi_2(N) \sim CN/(\log N)^2$ con una constante específica.

Primos en intervalos pequeños. El teorema de los números primos afirma

$$\pi(N) = \mathrm{Li}(N) + E \qquad \text{con } E = o\big(\mathrm{Li}(N)\big) \tag{7.32}$$

y nadie ha conseguido probar $E = O(N^\alpha)$ para ningún $\alpha < 1$, lo cual implicaría que no hay ceros de la función ζ en $\Re s > \alpha$. Por ello es sorprendente que en 1930 G. Hoheisel probase incondicionalmente

$$\pi(M+N) - \pi(M) \sim \frac{N}{\log M} \qquad \text{para } M^\alpha < N < M \tag{7.33}$$

con $\alpha = 0.9996\ldots$ Posteriormente se desarrolló una teoría que relacionaba este tipo de asintótica con resultados llamados de densidad, que establecen que los posibles ceros incumpliendo la hipótesis de Riemann son muy escasos (tienen poca densidad) en ciertos rectángulos. En 1972 M.N. Huxley [29] consiguió probar (7.33) para cualquier $\alpha > 7/12$ utilizando estas ideas y no ha habido otros avances desde entonces.

Nosotros vamos a establecer un caso particular del teorema de Brun-Titchmarsh que sustituye la asintótica (7.33) por una cota superior sin restricciones sobre el rango. Concretamente

$$\pi(M+N) - \pi(M) \le \big(2 + o(1)\big)\frac{N}{\log N} \qquad \text{para todo } 1 < N < M \tag{7.34}$$

donde $o(1)$ tiende a cero cuando N crece.

La observación clave es que el rango de sumación de n en (7.8) se puede cambiar de $[1, N]$ a $[M+1, M+N]$ porque ello equivale a multiplicar por $e(Mx_\nu)$ que tiene módulo 1. En consecuencia (7.25) se aplica igualmente a $[1, N]$ y a $[M+1, M+N]$.

En el conjunto

$$\mathcal{Z} = \big\{n \in [M+1, M+N] \,:\, n \not\equiv 0 \ (p) \quad \forall p \le Q\big\} \tag{7.35}$$

están todos los primos que son mayores que Q, por tanto

$$\pi(M+N)-\pi(M)\le Z+Q. \tag{7.36}$$

Como se excluye una clase por cada primo, $h(p)=1/(p-1)=p^{-1}+p^{-2}+p^{-3}+\dots$ por tanto $h(q)$ es la suma de los inversos de todos los números con los mismos factores primos que q, así pues

$$\sum_{q\le Q} h(q)\ge \sum_{q\le Q}\frac{1}{q}=\log Q+O(1). \tag{7.37}$$

La cota (7.25) con $Q=\sqrt{N}/\log N$ termina la prueba de (7.34).

Un Teorema de Linnik. Conjeturalmente los residuos y no residuos cuadráticos módulo un primo p están bien mezclados y el primer no residuo cuadrático $n(p)$ no debería tardar mucho en aparecer. De hecho se sabe [1] (ver también §9.2 [38]) que suponiendo la hipótesis de Riemann generalizada se cumple $n(p)\ll(\log p)^2$. El mejor resultado incondicional en esta línea es $n(p)\ll p^{\alpha}$ para todo $\alpha>(16e)^{-1/2}$ y tiene más de 50 años [7].

Linnik probó que fijado $\epsilon>0$ existe C_ϵ tal que

$$\#\big\{p\le N\,:\,n(p)>N^{\epsilon}\big\}<C_\epsilon. \tag{7.38}$$

Este resultado es uno de los que originaron la gran criba.

La demostración se sale del esquema de los ejemplos anteriores pues emplearemos Z para acotar el denominador de (7.25) en vez de ser al revés. Además ahora no se cribará con todos los primos hasta Q sino sólo con los de

$$\mathcal{P}=\#\big\{p\le Q\,:\,\text{todo } n\le N^{\epsilon} \text{ es residuo cuadrático módulo } p\big\}. \tag{7.39}$$

En la práctica esto es como poner $\omega(p)=0$ para el resto de los primos en $[2,Q]$.

Consideramos el conjunto

$$\mathcal{Z}=\big\{n\le N^2\,:\,n \text{ no es no residuo módulo } p,\quad \forall p\in\mathcal{P}\big\}. \tag{7.40}$$

Igual que en el ejemplo de prueba con los cuadrados, excluir los no residuos conduce a $h(p)=(p-1)/(p+1)$. Nótese que, a diferencia de los ejemplos anteriores, aquí el nombre de "gran criba" está justificado pues $\omega(p)$ es comparable a p. El denominador de (7.25) para $Q=N$ es

$$\sum_{q\le N} h(q)\ge\sum_{p\in\mathcal{P}} h(p)\ge\frac{1}{3}\#\big\{p\le N\,:\,n(p)>N^{\epsilon}\big\}. \tag{7.41}$$

Entonces (nótese que partimos del intervalo $[1,N^2]$ en vez de trabajar en $[1,N]$)

$$\#\big\{p\le N\,:\,n(p)>N^{\epsilon}\big\}\ll\frac{N^2}{Z}. \tag{7.42}$$

En $\mathcal{Z}$ están todos los números de $[1, N^2]$ que tienen todos sus factores primos $\leq N^\epsilon$ porque los primos de este tamaño son residuos cuadráticos módulo cualquier $p \in \mathcal{P}$ (por definición de $\mathcal{P}$) y el producto de residuos es residuo. Es posible dar una fórmula asintótica para los números con este tipo de factores primos, pero siguiendo §7.4 [32] utilizamos la desigualdad elemental:

$$Z \geq \#\big\{n \leq N^2 \,:\, p \leq N^\epsilon \quad \forall p \mid n\big\} \geq \#\big\{mp_1p_2\dots p_{[2\epsilon^{-1}]} \leq N^2 \,:\, N^{\epsilon-\epsilon^2/2} < p_j \leq N^\epsilon\big\} \tag{7.43}$$

que implica

$$Z \gg \sum_{\substack{N^{\epsilon-\epsilon^2/2}<p_j\leq N^\epsilon \\ j=1,2,\dots[2\epsilon^{-1}]}} \frac{N^2}{p_1p_2\cdots p_{[2\epsilon^{-1}]}} \gg N^2 \tag{7.44}$$

ya que la suma de los inversos de los primos en $[N^\alpha, N^\beta]$ con $\alpha < \beta$ es $\gg 1$. Sustituyendo en (7.42) se obtiene (7.38).

Capítulo 8

Ideas sobre el método del círculo

8.1. Introducción

Normalmente trabajar con congruencias es más fácil que trabajar con los enteros, por ello sería conveniente poder considerar las soluciones enteras de una ecuación como algún tipo de "límite algebraico" de soluciones de congruencias. Cuando se logra tal cosa en teoría algebraica de números se dice que se tiene un principio local-global [36]. El ejemplo más conocido, y a decir verdad uno de los pocos que existen, es el teorema de Hasse-Minkowski [4] que se aplica en el contexto de las formas cuadráticas. En teoría analítica de números hay métodos que retoman esta idea sin buscar en general expresiones exactas. Uno de ellos es el *método del círculo*, también llamado *método de Hardy-Littlewood* pues fue desarrollado por estos dos matemáticos en los años 20 y 30 del siglo XX.

El origen del método está en un artículo de Hardy y Ramanujan en el que estudiaban la función partición $p(n)$, que cuenta el número de maneras en que n se puede escribir como suma de enteros positivos permitiendo repeticiones y sin importar el orden. Con razonamientos elementales [26] se llega a la siguiente expresión para su función generatriz:

$$F(z) = \prod_{k=1}^{\infty} \left(1 - z^k\right)^{-1} = 1 + p(1)z + p(2)z^2 + p(3)z^3 + p(4)z^4 + \dots \tag{8.1}$$

La fórmula integral de Cauchy permite despejar $p(n)$ en función de $F(z)$ como

$$p(n) = \frac{1}{2\pi i} \int_C F(z) \frac{dz}{z^{n+1}} \tag{8.2}$$

para cualquier curva $C \subset \{|z| < 1\}$ que sea la frontera de una región D que contiene al origen. Gracias a (8.1) es posible aproximar F cerca del "polo de orden infinito" en $z = 1$ y lo mismo ocurre cerca de $z = e(a/q)$, donde $e(t)$ abrevia $e^{2\pi it}$, estando la cercanía necesaria para

sentir la singularidad en relación con el tamaño del denominador q. La función F tiene muchas simetrías, que provienen de cierta forma modular asociada, estableciendo relaciones entre los valores de F en diferentes puntos. Esta particularidad fue aprovechada por H. Rademacher [43] para lograr una fórmula "exacta" para $p(n)$ como una serie infinita cuyo término principal es $(4n\sqrt{3})^{-1}e^{\pi\sqrt{2n/3}}$, que ya habían obtenido Hardy y Ramanujan como asintótica de $p(n)$ (véase [32, §20]).

Otras funciones generatrices no gozan de estas simetrías especiales y una curva C cercana a la circunferencia unidad queda dividida en arcos en los que el integrando se puede aproximar bien por estar cerca de $e(a/q)$ con q pequeño y otros en los que una aproximación no es posible y se emplea una acotación. Los primeros reciben el nombre de *arcos mayores* y los segundos, *arcos menores*. Los adjetivos se refieren a su mayor o menor contribución, no a su tamaño.

Hardy y Littlewood aplicaron esta idea, el método del círculo, a diversos ejemplos que aquí abstraemos en el problema general consistente en dar una aproximación asintótica del número de representaciones, $r_k(N)$, de un número N grande como suma de k elementos de un conjunto $\mathcal{B}$ de enteros no negativos. Es decir, se busca una fórmula asintótica para

$$r_k(N) = \#\{(b_1, b_2, \dots, b_k) \in \mathcal{B}^k \,:\, N = b_1 + b_2 + \cdots + b_k\}. \tag{8.3}$$

El análogo de (8.2) es ahora

$$r_k(N) = \frac{1}{2\pi i}\int_C \Big(\sum_{b\in\mathcal{B}} z^b\Big)^k \frac{dz}{z^{N+1}} \qquad \text{ya que} \quad \Big(\sum_{b\in\mathcal{B}} z^b\Big)^k = \sum_{n=0}^{\infty} r_k(n) z^n. \tag{8.4}$$

En realidad podemos eliminar de $\mathcal{B}$ todos los números mayores que N pues no contribuyen a $r_k(N)$, con ello nos libramos de la serie infinita, que causaba complicaciones en los trabajos de Hardy y Littlewood y cuya razón de ser en el caso de las particiones era emplear las simetrías. Entonces $\sum_{b\in\mathcal{B}} z^b$ es un polinomio y nada impide escoger C como la propia circunferencia unidad. Parametrizando $z = e(x)$ nos olvidamos del uso artificioso de la variable compleja en (8.4) y tomamos como punto de partida la sencilla igualdad:

$$r_k(N) = \int_I S^k(x)\, e(-Nx)\, dx \qquad \text{con} \quad S(x) = \sum_{b\in\mathcal{B}} e(bx) \tag{8.5}$$

donde I es cualquier intervalo de longitud uno (si se prefiere, integramos en el toro unidimensional $\mathbb{T}$). Los arcos mayores serán ahora los subintervalos de I en los que tengamos una buena aproximación para $S(x)$, de la que saldrá un término principal; mientras que en el resto, los arcos menores, confiamos en que acotaciones de la suma trigonométrica $S(x)$ sean suficientes para acumular su contribución en un témino de error.

A partir de las propiedades de distribución de los $b \in \mathcal{B}$ en progresiones aritméticas (soluciones de congruencias) se extraen fórmulas aproximadas para $S(x)$ cuando x está cerca de

racionales con denominador moderadamente pequeño y por tanto estos conjuntos constituirán los arcos mayores que generan el término principal. Es de esta forma cómo se traspasan propiedades locales a una cuantificación de $r_k(N)$.

A pesar de la generalidad en el planteamiento, aquí nos centraremos en el caso de sumas de primos. La monografía [53] es una buena referencia para otros ejemplos y para el método del círculo en general.

8.2. La serie singular

Para estudiar problemas del tipo de la conjetura de Goldbach, debemos escoger en (8.5)

$$S(x) = \sum_{p \le N} e(px). \tag{8.6}$$

De esta forma $r_k(N)$ será el número de representaciones de N como suma de k primos.

De alguna forma, el teorema de los números primos afirma que la probabilidad de que n sea primo es aproximadamente $1/\log n$. Por tanto para x muy cerca de cero debería cumplirse

$$S(x) \sim \sum_{1<n\le N} \frac{e(nx)}{\log n} \tag{8.7}$$

y de hecho no es difícil probarlo comprobando que $\sum(a_n - 1/\log n)e(nx)$ es pequeño, sumando por partes, donde $a_p = 1$ y $a_n = 0$ si n es compuesto. Para simplificar la exposición, excusando el rigor, escribiremos la fórmula menos precisa

$$S(x) = \frac{D(x)}{\log N} + \text{error} \qquad \text{con} \qquad D(x) = \sum_{n \le N} e(nx), \tag{8.8}$$

que equivale a la anterior cuando x es muy pequeño. Podemos entender que $1/\log N$ proviene de que la densidad de los primos hasta N es $\pi(N)/N \sim 1/\log N$.

Por otro lado, muy cerca por ejemplo de $x = 1/2$, la aproximación debe ser

$$S(x) = -\frac{D(x-1/2)}{\log N} + \text{error} \tag{8.9}$$

simplemente porque todos los primos $p \neq 2$ son impares y por tanto $e(px) = -e(p(x-1/2))$. Si $x = 1/3$, debemos considerar el hecho de que $e(p/3)$ es $e(1/3)$ ó $e(2/3)$ dependiendo de si $p \equiv 1\ (3)$ o $p \equiv 2\ (3)$. Como la "mitad" de los primos es de cada uno de estos dos tipos,

$$S(x) = \left(\frac{e(1/3)}{2\log N} + \frac{e(2/3)}{2\log N}\right) D(x-1/3) + \text{error} = -\frac{D(x-1/3)}{2\log N} + \text{error}. \tag{8.10}$$

En general, el teorema de los números primos en progresiones aritméticas asegura que los primos están equidistribuidos en cada una de las $\phi(q)$ progresiones aritméticas módulo q que contienen infinitos primos y procediendo como antes se llega a que muy cerca de la fracción irreducible $x = a/q$ se cumple

$$S(x) = \frac{\mu(q)}{\phi(q)\log N} D(x - a/q) + \text{error}. \tag{8.11}$$

donde $\mu(q)$ proviene de evaluar $\sum_{(n,q)=1} e(n/q)$ con (n, q) representando el máximo común divisor de n y q. De hecho en general se tiene la siguiente evaluación de las llamadas *sumas de Ramanujan* [32, §3.2]

$$c_q(N) = \sum_{\substack{n=1 \\ (n,q)=1}}^{q} e\big(N\frac{n}{q}\big) = \sum_{d|(q,N)} \mu\big(\frac{q}{d}\big) d = \mu(q')\frac{\phi(q)}{\phi(q')} \qquad \text{con} \quad q' = \frac{q}{(q,N)}. \tag{8.12}$$

Por otro lado, la función $D(x)$ es pequeña si x no está próxima a un entero, lo que sugiere que no se pierde mucho aproximando $\int_{|x|<\epsilon} \big(D(x)\big)^k e(-Nx)\,dx$ por

$$\int_{-1/2}^{1/2} \big(D(x)\big)^k e(-Nx)\,dx = \int_0^1 \big(D(x)\big)^k e(-Nx)\,dx = \binom{N-1}{k-1} \sim \frac{N^{k-1}}{(k-1)!}. \tag{8.13}$$

Teniendo esto en cuenta, al sustituir todas estas aproximaciones en (8.5) y descartar los términos de error, se llega a

$$\sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \frac{\mu^k(q) e(-Na/q)}{\phi^k(q)(\log N)^k} \int_0^1 \big(D(x)\big)^k e(-Nx)\,dx \sim \sum_{q=1}^{\infty} \frac{\mu^k(q) c_q(-N)}{\phi^k(q)(\log N)^k} \frac{N^{k-1}}{(k-1)!} \tag{8.14}$$

Usando la evaluación (8.12) y la multiplicatividad de las funciones de q que aparecen en la última suma, no es difícil escribirla como

$$\frac{\mathfrak{S}_k(N) N^{k-1}}{(k-1)!(\log N)^k} \qquad \text{con} \quad \mathfrak{S}_k(N) = \prod_{p\nmid N} \left(1 + \frac{(-1)^{k+1}}{(p-1)^k}\right) \prod_{p|N} \left(1 + \frac{(-1)^k}{(p-1)^{k-1}}\right). \tag{8.15}$$

Es interesante notar que incluso si hay obstáculos insuperables para completar este análisis con rigor, fórmulas como (8.15) nos dan mucha intuición acerca de lo que podemos esperar. En otras palabras, el planteamiento del método del círculo es útil incluso cuando no funciona.

A $\mathfrak{S}_k(N)$ se le llama *serie singular* e incluye la información local (sobre congruencias). Aunque la relación local-global es más evidente cuando se tratan sumas de potencias [32,

§20.2], explicaremos cómo se manifiesta en (8.15). En nuestro caso, $\mathcal{B}$ en (8.3) es el conjunto de primos y su análogo natural módulo q son los números coprimos con q, es decir, el grupo de unidades $\mathcal{U}(\mathbb{Z}/q\mathbb{Z})$. La "densidad" de las soluciones locales es

$$\delta_q = \frac{\#\big\{(n_1,\dots,n_k)\in \mathcal{U}(\mathbb{Z}/q\mathbb{Z})^k \,:\, n_1+\cdots+n_k\equiv N \pmod q\big\}}{\phi^k(q)}. \tag{8.16}$$

Por otro lado, si seguimos el rastro a las expresiones que dieron lugar a la serie singular, se tiene

$$\mathfrak{S}_k(N) = \sum_{q=1}^{\infty}\sum_{\substack{a=1\\(a,q)=1}}^{q}\Big(\frac{1}{\phi(q)}\sum_{\substack{n=1\\(n,q)=1}}^{q} e\big(\frac{na}{q}\big)\Big)^k e\big(-\frac{Na}{q}\big) \tag{8.17}$$

$$= \sum_{q=1}^{\infty}\frac{1}{\phi^k(q)}\sum_{\substack{n_1=1\\(n_1,q)=1}}^{q}\cdots\sum_{\substack{n_k=1\\(n_k,q)=1}}^{q}\sum_{\substack{a=1\\(a,q)=1}}^{q} e\big(\frac{(n_1+\cdots+n_k-N)a}{q}\big) \tag{8.18}$$

Empleando (8.12), la suma interior es $\sum \mu(q/d)d$ con la sumación restringida a los $d\mid q$ tales que $n_1+\cdots+n_k\equiv N \pmod d$. Cada $n_i\in\mathcal{U}(\mathbb{Z}/d\mathbb{Z})$ corresponde a $\phi(q)/\phi(d)$ elementos en $\mathcal{U}(\mathbb{Z}/q\mathbb{Z})$, entonces

$$\mathfrak{S}_k(N) = \sum_{q=1}^{\infty}\frac{1}{\phi^k(q)}\cdot\sum_{d\mid q}\mu(q/d)d\cdot\Big(\frac{\phi(q)}{\phi(d)}\Big)^k\phi^k(d)\delta_d = \sum_{q=1}^{\infty}\sum_{d\mid q}\mu(q/d)d\delta_d. \tag{8.19}$$

Esta expresión es la suma de una convolución de funciones multiplicativas, además para $q=p^\alpha$, $\mu(q/d)=0$ excepto cuando $d=p^{\alpha-1}$ o $d=p^\alpha$. Por tanto, con la definición natural $\delta_1=0$,

$$\mathfrak{S}_k(N) = \prod_p\sum_{\alpha=1}^{\infty}\big(p^\alpha\delta_{p^\alpha}-p^{\alpha-1}\delta_{p^{\alpha-1}}\big) = \prod_p p\delta_p. \tag{8.20}$$

Para la última igualdad se usa que los términos entre paréntesis son nulos si $\alpha>1$.

8.3. La división en arcos mayores y menores

Sin necesidad de poner en rigor las aproximaciones de la sección anterior para $S(x)$, vamos a dar indicios de sus rangos de validez.

Pensemos por ejemplo en la aproximación de $S(x)$ para $x=\epsilon$ próximo a cero. Sumando por partes

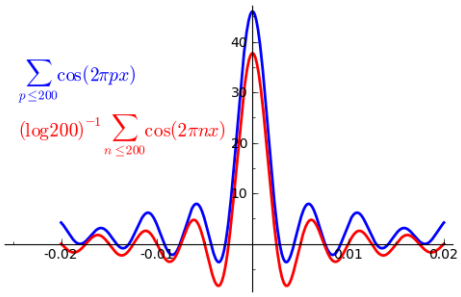
$$S(\epsilon) = \sum_{p\le N} e(p\epsilon) = \pi(N)e(N\epsilon) - 2\pi i\epsilon\int_1^N \pi(t)e(\epsilon t)\,dt. \tag{8.21}$$

Por el teorema de los números primos sabemos que $\pi(t) = \mathrm{Li}(x) + O\big(t/(\log t)^B\big)$ para cualquier $B > 0$ (hay un error mejor pero no muchísimo mejor y elegimos éste para simplificar). Entonces el término de error en $S(\epsilon)$ será comparable a

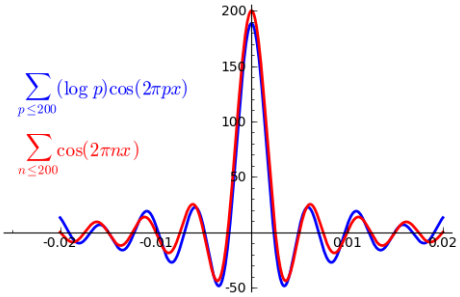
$$\epsilon \int_2^N \frac{t}{(\log t)^B}\, dt \ll \frac{\epsilon N^2}{(\log N)^B}. \tag{8.22}$$

Si queremos que esto sea de orden menor que $\pi(N)$ que es la cota trivial para $S(\epsilon)$, debe cumplirse $|\epsilon| < N^{-1}(\log N)^A$ para algún $A < B$. En el caso $S(a/q + \epsilon)$ se llega a una condición similar, con el agravante de que nuestro poco conocimiento sobre la uniformidad en el módulo para primos en progresiones aritméticas sólo permite considerar q menor que una potencia de logaritmo de N (recuérdese el teorema de Siegel-Walfisz).

Nótese que estas consideraciones a partir de (8.21) conducen a un término principal demasiado complicado. Al igual que en la demostración del teorema de los números primos la función que aparece naturalmente es $\psi(x)$ en vez de $\pi(x)$, aquí es técnicamente más simple trabajar con $S^*(x) = \sum_{p \le N} (\log p) e(px)$. La aproximación de $S^*(\epsilon)$ mediante $D(\epsilon)$ da lugar a la ganancia indicada de una potencia de logaritmo, mientras que la de $S(\epsilon)$ mediante $D(\epsilon)/\log N$, sugerida en la sección anterior, es más débil. La situación es paralela a las acotaciones de la teoría de la distribución de los primos $\psi(x) \ll x/(\log x)^B$ para todo $B > 0$ mientras que $\pi(x) - x/\log x \gg x/(\log x)^2$.



$S(x)$ y $D(x)/\log N$ para $N = 200$



$S^*(x)$ y $D(x)$ para $N = 200$

Más allá de la realización técnica, el análisis anterior sugiere que la elección natural de los arcos mayores es

$$\mathfrak{M} = \Big\{0 \le x \le 1 \,:\, \Big|x - \frac{a}{q}\Big| < N^{-1}(\log N)^A \text{ con } q \le (\log N)^A\Big\} \tag{8.23}$$

donde se entiende que $a/q \in [0,1]$ es una fracción irreducible. La medida de este conjunto es

$$|\mathfrak{M}| \le 2 \sum_{q \le (\log N)^A} \sum_{a=1}^{q} N^{-1}(\log N)^A \le 2N^{-1}(\log N)^{3A} \tag{8.24}$$

y por tanto tiende a cero cuando $N \to \infty$. En este cálculo se ha usado implícitamente que los intervalos que componen los arcos mayores son disjuntos. Esto es elemental ya que para dos fracciones irreducibles distintas

$$\Big|\frac{a}{q} - \frac{a'}{q'}\Big| \geq \frac{1}{qq'} \geq (\log N)^{-2A} \gg (\log N)^A \tag{8.25}$$

Por otro lado, los arcos menores conforman el conjunto $\mathfrak{m} = [0,1] - \mathfrak{M}$ y constituyen casi todos los puntos en el sentido de la medida, puesto que $|\mathfrak{m}| \to 1$ cuando $N \to \infty$. No obstante, esperamos que su contribución sea menor que la del raquítico conjunto $\mathfrak{M}$.

8.4. Aplicación y limitaciones del método del círculo

El esquema del método del círculo responde a la fórmula

$$r_k(N) = \int_{\mathfrak{M}} S^k(x)\, e(-Nx)\, dx + O\Big(\int_{\mathfrak{m}} |S(x)|^k\, dx\Big) = I_{\mathfrak{M}} + O\big(I_{\mathfrak{m}}\big) \tag{8.26}$$

ya que, partiendo de (8.5), habíamos definido $\mathfrak{m}$ como el conjunto donde no sabíamos aproximar $S(x)$ y por tanto no se puede aprovechar la oscilación de $e(-Nx)$.

Para el caso de sumas de primos, la heurística desplegada en relación con la serie singular funciona en los arcos mayores y entonces se tiene

$$I_{\mathfrak{M}} \sim \frac{\mathfrak{G}_k(N) N^{k-1}}{(k-1)!(\log N)^k} \qquad \text{con} \quad \mathfrak{G}_k(N) = \prod_{p\nmid N}\Big(1 + \frac{(-1)^{k+1}}{(p-1)^k}\Big)\prod_{p\mid N}\Big(1 + \frac{(-1)^k}{(p-1)^{k-1}}\Big). \tag{8.27}$$

I.M. Vinogradov consiguió en 1937 utilizar incondicionalmente el método del círculo para sumas de primos probando la acotación:

$$S(x) \ll Nq^{-1/2} + N^{1/2}q^{1/2} + Ne^{-\frac{1}{2}\sqrt{\log N}} \qquad \text{para} \quad \Big|x - \frac{a}{q}\Big| < \frac{1}{q^2} \ \text{ con } \ \frac{a}{q} \ \text{ irreducible.} \tag{8.28}$$

Todo x admite una aproximación de este tipo por un teorema elemental de P.G.L. Dirichlet [26, §11.3], con lo cual esta acotación es general. De hecho ese teorema asegura que, fijado cierto Q, para cualquier x existe a/q con $q \leq Q$ tal que $|x - a/q| < 1/qQ$. Si escogemos $Q = N(\log N)^{-A}$ tendremos $(\log N)^A \leq q \leq Q$ para $x \in \mathfrak{m}$ y (8.28) implica

$$S(x) \ll N(\log N)^{-A/2} \qquad \text{para todo } x \in \mathfrak{m}. \tag{8.29}$$

Aplicando esta cota directamente obtenemos $I_{\mathfrak{m}} = O\big(N^k(\log N)^{-Ak/2}\big)$ que es terriblemente malo porque supera a (8.27). En términos analíticos el problema es que estamos usando sólo

una acotación L^∞ para lograr una acotación L^k y eso no es buena idea, es mejor ayudarse (interpolar) con otra acotación L^p. Sin tanta palabrería, empleando

$$\int_0^1 |S(x)|^2\,dx = \sum_{p\le N} 1^2 = \pi(N) \tag{8.30}$$

se tiene

$$I_{\mathfrak{m}} \ll \big(N(\log N)^{-A/2}\big)^{k-2} \int_0^1 |S(x)|^2\,dx = O\big(N^{k-1}(\log N)^{-B}\big) \tag{8.31}$$

para B arbitrariamente grande cuando A crece, siempre que $k>2$. Entonces hemos probado

$$r_k(N) \sim \frac{\mathfrak{G}_k(N)N^{k-1}}{(k-1)!(\log N)^k} \qquad \text{para } k>2. \tag{8.32}$$

Nos hemos dejado fuera exactamente la (famosa) conjetura binaria de Goldbach. Las malas noticias son que hay dificultades teóricas para alcanzarla.

Los términos $e(px)$ en (8.6) no tienen ninguna razón por la que deban resonar para $x\notin\mathbb{Q}$ y es lógico pensar que no difieren mucho de variables aleatorias independientes (aunque realmente no lo sean). Por el teorema central del límite, la parte real e imaginaria de $S(N)/\pi(N)$ cuando N crece se deberían comportan como una distribución normal lo que hace sospechar que no se puede mejorar la acotación $S(x)\ll N^{1/2+\epsilon}$ y eso ya es demasiado grande como para tratar el caso $k=2$, la conjetura binaria de Goldbach. De hecho es lógico, porque $\mathfrak{m}$ difiere en tan poco de $[0,1]$ que es demasiado optimista pensar que podemos mejorar el valor de $\int_0^1 |S(x)|^2\,dx$.

Ésta es una limitación que aparece en el método del círculo para $k=2$ siempre que $\mathcal{B}$ tenga densidad positiva o densidad $O\big((\log N)^{-r}\big)$, como en el caso de los primos, en los enteros $[1,N]$. Se dice que el método del círculo falla para problemas binarios, aunque este fenómeno va más allá del caso $k=2$. Por ejemplo, tampoco podemos hallar la asintótica del número de representaciones como suma de cuatro cuadrados, porque aunque $k=4\neq 2$, definiendo $x=n_1^2+n_2^2$, $y=n_3^2+n_4^2$ estamos resolviendo el problema binario $x+y=N$ sobre el conjunto de números representables como suma de dos cuadrados, que tiene densidad $O\big((\log N)^{-1/2}\big)$ [32, p.24].

El fracaso del método del círculo en problemas binarios no significa que la contribución heurística de los arcos mayores sea incorrecta, porque al estimar la contribución de los arcos menores mediante una cota superior estamos perdiendo el signo y es de esperar que al integrar $S^2(x)e(-Nx)$ sobre los arcos menores haya mucha cancelación aunque no sepamos medirla.

8.5. El método del círculo de Kloosterman

H.D. Kloosterman en un famoso trabajo [33] superó la limitación teórica del método del círculo en cierto problema estudiando la cancelación entre diferentes arcos menores. Visto de

otra forma, fingió que no existían los arcos menores y probó que había cierto grado de cancelación de errores más que una mera acumulación de ellos.

El problema que trató Kloosterman está relacionado con formas cuadráticas de cuatro variables. Para dar una idea, aplicamos el método del círculo a sumas de cuatro cuadrados partiendo de la formulación con series infinitas análoga a (8.2)

$$r_4(n) = \frac{1}{2\pi i}\int_C \Big(\sum_{n=-\infty}^{\infty} z^{n^2}\Big)^4 \frac{dz}{z^{n+1}}. \tag{8.33}$$

La función entre paréntesis se relaciona tras un cambio de variable con la función clásica

$$\theta(z) = \sum_{n=-\infty}^{\infty} e(n^2 z). \tag{8.34}$$

Una aplicación ingeniosa de la fórmula de sumación de Poisson tras dividir en progresiones aritméticas [32, Lem.20.11], conduce a

$$\theta\big(z+\frac{a}{q}\big) = \sqrt{\frac{i}{2z}} \sum_{n=-\infty}^{\infty} \Big(\frac{1}{q}\sum_{h=1}^{q} e\big(\frac{\overline{a}}{q}(h^2+hn)\big)\Big) e\big(-\frac{n^2}{4q^2 z}\big) \tag{8.35}$$

donde $\overline{a}$ indica el inverso de a módulo q.

Esta fórmula permite estudiar $\theta(w)$ con w cercano a a/q. Si $-1/q^2 z$ tiene parte imaginaria grande (lo que requiere que $q^2 z$ sea pequeño), la serie tiene un decaimiento exponencial y sólo el primer término es relevante. Por otro lado, si $q^2 z$ es grande, hay muchos términos que contribuyen significativamente a la serie anterior y no hay una asintótica clara. En principio deberíamos asociar esos valores a un arco menor. Kloosterman probó que hay cancelación en sumas del tipo

$$\sum_{\substack{a=1\\(a,q)=1}}^{q} e\big(\frac{\overline{a}M + aN}{q}\big), \tag{8.36}$$

hoy llamadas *sumas de Kloosterman*, y esta cancelación es suficiente para conseguir que los términos incontrolados sean absorbidos por el error cuando se suman las contribuciones de los arcos correspondientes a un mismo denominador.

Aunque en la explicación anterior la aparición del inverso parece deberse a (8.35), la división del círculo o del intervalo $[0,1]$ por medio de las fracciones de Farey (fracciones irreducibles con denominador acotado) lleva aparejada intrínsecamente una relación con la distribución de los inversos (véase [32, Prop. 20.7]).

8.6. El método del círculo de Davenport y Heilbronn

Con cierta laxitud se califica como método del círculo a una variante aplicable a desigualdades diofánticas que sólo emplea un arco mayor y no conduce a fórmulas asintóticas. Fue introducida por H. Davenport y H. Heilbronn en [15] al estudiar el número de soluciones $\mathcal{N}$ de

$$\left|\lambda_1 n_1^2 + \lambda_2 n_2^2 + \lambda_3 n_3^2 + \lambda_4 n_4^2 + \lambda_5 n_5^2\right| < \epsilon \qquad \text{con} \quad 1 \le n_i \le N \tag{8.37}$$

donde $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$ son números reales no nulos fijados tales que no todos tienen el mismo signo y algún cociente λ_i/λ_j es irracional. Nótese que ambas condiciones son necesarias para que exista solución con ϵ arbitrariamente pequeño.

Tomando una función $\phi \in C^\infty$, $0 \le \phi \le 1$ y con soporte incluido en el intervalo $[-\epsilon, \epsilon]$, la fórmula de inversión para la transformada de Fourier $\widehat{\phi}$ asegura

$$\mathcal{N} \ge \int_{-\infty}^{\infty} S_1(x) S_2(x) S_3(x) S_4(x) S_5(x) \widehat{\phi}(x)\, dx \qquad \text{con} \quad S_i(x) = \sum_{n=1}^{N} e(n^2 \lambda_i x). \tag{8.38}$$

Si $|x| \ll N^{-2}$ está claro que no es posible ir más allá de la estimación trivial para S_i, resultando de ello

$$\mathcal{N} \ge C(\epsilon) N^5 \cdot N^{-2} + \int_{|x| \gg N^{-2}} S_1(x) S_2(x) S_3(x) S_4(x) S_5(x) \widehat{\phi}(x)\, dx. \tag{8.39}$$

El análisis se puede extender a $N^{-2} \ll x \ll N^{-1}$ donde las fases $f(n) = n^2 \lambda_i x$ verifican $f'(n) \ll 1$ y la fórmula de sumación de Poisson se aplica con éxito para aproximar $S_i(x)$. El problema está en saber qué ocurre más allá del único "arco mayor" $x \ll N^{-1}$. A la larga, el decaimiento de $\widehat{\phi}$ acabará con las dificultades pero éste sólo se manifiesta para $x \gg \epsilon^{-1}$ y el rango $N^{-1} \ll x \ll \epsilon^{-1}$ sigue siendo problemático. En dicho rango, $S_i(x)$ es de hecho grande cuando $\lambda_i x$ toma ciertos valores racionales, pero esta situación no puede darse para todos los i porque entonces λ_i/λ_j sería racional en contra de las hipótesis. Naturalmente los valores tales que $\lambda_i x$ está muy cerca de un racional (de denominador pequeño) son igualmente conflictivos y en el argumento entra la fracción continua (la mejor aproximación racional) de $\lambda_i/\lambda_j \notin \mathbb{Q}$. En cualquier caso, fuera del arco mayor la ventaja se consigue acotando el mínimo de los $|S_i(x)|$ en lugar de centrarse en uno en particular. El razonamiento preciso está en el artículo original [15] y en [14, §20]. Allí se muestra que para el último paso se necesita restringir N a cierta sucesión de valores. Para ellos la contribución del arco mayor es dominante y se tiene $\mathcal{N} \ge C(\epsilon) N^3$, la conclusión es entonces que el límite superior de $\mathcal{N} N^{-3}$ no es nulo.

Apéndice A

Programas

Para elaborar las figuras y tablas se han usado los siguientes programas escritos, salvo una excepción, para ser ejecutados con el *software* matemático SAGE:

Tabla de la página 1

```
def print_table(n):
        print '\t','pi(n) =',prime_pi(n)
        print '\t','pi(n)/(n/log(n))-1 =',((prime_pi(n))/(n/log(n))-1).n()
        print '\t','pi(n)/Li(n)-1 =',((prime_pi(n))/Li(n)-1).n()


for k in range(1,6):
        print '-----------------------------------------------'
        print 2*k
        print_table(10^(2*k))
```

Figuras de la página 4

```
# I cannot control the ticks. Old version?
# import matplotlib
# from matplotlib import ticker


###############
# FIRST PLOT #
###############
P = line([(-3,3),(1,3),(1,-3),(-3,-3)], thickness=3)
P += plot(3,x, -3,1, fill=-3, fillcolor='red', fillalpha=0.2, axes=False)
P += point((0,0), pointsize=80)
```

```
# axes
P += line([(-3.2,0),(1.2,0)],color='black', thickness=1)
P += line([(0,-3.2),(0,3.2)],color='black', thickness=1)

# text
T = text("$s=0$", (-0.4,0.3),fontsize=18)
T += text("$c$", (1.2,-0.3),fontsize=16)
T += text("$R$", (0.2,3.2),fontsize=15)
T += text("$-R$", (0.2,-3.3),fontsize=15)
T += text("$\mathcal{R}_1$", (-3,3.3),fontsize=18)

P += T
P.set_axes_range(-3.3, 1.3, -3.3, 3.3)
show(P,aspect_ratio=1)



###############
# SECOND PLOT #
###############
P = line([(5,3),(1,3),(1,-3),(5,-3)], thickness=3)
P += plot(3,x, 1,5, fill=-3, fillcolor='red', fillalpha=0.2, thickness=2, axes=False)

# axes
P += line([(-0.2,0),(4.1,0)],color='black', thickness=1)
P += line([(0,-3.2),(0,3.2)],color='black', thickness=1)

# text
T = text("$c$", (0.8,0.2),fontsize=16)
T += text("$R$", (0.2,3.2),fontsize=15)
T += text("$-R$", (0.2,-3.3),fontsize=15)
T += text("$\mathcal{R}_2$", (1.4,3.3),fontsize=18)

P += T
P.set_axes_range(-0.1, 5.1, -3.3, 3.3)
show(P,aspect_ratio=1)
```

Figuras de la página 14

```
# I cannot control the ticks. Old version?
# import matplotlib
# from matplotlib import ticker


###############
# FIRST PLOT #
```

```
###############
lb = 0.6
ub = 0.9
f = log(1+ ( exp( (1-x)^(-1) ) - exp( (1-lb)^(-1) ) )/1 )
P =  plot(f(x=ub), x, ub, 1, fill=-f(x=ub), fillcolor='red',
                                    fillalpha=0.2, axes=False, thickness=0)
P += plot(f, x, lb, ub, fill=0, fillcolor='red', fillalpha=0.2, axes=False)
P += plot(-f, x, lb, ub, fill=0, fillcolor='red', fillalpha=0.2, axes=False)
P += point((0.7,0), pointsize=80)

# axes
Q = line([(0.49,0),(1.01,0)],color='black', thickness=1)
Q += line([(0.5,f(x=ub)),(0.5,-f(x=ub))],color='black', thickness=1)

# text
T = text("$1/2$", (0.45,-0.5),fontsize=16)
T += text("$1$", (1.02,-0.5),fontsize=16)

Q += T
show(P+Q, figsize=[3.5,8])

###############
# SECOND PLOT #
###############
lb = 0.6
ub = 0.9
f = log(1+ ( exp( (1-x)^(-1) ) - exp( (1-lb)^(-1) ) )/1 )
P =  plot(f(x=ub), x, ub, 1, fill=-f(x=ub), fillcolor='red',
                                    fillalpha=0.2, axes=False, thickness=0)
P += plot(f, x, lb, ub, fill=0, fillcolor='red', fillalpha=0.2, axes=False)
P += plot(-f, x, lb, ub, fill=0, fillcolor='red', fillalpha=0.2, axes=False)
P += point((0.7,0), pointsize=70)
print f(x=ub)

lb = 0.75
ub = 0.912
f = log(1+ ( exp( (1-x)^(-1) ) - exp( (1-lb)^(-1) ) )/4 )
P +=  plot(f(x=ub), x, ub, 1, fill=-f(x=ub), fillcolor='red',
                                    fillalpha=0.2, axes=False,thickness=0)
P += plot(f, x, lb, ub, fill=0, fillcolor='red', fillalpha=0.2, axes=False)
P += plot(-f, x, lb, ub, fill=0, fillcolor='red', fillalpha=0.2, axes=False)
P += point((0.82,0), pointsize=70)
print f(x=ub)

lb = 0.85
ub = 0.9217
f = log(1+ ( exp( (1-x)^(-1) ) - exp( (1-lb)^(-1) ) )/16 )
P +=  plot(f(x=ub), x, ub, 1, fill=-f(x=ub), fillcolor='red', fillalpha=0.2, axes=False,
thickness=0)
```

```
P += plot(f, x, lb, ub, fill=0, fillcolor='red', fillalpha=0.2, axes=False)
P += plot(-f, x, lb, ub, fill=0, fillcolor='red', fillalpha=0.2, axes=False)
P += point((0.915,0), pointsize=70)
print f(x=ub)


lb = 0.95
ub = 0.950007
f = log(1+ ( exp( (1-x)^(-1) ) - exp( (1-lb)^(-1) ) )/64 )
P +=  plot(f(x=ub), x, ub, 1, fill=-f(x=ub), fillcolor='red', fillalpha=0.2, axes=False,
thickness=0)
P += plot(f, x, lb, ub, fill=0, fillcolor='red', fillalpha=0.2, axes=False)
P += plot(-f, x, lb, ub, fill=0, fillcolor='red', fillalpha=0.2, axes=False)
P += point((0.97,0), pointsize=70)
print f(x=ub)

show(P+Q, figsize=[3.5,8])
```

Tabla de la página 20

```
# PRIMES REPRESENTED 2x^2+7y^2=p, 2x^2+7y^2=2p

def q1(p):
        y=1
        while p>=7*y*y:
                if is_square( (p-7*y*y)/2 ):
                        return True
                y +=2
        return False

def q2(p):
        y=2
while 2*p>=7*y*y:
                if is_square( (2*p-7*y*y)/2 ):
                        return True
                y +=2
        return False

N = 100
shi = [1,9,15,23,25,39]
for  i in range(N):
        for j in shi:
                p = j+56*i
                if is_prime(p):
                        print p,'-->',
                        if q1(p):
                                print 'I'
```

```
elif q2(p):
        print 'II'
else:
        print 'ERROR'
```

Figuras de la página 22

```
# Restricted Fundamental Domain and generators


x= var('x')

#############
# FIGURE 1 #
#############
up_l=3 #upper limit
P = line([(-0.8,0),(0.8,0)], thickness=1)
P += plot(sqrt(1-x^2),x, -1/2,1/2, fill=up_l, fillcolor='blue',
                                    fillalpha=0.2, axes=False,thickness=0)
P += line([(-0.5,up_l),(-0.5,0.866)], thickness=3)
P += line([(0.5,up_l),(0.5,0.866)], thickness=3, linestyle='--')
P += plot(sqrt(1-x^2),x, -1/2,0, axes=False, thickness=3)
P += plot(sqrt(1-x^2),x, 0,1/2, axes=False, thickness=3, linestyle='--')
P += point((0,1), pointsize=50)


# text
T = text("$\mathcal{D}$", (-0.65,up_l),fontsize=18)
T +=text("$\\frac{-1+i\\sqrt{3}}{2}$", (-0.5,0.65), fontsize=18)
T +=text("$\\frac{1+i\\sqrt{3}}{2}$", (0.5,0.65), fontsize=18)
T +=text("$i$", (0,1.2), fontsize=18)

P += T
show(P,aspect_ratio=1)


#############
# FIGURE 2 #
#############
up_l=4 #upper limit
P = line([(-0.5,up_l),(-0.5,0.866)], thickness=3)
P += line([(0.5,up_l),(0.5,0.866)], thickness=3)
P += line([(1.5,up_l),(1.5,0.866)], thickness=3)
P += plot(sqrt(1-x^2),x, -1/2,1/2, fill=up_l, fillcolor='blue',
                                    fillalpha=0.2, axes=False,thickness=3)
P += plot(sqrt(1-x^2),x, -1/2,1/2, axes=False, thickness=3)
P += plot(sqrt(1-(x-1)^2),x, 1/2,3/2, fill=up_l, fillcolor='blue',
```

```
                                        fillalpha=0.2, axes=False, thickness=3)
P += plot(sqrt(1-(x-1)^2),x, 1/2,3/2, axes=False, thickness=3)
P += arrow((0, .4+up_l/2), (1, .4+up_l/2), color='black', width=4)

# text
T = text("$T$", (0.7,.1+up_l/2+.45),fontsize=20, color='black')

P += T
show(P,aspect_ratio=1)


#############
# FIGURE 3 #
#############
up_l=3 #upper limit
P = line([(-1.1,0),(1.1,0)], thickness=1)
P += plot(sqrt(1-x^2),x, -1/2,1/2, fill=up_l, fillcolor='blue',
                                        fillalpha=0.2, axes=False, thickness=3)
P += line([(-0.5,up_l),(-0.5,0.866)], thickness=3)
P += line([(0.5,up_l),(0.5,0.866)], thickness=3)
P += plot(sqrt(1-x^2),x, -1,-1/2, axes=False, thickness=2, linestyle='--')
P += plot(sqrt(1-x^2),x, 1/2,1, axes=False, thickness=2, linestyle='--')

P += arrow((-0.2, 0.5), (0.88,1.47), color='black', width=4)
P += arrow( (0,2),(0, 0.3), color='black', width=4)


# text
T = text("$S$", (-0.18,0.35),fontsize=20, color='black')

P += T
show(P,aspect_ratio=1)
```

Figura de la página 30

```
a = 10
b = 7
f = b*sqrt(1-x^2/a^2)
c = plot(f,x,-a,a, thickness=4)
c += plot(-f,x,-a,a, thickness=4)
c += line([(-a-1,0),(a+1,0)],color='black')+line([(0,-b-1),(0,b+1)],color='black')
for i in range(-a,a+1):
    for j in range(-b,b+1):
        if (i*i/a^2+j*j/b^2)<1:
            c += point((i,j), pointsize=10)
            c=c+line([(i-1/2,j-1/2), (i+1/2,j-1/2), (i+1/2,j+1/2),
                                        (i-1/2,j+1/2), (i-1/2,j-1/2)], thickness=1)
```

```
show(c, aspect_ratio = 1, axes=False)
```

Figura de la página 35

```
# Plot a periodic partition of the unity multplied by cos(pi*x)


x= var('x')
# spline conecting (-3/2,0), (-1/2,1), (1/2,1), (3/2,0)
def m_spline(x):
        if x^2>2.25: return 0
        if x>0: x=-x
        if x>-0.5: return 1
        x += 1.5
        return 6*x^5-15*x^4+10*x^3
        return (3-2*x)*x^2

s_width = 0.12

def phi0(x):
        return cos(pi*x)*m_spline(x/s_width/2)

def phi1(x):
        return cos(pi*x)*m_spline((x-1)/s_width/2)

def m_psi(x):
        if x< -0.5:
                x +=2
        return cos(pi*x)-phi0(x)-phi1(x)

def m_psie(x):
        if x>1: return 0
        return m_psi(x)

P = plot(cos(pi*x), x, -1, 1.5, linestyle='--', thickness=1)
P += plot(phi0, x,-0.5, 1.5, thickness=3, color='red')
P += plot(phi1, x,-0.5, 1.5, thickness=3, color='green')
P += plot(m_psie, x,-1, 1.5, thickness=3, color='blue')
P += plot(m_psi, x,1, 1.5, linestyle='--', thickness=3, color='blue')

# text
T = text("$y=\cos(\pi x)$", (-0.51,-0.9),fontsize=16)
T += text("$\phi_0$", (0.2,1),fontsize=18)
T += text("$\phi_1$", (1.2,-1),fontsize=18)
T += text("$\psi$", (-0.9,0.15),fontsize=18)
```

```
P += T

show(P, figsize=[8,3])
```

Figuras de la página 38

```
# Plot exponential sums

def plot_exps(N,f):
        L = []
        x, y = 0, 0
        for k in range(1,N+1):
        x += cos( (2*pi*f(k)).n() )
        y += sin( (2*pi*f(k)).n() )
        L.append( [x,y] )

        P = list_plot(L, pointsize=10)
        show(P,aspect_ratio=1)

def plot_numbers(alp,N):
        L = [ (sqrt(n)/pi/alp*sin(2*pi*alp*sqrt(n)),
        sqrt(n)/pi/alp*cos(2*pi*alp*sqrt(n))) for n in range(N)]
        T =line(L)
#       T +=list_plot(L)
        for n in range(len(L)):
                T += text(n, L[n],fontsize=12)
        show(T, aspect_ratio=1, axes=False)


#plot_numbers(0.5,200)

def f1(n):
        return n^(3/2)

def f2(n):
        return n^(1/2)

def f3(n):
        return n^(1/2)/2


N = 300
N = 500
#plot_exps(N,f3)
#plot_exps(N,f2)
plot_exps(4*N,f1)
```

Figura de la página 42

```
# Exponent pairs

def A(p):
        return ( p[0]/(p[0]+1)/2,  (p[0]+p[1]+1)/(p[0]+1)/2 )

def B(p):
        return ( p[1]-1/2, p[0]+1/2 )

L= []
# L.append( [ "B", B([0,1]) ] )
L.append( [ "trivial", A([0,1]) ] )
L.append( [ "$AB$", A(B([0,1])) ] )
L.append( [ "$A^2B$", A(A(B([0,1]))) ] )
L.append( [ "$ABA^2B$", A(B(A(A(B([0,1])))))  ] )
L.append( [ "$A^3B$", A(A(A(B([0,1])))) ] )
L.append( [ "$A^4B$", A(A(A(A(B([0,1])))))  ] )
L.append( [ "$A^2BA^2B$", A(A(B(A(A(B([0,1])))))) ] )
L.append( [ "$ABABA^2B$", A(B(A(B(A(A(B([0,1])))))))  ] )

l_points =[]
for item in L:
        l_points.append( item[1] )

P = list_plot(l_points,  pointsize=25)
P += line(sorted(l_points),  thickness=1)
for item in L:
        P += text(item[0], (item[1][0]+0.0025, item[1][1]+0.015),fontsize=15,
                                                        horizontal_alignment='left')

P.set_axes_range(-0.01, 0.18, 2/3, 1.01)
show(P)
```

Figura de la página 49

```
 # This file is automatically created by goldbach.c

 r2 = [
         10804,16400,8320,9742,18760,11902,8750,16266,8084,8122,
         ..................................................
         40112,15874,17518,29044,15064,19518,29342,14494,15064,43700]

 N = 2*len(r2)
```

```
P = list_plot(zip(range(N,2*N,2), r2), pointsize=1)

show(P, figsize=[15.8,10])

show(P, figsize=[31.6,20])
```

```
/*
This is goldbach.c
# r_2(n) for Goldbach in [N,2N)
*/

/*
run  a.out > plot_goldbach.sage
*/

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


/*
primes.h = Table of prime numbers
*/
#include "primes.h"

#define N 1000000

int f_pi(int n){
        int i=0, k=0;
        while( primes[i]<= n){
                ++k;
                ++i;
        }
        return k;
}


int main(int argc, char **argv)
{
        int i,j,n;
        int r2[N]={0};
        int limit = f_pi(2*N);
        for(i=1;i<limit; ++i){
                for(j=i;j<limit; ++j){
                        n = primes[i] + primes[j] -N;
```

```
                        if( (n>=0)&& (n<N) ){
                                r2[n]+=2;
                        }
                }
        }

//        for(i=0;i<N; i+=2)
//                printf("r2[%d]=%d\n",i+N,r2[i]);

        printf("# This file is automatically created by goldbach.c\n\n");
        printf("r2 = [");

        for(i=0;i<N-2; i+=2){
                if ( (i%20)==0 ) printf("\n\t");
                printf("%d,",r2[i]);
        }
        printf("%d]\n\n",r2[N-2]);

        printf("N = 2*len(r2)\n\n");
        printf("P = list_plot(zip(range(N,2*N,2), r2), pointsize=1)\n\n");
        printf("show(P, figsize=[15.8,10])\n\n");
        printf("show(P, figsize=[31.6,20])\n\n");

        return EXIT_SUCCESS;
}
```

Figuras de la página 62

```
#
# Plot an approximation of Beurling Selberg function in [-a,a]
#

N = 1000 # number of knots
a = 3
L = []
for n in range(-N,N):
        z = (10^-6+a*n/N ).n()
        s = (2*z+1)/z/z
        for k in range(1,101):
                s += (1/(z-k)^2-1/(z+k)^2)
        s *= (sin(pi*z)/pi)^2
        L.append(s.n())

P = list_plot( zip(srange(-a,a,a/N),L),  plotjoined=True, thickness=3)
P += line([(0,1),(a,1)], thickness=1.5, color='green')
P += line([(-a,-1),(0,-1)], thickness=1.5, color='green')
P += point((0,0), color='green', pointsize=20)
```

```
show(P, figsize=[8,4])
# show(P)


# approximate beurling
def app_beurling(x):
        # scaling
        x = 2*N*(x+a)/2/a
        if x<0: return -1
        if x>len(L)-1: return 1
        n = floor(x)
        return L[n]*(n+1-x)+L[n+1]*(x-n)

# approximate characteristic function of [-1/2,1/2] with \delta = 1/4
def app_c_fp(x):
        a1 = -1/2
        a2 = 1/2
        delta = 1/4
        return (app_beurling((x-a1)/delta) + app_beurling((a2-x)/delta) )/2

def app_c_fm(x):
        a1 = -1/2
        a2 = 1/2
        delta = 1/4
        return -(app_beurling((a1-x)/delta) + app_beurling((x-a2)/delta) )/2


P = plot(app_c_fp, x, -1.5,1.5, thickness=2)
P += plot(app_c_fm, x, -1.5,1.5, color= 'red', thickness=2)
P += line([(-1.5,0),(-1/2,0),(-1/2,1),(1/2,1),(1/2,0),(1.5,0)],
                                                thickness=1.5, color='green')
show( P, figsize=[8,4])
```

Figuras de la página 68

```
    # FIRST FIGURE

    def rota( x,y ):
    alpha = -15*pi/180
    return (cos(alpha)*x-sin(alpha)*y, sin(alpha)*x+cos(alpha)*y)


    # vector
    corr = 0.1
    P = arrow((0,0), rota(4+corr,6+corr), color='red', width=4, arrowsize=10)

    # axes
```

```
P += line([(0,0),rota(0,10)], thickness=2)
P += line([(0,0),rota(10,0)], thickness=2)
P += line([(0,0),rota(-3.5,-3.5)], thickness=2)

# projections
P += line([(0,0),rota(0,8)], color='black', thickness=4)
P += line([(0,0),rota(6,0)], color='black', thickness=4)
P += line([(0,0),rota(-2,-2)], color='black', thickness=4)

#frame
P += line([rota(0,8),rota(6,8),rota(4,6),rota(-2,6),rota(0,8)], color='black',
                                                    linestyle='--', thickness=2)
P += line([rota(4,6),rota(4,-2),rota(6,0),rota(6,8)], color='black',
                                                    linestyle='--', thickness=2)
P += line([rota(-2,6),rota(-2,-2),rota(4,-2)], color='black',
                                                    linestyle='--', thickness=2)

# points
P += point([rota(0,8)], color='black', size=90)
P += point([rota(6,0)], color='black', size=90)
P += point([rota(-2,-2)], color='black', size=90)

# text
P += text("$\\vec{n}$", rota(2,4.3),fontsize=30, color='black')

show(P,aspect_ratio=1, axes=False)

# SECOND FIGURE

# vector
corr = .1
P = arrow((0,0), (7+corr,1), color='red', width=4, arrowsize=10)

# axes
# P += line([(0,0),(10,0)], thickness=2)
P += line([(0,0),(10,0.5)], thickness=2)
P += line([(0,0),(90/sqrt(82),-10/sqrt(82))], thickness=2)
P += line([(0,0),(70/sqrt(53),20/sqrt(53))], thickness=2)


# projections
#P += line([(0,0),(7,0)], color='black', thickness=4)
P += line([(0,0),(70.5*(10+corr/2.11)/100.25,70.5*(0.5+corr/2.11)/100.25)],
                                                    color='black',thickness=4)
P += line([(0,0),(51*7/53,51*2/53)], color='black', thickness=4)
P += line([(0,0),(62*9/82,62*(-1)/82)], color='black', thickness=4)


#frame
```

```
#P += line([(7,0),(7,1)], color='black', linestyle='--', thickness=2)
P += line([(70.5*(10+corr/2.11)/100.25,70.5*(0.5+corr/2.11)/100.25),(7,1)],
                                        color='black', linestyle='--', thickness=2)
P += line([(51*7/53,51*2/53),(7,1)], color='black', linestyle='--', thickness=2)
P += line([(62*9/82,62*(-1)/82),(7,1)], color='black', linestyle='--', thickness=2)


# points
P += point([(70.5*(10+corr/2.11)/100.25,70.5*(0.5+corr/2.11)/100.25)],
                                                    color='black', size=90)
P += point([(51*7/53,51*2/53)], color='black', size=90)
P += point([(62*9/82,62*(-1)/82)], color='black', size=90)

# text
P += text("$\\vec{n}$", (7.6,1.1),fontsize=30, color='black')

show(P,aspect_ratio=1, axes=False)
```

Figuras de la página 82

```
N = 200
prim = prime_range(1,N)
delta = 1/log(N)
Pdir = plot(dir, x, -4/N, 4/N, color='red', thickness=2)
delta = 1
Pdir2 = plot(dir, x, -4/N, 4/N, color='red', thickness=2)

def s(x):
    su = 0.0
    for p in prim:
        su += cos(2*pi*p*x)
    return su

def s2(x):
    su = 0.0
    for p in prim:
        su += log(p)*cos(2*pi*p*x)
    return su

# Dirichlet kernel from 1 to N
def dir(x):
    return delta*(sin(2*pi*(N+1/2)*x) - sin(pi*x) )/2/sin(pi*x)


# FIRST PLOT
t = cputime()
P = plot(s, x, -4/N, 4/N, thickness=2)
```

```
P += Pdir
te = " $\\sum_{p\\leq "+str(N)+"}\\cos(2\\pi px)$"
P += text(te, (-5/N,0.9*s(0.001)), fontsize=14, horizontal_alignment="left")
te = " $(\\log "+str(N)+")^{-1}\\sum_{n\\leq "+str(N)+"}\\cos(2\\pi nx)$"
P += text(te, (-5/N,0.6*s(0.001)), fontsize=14,
                                    horizontal_alignment="left", color='red')
P.show()
print cputime(t)


# SECOND PLOT
t = cputime()
P = plot(s2, x, -4/N, 4/N, thickness=2)
P += Pdir2
te = " $\\sum_{p\\leq "+str(N)+"}(\\log\ p)\\cos(2\\pi px)$"
P += text(te, (-5/N,0.9*s2(0.001)), fontsize=14, horizontal_alignment="left")
te = "\n\n $\\sum_{n\\leq "+str(N)+"}\\cos(2\\pi nx)$"
P += text(te, (-5/N,0.6*s2(0.001)), fontsize=14,
                                    horizontal_alignment="left", color='red')
P.show()
print cputime(t)
```

Bibliografía

[1] N. C. Ankeny. The least quadratic non residue. *Ann. of Math. (2)*, 55:65–72, 1952.

[2] R. C. Baker. Metric number theory and the large sieve. *J. London Math. Soc. (2)*, 24(1):34–40, 1981.

[3] E. Bombieri. *Le grand crible dans la théorie analytique des nombres*. Société Mathématique de France, Paris, 1974. Avec une sommaire en anglais, Astérisque, No. 18.

[4] A. I. Borevich and I. R. Shafarevich. *Number theory*. Translated from the Russian by N. Greenleaf. Pure and Applied Mathematics, Vol. 20. Academic Press, New York, 1966.

[5] J. H. Bruinier, G. van der Geer, G. Harder, and D. Zagier. *The 1-2-3 of modular forms*. Universitext. Springer-Verlag, Berlin, 2008. Lectures from the Summer School on Modular Forms and their Applications held in Nordfjordeid, June 2004, Edited by K. Ranestad.

[6] D. A. Buell. *Binary quadratic forms*. Springer-Verlag, New York, 1989. Classical theory and modern computations.

[7] D. A. Burgess. The distribution of quadratic residues and non-residues. *Mathematika*, 4:106–112, 1957.

[8] F. Chamizo. Métodos analíticos en teoría de números. `http://www.uam.es/fernando.chamizo/libreria/fich/APtenumav02.pdf`, 2002.

[9] J. Cilleruelo and A. Córdoba. *La teoría de los números*. Biblioteca Mondadori. Mondadori España, Madrid, 1992.

[10] H. Cohn. *Advanced number theory*. Dover Publications Inc., New York, 1980. Reprint of *A second course in number theory*, 1962, Dover Books on Advanced Mathematics.

[11] 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.

[12] D. A. Cox. *Primes of the form $x^2 + ny^2$*. A Wiley-Interscience Publication. John Wiley & Sons Inc., New York, 1989. Fermat, class field theory and complex multiplication.

[13] H. Davenport. *Multiplicative number theory*, volume 74 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2000. Revised and with a preface by H. L. Montgomery.

[14] H. Davenport. *Analytic methods for Diophantine equations and Diophantine inequalities*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, second edition, 2005. With a foreword by R. C. Vaughan, D. R. Heath-Brown and D. E. Freeman, Edited and prepared for publication by T. D. Browning.

[15] H. Davenport and H. Heilbronn. On indefinite quadratic forms in five variables. *J. London Math. Soc.*, 21:185–193, 1946.

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

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

[18] P. Erdös and P. Turán. On a problem in the theory of uniform distribution. I. *Nederl. Akad. Wetensch., Proc.*, 51:1146–1154 = Indagationes Math. 10, 370–378 (1948), 1948.

[19] J. Friedlander and A. Granville. Limitations to the equi-distribution of primes. III. *Compositio Math.*, 81(1):19–32, 1992.

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

[21] J. B. Friedlander. Producing prime numbers via sieve methods. In *Analytic number theory*, volume 1891 of *Lecture Notes in Math.*, pages 1–49. Springer, Berlin, 2006.

[22] C. F. Gauss. *Disquisitiones arithmeticae*. Springer-Verlag, New York, 1986. Translated and with a preface by A. A. Clarke, Revised by W. C. Waterhouse, C. Greither and A. W. Grootendorst and with a preface by Waterhouse.

[23] A. O. Gel′fond and Yu. V. Linnik. *Elementary methods in the analytic theory of numbers*. Translated from the Russian by D. E. Brown. Translation edited by I. N. Sneddon. International Series of Monographs in Pure and Applied Mathematics, Vol. 92. Pergamon Press, Oxford, 1966.

[24] S. W. Graham and G. Kolesnik. *van der Corput's method of exponential sums*, volume 126 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1991.

[25] H. Halberstam and H.-E. Richert. *Sieve methods*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], London-New York, 1974. London Mathematical Society Monographs, No. 4.

[26] G. H. Hardy and E. M. Wright. *An introduction to the theory of numbers*. Oxford University Press, Oxford, sixth edition, 2008. Revised by D. R. Heath-Brown and J. H. Silverman.

[27] D. R. Heath-Brown. Lectures on sieves. In *Proceedings of the Session in Analytic Number Theory and Diophantine Equations*, volume 360 of *Bonner Math. Schriften*, page 50, Bonn, 2003. Univ. Bonn.

[28] L. K. Hua. *Introduction to number theory*. Springer-Verlag, Berlin, 1982. Translated from the Chinese by P. Shiu.

[29] M. N. Huxley. On the difference between consecutive primes. *Invent. Math.*, 15:164–170, 1972.

[30] M. N. Huxley. *Area, lattice points, and exponential sums*, volume 13 of *London Mathematical Society Monographs. New Series*. The Clarendon Press Oxford University Press, New York, 1996. Oxford Science Publications.

[31] 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)].

[32] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[33] H. D. Kloosterman. On the representation of numbers in the form $ax^2 + by^2 + cz^2 + dt^2$. *Acta Math.*, 49:407–464, 1926.

[34] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences*. Wiley-Interscience [John Wiley & Sons], New York, 1974. Pure and Applied Mathematics.

[35] J. E. Littlewood. The quickest proof of the prime number theorem. *Acta Arith.*, 18:83–86, 1971.

[36] B. Mazur. On the passage from local to global in number theory. *Bull. Amer. Math. Soc. (N.S.)*, 29(1):14–50, 1993.

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

[38] H. L. Montgomery. *Ten lectures on the interface between analytic number theory and harmonic analysis*, volume 84 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1994.

[39] H. L. Montgomery and R. C. Vaughan. The large sieve. *Mathematika*, 20:119–134, 1973.

[40] M. R. Murty. *Problems in analytic number theory*, volume 206 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2008. Readings in Mathematics.

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

[42] P. Pollack. *Not always buried deep*. American Mathematical Society, Providence, RI, 2009. A second course in elementary number theory.

[43] H. Rademacher. On the expansion of the partition function in a series. *Ann. of Math. (2)*, 44:416–422, 1943.

[44] O. Ramaré. *Arithmetical aspects of the large sieve inequality*, volume 1 of *Harish-Chandra Research Institute Lecture Notes*. Hindustan Book Agency, New Delhi, 2009. With the collaboration of D. S. Ramana.

[45] W. Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.

[46] W. Schmidt. *Equations over finite fields: an elementary approach*. Kendrick Press, Heber City, UT, second edition, 2004.

[47] A. Selberg. The general sieve-method and its place in prime number theory. In *Proceedings of the International Congress of Mathematicians, Cambridge, Mass., 1950, vol. 1*, pages 286–292, Providence, R. I., 1952. Amer. Math. Soc.

[48] A. Selberg. *Collected papers. Vol. II*. Springer-Verlag, Berlin, 1991. With a foreword by K. Chandrasekharan.

[49] J.-P. Serre. *A course in arithmetic*. Springer-Verlag, New York, 1973. Translated from the French, Graduate Texts in Mathematics, No. 7.

[50] D. E. Smith. *A source book in mathematics*. 2 vols. Dover Publications Inc., New York, 1959.

[51] M. Spivak. *Cálculo infinitesimal. Tomos I y II*. Editorial Reverté, 1984.

[52] J. D. Vaaler. Some extremal functions in Fourier analysis. *Bull. Amer. Math. Soc. (N.S.)*, 12(2):183–216, 1985.

[53] R. C. Vaughan. *The Hardy-Littlewood method*, volume 125 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, second edition, 1997.

[54] I. M. Vinogradov. *The method of trigonometrical sums in the theory of numbers*. Dover Publications Inc., Mineola, NY, 2004. Translated from the Russian, revised and annotated by K. F. Roth and Anne Davenport, Reprint of the 1954 translation.

Índice alfabético