El diablo de los números

Sección a cargo de

Javier Cilleruelo Mateo

La resolución de la conjetura ternaria de Goldbach es uno de esos logros matemáticos que hacen historia. Harald Helfgott, peruano afincado en París, ha tenido la amabilidad de contarnos las estrategias que le han llevado a la resolución de la conjetura. Espero que los lectores disfruten, como yo lo he hecho, de su estilo directo, transparente y con medidas dosis de humor.

No se me ocurre mejor artículo para despedirme de los lectores como responsable de esta sección, que he intentado llevar con entusiasmo y me ha dado muchas satisfacciones científicas y personales.

La conjetura débil de Goldbach

por

Harald Helfgott

1. Introducción

Leonhard Euler, uno de los matemáticos más importantes del siglo XVIII y de todos los tiempos, y su amigo, el amateur y polímata Christian Goldbach, tuvieron una regular y abundante correspondencia. En su célebre carta del 7 de junio de 1742, Goldbach en verdad dio un enunciado de apariencia un tanto confusa, o por lo menos poco familiar («todo número puede ser descompuesto en una suma de un número arbitrario de primos»). Euler rápidamente la redujo a la conjetura siguiente, que, según dijo, Goldbach ya le había expuesto anteriormente:

> «Todo entero positivo puede expresarse como suma de, como mucho, tres números primos».

Nosotros diríamos ahora «todo entero positivo mayor que 5», ya que en la actualidad no se considera al 1 número primo. Por otro lado, la conjetura se ha dividido, de manera natural, en dos:

- La conjetura *débil* (o ternaria) de Goldbach, que dice que todo entero impar mayor que 5 puede escribirse como suma de tres números primos, y

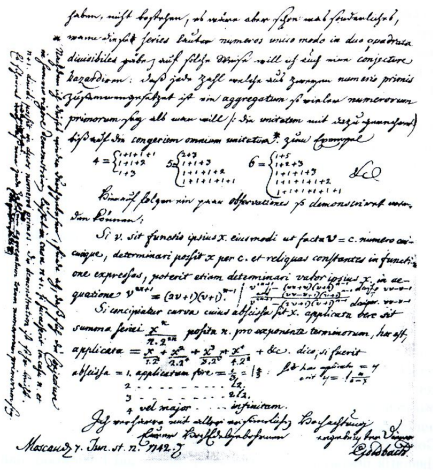

Figura 1: Euler y la carta de Goldbach.

- La conjetura *fuerte* (o binaria) de Goldbach, que afirma que todo entero par mayor que 2 puede expresarse como suma de dos números primos.

Como sus nombres sugieren, la conjetura fuerte implica a la débil (fácilmente: reste 3 a su número impar y después exprese $n-3$ como suma de dos primos).

Se puede consultar [8, Ch. XVIII] para conocer la historia temprana de la conjetura. En resumen, parece que Waring volvió a proponer por su cuenta la conjetura débil a finales del siglo XVIII, y que en el siglo XIX se hizo algo de trabajo computacional (comprobando la conjetura para los números enteros hasta $2 \cdot 10^6$ a mano), pero poco progreso de verdad.

La conjetura fuerte sigue fuera de nuestro alcance. Hace unos meses —mi *preprint* [12] apareció el 13 de mayo de 2013— probé la conjetura débil de Goldbach.

Los cimientos de la prueba descansan en los avances logrados a principios del siglo XX por Hardy, Littlewood y Vinogradov. En 1937, Vinogradov probó [29] que la conjetura es cierta para todos los números impares mayores que alguna constante C. Hardy y Littlewood [10] ya lo habían demostrado unos años antes, pero bajo la suposición de que la *hipótesis generalizada de Riemann* fuera cierta; hablaremos de esto más adelante. Desde entonces, la constante C ha sido especificada y gradualmente mejorada, pero el mejor valor (esto es, el más pequeño) de C del que se disponía era $C = e^{3100} > 10^{1346}$ (Liu-Wang [17]), lo cual era, de lejos, demasiado grande. Incluso $C = 10^{100}$ sería demasiado: como 10^{100} es más grande que el producto del número estimado de partículas subatómicas del universo por el número de segundos desde el Big Bang, no habría ninguna esperanza de comprobar cada caso hasta 10^{100} por ordenador (aun asumiendo que uno fuera un dictador alienígena usando el universo entero como una computadora muy altamente paralela).

Yo reduje C a 10^{29} (y podría bajarlo más si fuera necesario). D. Platt y yo habíamos comprobado la conjetura para todos los números impares hasta $8.8 \cdot 10^{30}$

por ordenador (y podríamos haber llegado más lejos), así que este fue el final de la historia.

* * *

Es justo que repasemos algunos de los principales avances entre la época de Vinogradov y la nuestra. En 1933, Schnirelmann probó [24] que todo entero $n > 1$ puede escribirse como la suma de, a lo más, K primos, donde K era una constante no especificada. Se trata de uno de los trabajos precursores de la combinatoria aditiva. En 1969, Klimov dio un primer valor para la constante ($K = 6 \cdot 10^9$); luego la mejoró a $K = 115$ (con G. Z. Piltay y T. A. Sheptickaja) y $K = 55$. Siguieron resultados de Vaughan [27] ($K = 27$), Deshouillers [7] ($K = 26$) y Riesel-Vaughan [23] ($K = 19$).

Ramaré mostró en 1995 que todo par $n > 1$ es la suma de a lo más 6 primos [21]; sus métodos se inscriben más en la tradición de Vinogradov que en la de Schnirelmann. Por último, en 2012, Tao probó [25] que todo impar $n > 1$ es la suma de a lo más 5 primos.

Hubo otras líneas de aproximación a la conjetura fuerte. Estermann [9] demostró, usando ideas cercanas a las de Vinogradov, que casi todo número par (es decir, un conjunto de densidad 1 en los pares) puede ser escrito como la suma de dos números primos. En 1973, J.-R. Chen llegó a probar [1] que todo par más grande que una cierta constante puede escribirse como la suma de un primo y el producto de dos primos. Por cierto, el mismo Chen, junto con T.-Z. Wang, es responsable de las mejoras cotas para C (en Goldbach ternario) antes de Liu y Wang: $C = \exp(\exp(11.503)) < 4 \cdot 10^{43000}$ [2] y $C = \exp(\exp(9.715)) < 6 \cdot 10^{7193}$ [3].

* * *

¿Cuáles son los elementos de la demostración? Demos primero un paso atrás y echemos una mirada a la estructura general del «método del círculo», introducido por Hardy y Littlewood.

2. El método del círculo: análisis de Fourier en los enteros

El análisis de Fourier es algo que usamos cada vez que sintonizamos una radio: hay una señal, y la descomponemos en sus componentes en diferentes frecuencias. En términos matemáticos: se nos da una función $f : \mathbb{R} \to \mathbb{C}$ (esto es, una función de una sola variable real; en el caso de la radio la variable es el tiempo) y definimos la transformada de Fourier $\widehat{f} : \mathbb{R} \to \mathbb{C}$ como $\widehat{f}(r) = \int_{\mathbb{R}} f(x) e(-xr)\, dx$, donde escribimos $e(t)$ por $e^{2\pi i t}$. Entonces, como se aprende en cualquier curso de análisis de Fourier, $f(x) = \int_{\mathbb{R}} \widehat{f}(r) e(xr)\, dr$, siempre que f decaiga suficientemente rápido y se comporte bien. (Esta es la «fórmula de inversión de Fourier».)

En otras palabras, $x \mapsto f(x)$ ha sido descompuesta como una suma de funciones exponenciales (complejas), con la función exponencial (compleja) $x \mapsto e(xr)$ presente con intensidad $\widehat{f}(r)$. (Esto es equivalente a una descomposición en ondas sinusoidales $x \mapsto \operatorname{sen}(2\pi x r)$ y $x \mapsto \cos(2\pi x r)$, ya que $e^{iz} = \cos(z) + i \operatorname{sen}(z)$.) Volviendo al

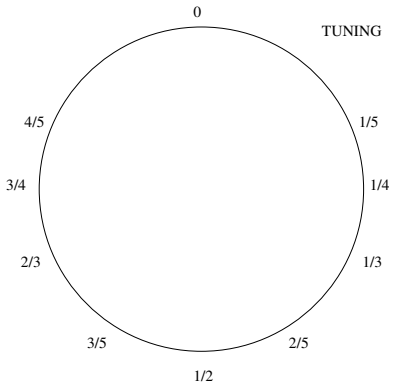

Figura 2: El dial de la radio de un verdadero especialista en teoría de números.

ejemplo de la radio: $\widehat{f}(r)$ es grande cuando r está cerca de la frecuencia de alguna estación de radio, y pequeño en otro caso. (Lo que la radio recibe es una superposición f de lo que transmiten todas las estaciones; el trabajo del receptor de radio consiste precisamente en descifrar la contribución de las frecuencias r alrededor de un r_0 dado.)

Podemos hacer lo mismo si f es una función que va de los enteros $\mathbb{Z}$ a $\mathbb{C}$. De hecho, las cosas son ahora más simples —se llega a definir $\widehat{f}$ como una suma en vez de como una integral: $\widehat{f}(\alpha) = \sum_n f(n)e(-\alpha n)$—. Algo interesante aquí es que $\widehat{f}(\alpha)$ no cambia en absoluto si sumamos 1, o cualquier otro entero m, a α. Esto es así porque, para m entero,

$$e(-(\alpha+m)n) = e^{-2\pi i\alpha n}(e^{-2\pi i})^{mn} = e(-\alpha n)\cdot 1^{-mn} = e(-\alpha n).$$

(Gracias de nuevo, Euler.) Por tanto, podemos restringir α al intervalo $[0,1]$ —o, de forma más abstracta, podemos pensar en α como un elemento del cociente $\mathbb{R}/\mathbb{Z}$—. Topológicamente, $\mathbb{R}/\mathbb{Z}$ es un círculo —lo cual es lo mismo que decir que, como no importa si sumamos o restamos 1 a nuestra frecuencia, podríamos también hacer que la aguja del dial de nuestra radio recorra un círculo marcado con números de 0 hasta 1, en vez de que se deslice en (un segmento de) la recta real (como en la radio sobre mi mesa)—. De allí viene el nombre de *método del círculo*.

La descomposición de f ahora se ve como sigue: $f(n) = \int_0^1 \widehat{f}(\alpha)e(\alpha n)\,d\alpha$, a condición de que f decaiga suficientemente rápido.

¿Por qué nos importa todo esto? La transformada de Fourier es útil inmediatamente si estamos trabajando en problemas aditivos, como las conjeturas de Goldbach. La razón detrás de esto es que la transformada de una convolución es igual al producto de las transformadas: $\widehat{f*g} = \widehat{f}\cdot\widehat{g}$. Recordemos que la convolución (aditiva) de $f, g : \mathbb{Z} \to \mathbb{C}$ está definida por

$$(f*g)(n) = \sum_{m\in\mathbb{Z}} f(m)g(n-m).$$

Figura 3: El dúo Hardy-Littlewood e I. M. Vinogradov.

Podemos ver entonces que $(f * g)(n)$ puede ser distinto de cero sólo si n puede ser escrito como $n = m_1 + m_2$ para algunos m_1, m_2 tales que $f(m_1)$ y $g(m_2)$ sean distintos de cero. De forma similar, $(f * g * h)(n)$ puede ser distinto de cero sólo si n puede escribirse como $n = m_1 + m_2 + m_3$ para algunos m_1, m_2 y m_3 tales que $f(m_1)$, $g(m_2)$ y $h(m_3)$ sean todos distintos de cero. Ello sugiere que, para estudiar la conjetura ternaria de Goldbach, sea conveniente elegir f, g, h de forma que tomen valores distintos de cero sólo en los primos. Por ejemplo, si $f = g = h$ es la función que vale 1 en los primos y 0 en el resto, es claro que $(f*g*h)(n)$ coincide exactamente con el número de representaciones de n como suma de tres primos.

Hardy y Littlewood definieron $f(n) = g(n) = h(n) = 0$ para n compuesto (o cero o negativo) y $f(n) = g(n) = h(n) = (\log n)e^{-n/x}$ para n primo (donde x es un parámetro que será fijado más adelante). Aquí el factor $e^{-n/x}$ está para proporcionar «decaimiento rápido», por lo que todo converge. Como veremos más adelante, la elección de Hardy y Littlewood de $e^{-n/x}$ (en vez de alguna otra función de decaimiento rápido) es de hecho muy inteligente, aunque no la mejor posible. El término $\log n$ aparece por razones técnicas (básicamente, resulta que tiene sentido ponderar un primo p por $\log p$ porque aproximadamente uno de cada $\log p$ enteros alrededor de p es primo).

Vemos que $(f * g * h)(n) \neq 0$ si y sólo si n puede ser escrito como la suma de tres primos. Nuestra tarea es, entonces, mostrar que $(f*g*h)(n)$ (es decir, $f*f*f(n)$) es distinto de cero para todo n impar mayor que una constante. Como la transformada de una convolución es igual al producto de las transformadas, tenemos que

$$(f * g * h)(n) = \int_0^1 \widehat{f * g * h}(\alpha) e(\alpha n)\, d\alpha = \int_0^1 (\widehat{f}\,\widehat{g}\,\widehat{h})(\alpha) e(\alpha n)\, d\alpha.$$

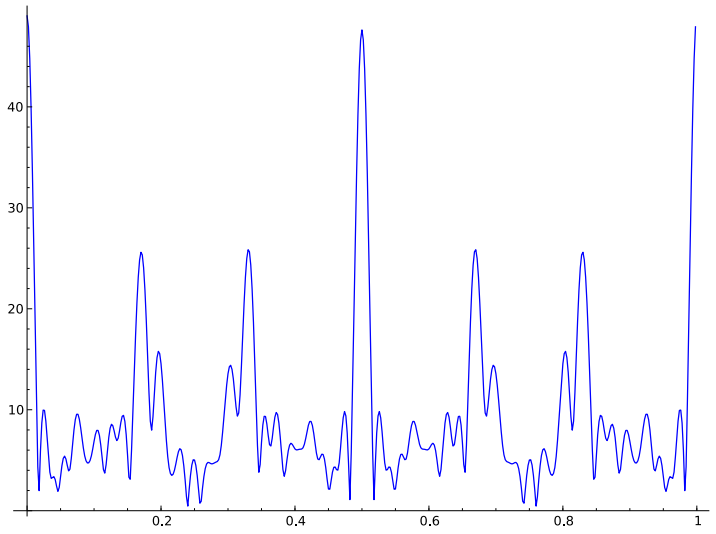

Figura 4: Gráfica de $|f(\alpha)|$.

Nuestro trabajo es, por lo tanto, mostrar que la integral

$$\int_0^1 (\widehat{f}\,\widehat{g}\,\widehat{h})(\alpha)e(\alpha n)\,d\alpha = \int_0^1 (\widehat{f}(\alpha))^3 e(\alpha n)\,d\alpha$$

es distinta de cero.

Resulta que $\widehat{f}(\alpha)$ es particularmente grande cuando α está cerca de un racional con denominador pequeño; es como si realmente hubiera estaciones de radio transmitiendo las frecuencias (de denominador pequeño) marcadas en el dial dibujado arriba —cuando la aguja del dial está cerca de una de ellas, hay una señal fuerte y clara (i.e., la intensidad $\widehat{f}(\alpha)$ es grande), y cuando estamos lejos de todas ellas, podemos escuchar sólo un leve zumbido—. En la figura 4 se representa el valor de $|f(\alpha)|$ cuando $f(\alpha) = \sum_{p\leq 60} \log p\, e(\alpha p)$. Esto sugiere la siguiente estrategia: calcular $\widehat{f}(\alpha)$ para todo α dentro de arcos pequeños alrededor de los racionales con denominadores pequeños (los arcos mayores —llamados así porque aportan una mayor contribución, a pesar de ser pequeños—); acotar $\widehat{f}(\alpha)$ para α fuera de los arcos mayores (todo lo que hay fuera de los arcos mayores se denomina arcos menores); por último, mostrar que la contribución de los arcos menores a la integral es menor, en valor absoluto, que la contribución de los arcos mayores, forzando así que la integral $\int_0^1 (\widehat{f}(\alpha))^3 e(\alpha n)\,d\alpha$ sea distinta de cero.

Es a esta estrategia general a la que se denomina *método del círculo*. Hardy y Littlewood la introdujeron para tratar una amplia variedad de problemas aditivos; por ejemplo, fue también parte de su enfoque sobre el problema de Waring, que trata de enteros que son suma de potencias k-ésimas de enteros. El método fue desarrollado plenamente por Vinogradov, quien fue el primero en dar buenas cotas

incondicionales para $\widehat{f}(\alpha)$ cuando α está en los arcos menores (un logro considerado muy notable en su tiempo). El método del círculo es también mi estrategia general: lo que he hecho es dar estimaciones mucho mejores para los arcos mayores y menores que las que teníamos previamente, para unas funciones f, g y h elegidas con mucho cuidado.

(Incidentalmente: si quisiéramos tratar la conjetura binaria, o fuerte, de Goldbach con el método del círculo nos toparíamos pronto con un obstáculo mayúsculo: el «ruido» procedente de los arcos menores abruma la contribución de los arcos mayores. Ver la exposición de este problema en el artículo [26], en el blog de T. Tao.)

3. Funciones L de Dirichlet y sus ceros

Antes de que podamos comenzar a trabajar en los arcos mayores, necesitamos hablar sobre las funciones L. La reina de estas funciones es la función zeta, $\zeta(s)$, estudiada para s complejo por Riemann, cuyo nombre ahora lleva. Está dada por

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

cuando la parte real $\Re(s)$ de s es mayor que 1. Cuando $\Re(s) \leq 1$, la serie diverge, pero la función puede definirse (de forma única) por continuación analítica (y esto puede hacerse explícitamente usando, por ejemplo, Euler-Maclaurin, como en [5, p. 32]), con un polo en $s = 1$.

La conexión entre la función $\zeta(s)$ (pero con $s > 1$ y real) y los primos había sido descubierta anteriormente por Euler, quien demostró la notable identidad

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

donde el producto se extiende sobre todos los primos. Euler dedujo fácilmente de esta identidad que la suma de los inversos de los primos es infinita. Se debe a Riemann, sin embargo, la generalización de la función $\zeta(s)$ al plano complejo y el esbozo de la estrategia para demostrar lo que hoy conocemos como el *teorema de los números primos*, felizmente demostrado por Hadamard y de la Vallée-Poussin:

$$|\{p \text{ primo } : \ p \leq x\}| \sim \frac{x}{\log x},$$

donde $|\cdot|$ representa el cardinal de un conjunto. Si la función $\zeta(s)$ es la que permite estudiar la distribución de los primos, las funciones L de Dirichlet son las que nos proporcionan información sobre la distribución de los primos en progresiones aritméticas. Están definidas por

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


Figura 5: Dirichlet y Riemann.

para $\Re(s) > 1$, y por continuación analítica para $\Re(s) \leq 1$. Aquí χ es cualquier carácter de Dirichlet; para cada χ dado, $L(s, \chi)$ es una función de s. Un carácter de Dirichlet χ (de módulo q) es una función $\chi : \mathbb{Z} \to \mathbb{C}$ de período q (esto es, $\chi(n) = \chi(n+q)$ para todo n), con las propiedades adicionales de que es multiplicativa ($\chi(ab) = \chi(a)\chi(b)$ para a, b cualesquiera) y que $\chi(n) = 0$ cuando n y q no son coprimos. (La forma sofisticada de decir todo esto es que χ es un carácter de $(\mathbb{Z}/q\mathbb{Z})^*$ en $\mathbb{Z}$.)

Un *cero* de una función f es un $s \in \mathbb{C}$ tal que $f(s) = 0$. Un *cero no trivial* de $\zeta(s)$, o de $L(s, \chi)$, es un cero de $\zeta(s)$, o de $L(s, \chi)$, tal que $0 < \Re(s) < 1$. (Los otros ceros son llamados triviales porque es fácil decir dónde están, a saber, en ciertos enteros no positivos.) La hipótesis de Riemann asevera que todos los ceros no triviales de la función zeta de Riemann «yacen en la recta crítica», lo cual significa que $\Re(s) = 1/2$. La hipótesis generalizada de Riemann para funciones L de Dirichlet dice que, para todo carácter de Dirichlet χ, todo cero no trivial de $L(s, \chi)$ satisface $\Re(s) = 1/2$.

Como tanto la hipótesis de Riemann (HR) como la hipótesis generalizada de Riemann (HGR) siguen sin ser demostradas, cualquier resultado probado usando cualquiera de ellas será condicional; ahora bien, nosotros queremos probar resultados incondicionales. Lo que sí puede ser demostrado, y utilizado, son resultados parciales en la dirección de la HGR. Tales resultados son de dos tipos:

- Regiones libres de ceros. Desde finales del siglo XIX (de la Vallée-Poussin) sabemos que hay regiones con forma de reloj de arena (más precisamente, de la forma $c/\log t \leq \sigma \leq 1 - c/\log t$, donde c es una constante y donde escribimos $s = \sigma + it$) fuera de las cuales no pueden yacer ceros no triviales.
- Verificaciones finitas de HGR. Es posible (usando un ordenador) probar pedazos finitos y no muy grandes de la HGR, en el sentido de verificar que todos los ceros s no triviales de una función $L(s, \chi)$ (χ dado) con parte imaginaria $\Im(s)$ menor que alguna constante H yacen en la recta crítica $\Re(s) = 1/2$.

La mayor parte de los trabajos hasta la fecha sigue la primera alternativa. Yo elegí la segunda, y esto tuvo consecuencias para la manera en la que definí los arcos mayores y menores: conseguí resultados muy precisos en los arcos mayores, pero tuve que definirlos de tal manera que fueran pocos y muy estrechos; si no, el método no hubiera funcionado. Esto significó que los métodos para los arcos menores tenían que ser particularmente potentes, ya que una parte del círculo más grande de lo habitual quedó para ser tratada con ellos.

Vamos a ver más detenidamente cómo se puede lidiar con los arcos mayores usando resultados parciales de la HGR y, en particular, verificaciones finitas de la HGR.

4. Estimaciones en los arcos mayores

Recordemos que queremos calcular sumas del tipo $\widehat{f}(\alpha) = \sum f(n)e(-\alpha n)$, donde $f(n)$ es algo como $(\log n)e^{-n/x}$ para n primo y 0 para n compuesto. Vamos a modificar esto sólo un poco; de hecho calcularemos

$$S_\eta(\alpha, x) = \sum \Lambda(n)e(\alpha n)\eta(n/x),$$

donde Λ es la función de Mangoldt: $\Lambda(n) = \log p$ si n es de la forma p^k, con $k \geq 1$, y $\Lambda(n) = 0$ de lo contrario. (El uso de α en vez de $-\alpha$ es sólo una concesión a la tradición, como lo es el uso de la letra S, de «suma». Por otra parte, el uso de $\Lambda(n)$ en lugar de simplemente $\log p$ simplifica las cosas cuando hay que trabajar con las así llamadas *fórmulas explícitas*, que veremos enseguida.) Aquí $\eta(t)$ es alguna función de decaimiento rápido; puede ser e^{-t}, como en el trabajo de Hardy y Littlewood, o (como en mi trabajo) alguna otra función. (Podría incluso ser el «truncamiento brutal» $1_{[0,1]}(t)$, igual a 1 cuando $t \in [0, 1]$ y a 0 de lo contrario; esto sería bueno para los arcos menores, pero, como veremos, resulta ser una mala idea cuando se tratan los arcos mayores.)

Asumamos que α está en un arco mayor, es decir, que podemos escribir α de la forma $\alpha = a/q + \delta/x$ para algún a/q (q pequeño) y algún δ (con $|\delta|$ pequeño). Podemos expresar $S_\eta(\alpha, x)$ como una combinación lineal (esto es, una suma de múltiplos) de términos de la forma $S_{\eta,\chi}(\delta/x, x)$, donde

$$S_{\eta,\chi}(\delta/x, x) = \sum_n \Lambda(n)\chi(n)e(\delta n/x)\eta(n/x)$$

y χ recorre los caracteres de Dirichlet de módulo q.

¿Por qué son las sumas $S_{\eta,\chi}$ mejores que las sumas S_η? El argumento se ha convertido en δ/x, donde antes era α. Aquí δ es pequeño —más pequeño que una constante, en nuestro tratamiento—. En otras palabras, $e(\delta n/x)$ se moverá alrededor del círculo un número acotado de veces a medida que n vaya de 1 hasta, digamos, $10x$ (para cuando $\eta(n/x)$ es ya muy pequeño). Esto hace que la suma $S_{\eta,\chi}$ sea mucho más fácil de calcular.

Es un hecho estándar que podemos expresar $S_{\eta,\chi}$ mediante una fórmula explícita (sí, la frase tiene un significado técnico, como el *Jugendtraum* de Kronecker):

$$S_{\eta,\chi}(\delta/x, x) = [\widehat{\eta}(-\delta)]x - \sum_{\rho} F_\delta(\rho)x^\rho + \text{error pequeño.}$$

Aquí el término entre corchetes aparece sólo para $q = 1$. En la suma, ρ recorre todos los ceros no triviales de $L(s, \chi)$, y F_δ es la transformada de Mellin de $e(\delta t)\eta(t)$:

$$F_\delta(s) = \int_0^\infty e(\delta t)\eta(t)t^s \frac{dt}{t}.$$

Lograremos nuestro objetivo si llegamos a demostrar que la suma $\sum_\rho F_\delta(\rho)x^\rho$ es pequeña.

La cuestión es esta: si comprobamos la HGR hasta parte imaginaria H, entonces sabemos que todo ρ con $|\Im(\rho)| \leq H$ satisface $\Re(\rho) = 1/2$, y por lo tanto $|x^\rho| = \sqrt{x}$. En otras palabras, x^ρ es entonces muy pequeño (comparado con x). Sin embargo, para cualquier ρ cuya parte imaginaria tenga valor absoluto mayor que H no sabemos nada sobre su parte real aparte de que $0 < \Re(\rho) < 1$. (De acuerdo, podríamos usar una región libre de ceros, pero las regiones libres de ceros conocidas son notoriamente débiles para $\Im(\rho)$ grande —es decir, nos servirían de poco en la práctica—.) Por lo tanto, nuestra única opción es asegurarnos que $F_\delta(\rho)$ sea pequeña cuando $|\Im(\rho)| \geq H$.

Esto, claro está, tendría que ser cierto para δ muy pequeño (incluyendo $\delta = 0$) y para δ no tan pequeño (δ entre 1 y una constante). Si se juega con el *método de la fase estacionaria*, se consigue ver que $F_\delta(\rho)$ se comporta como $M_\eta(\rho)$ para δ muy pequeño (aquí M_η es la transformada de Mellin de η) y como $\eta(t/|\delta|)$ para δ no tan pequeño (donde $t = \Im\rho$). Por tanto, estamos en un dilema clásico, a menudo llamado «principio de incertidumbre» porque es el hecho matemático subyacente al principio físico del mismo nombre: no se puede tener una función η que decrezca muy rápidamente y cuya transformada de Fourier (o, en este caso, su transformada de Mellin) también decaiga muy rápidamente.

¿Qué significa aquí «muy rápidamente»? Significa «más rápido que cualquier exponencial e^{-Ct}». Por tanto, la elección $\eta(t) = e^{-t}$ de Hardy y Littlewood parecería ser esencialmente óptima.

¡No tan deprisa! Lo que podemos hacer es elegir η de tal manera que M_η decrezca exponencialmente (con una constante C un poco peor que antes) y que η decrezca más rápido que exponencialmente. Esto es lo crucial, ya que $t/|\delta|$ (y no tanto t en sí) corre el riesgo de ser bastante pequeño.

Una elección de η que obedece a esta descripción es la *función gaussiana* $\eta(t) = e^{-t^2/2}$. La transformada de Mellin F_δ es entonces una función *cilíndrica parabólica*, con valores imaginarios para uno de sus parámetros. Las funciones cilíndricas parabólicas parecen ser muy apreciadas y estudiadas en el mundo aplicado —pero más que nada para valores reales del citado parámetro—. Hay algunos desarrollos asintóticos de F_δ en la literatura para parámetros generales (notablemente por F. W. J. Olver), pero ninguna que sea suficientemente explícita para mis propósitos. Por

tanto, tenía que proporcionar estimaciones totalmente explícitas yo mismo, usando el método del punto de silla. Esto me llevó un buen rato, pero los resultados seguramente serán de aplicación general —hola, ingenieros— y también es de esperar que la función gaussiana se vuelva un poco más popular en trabajos explícitos en teoría de números.

A propósito, estas estimaciones de funciones cilíndricas parabólicas nos permiten tomar no sólo $\eta(t) = e^{-t^2/2}$, sino también, más generalmente, $\eta(t) = h(t)e^{-t^2/2}$, donde h es cualquier función de banda limitada, lo que significa, en este contexto, cualquier función h cuya transformada de Mellin restringida al eje y tenga soporte compacto. Deseamos optimizar la elección de $h(t)$ —hablaremos de ello más adelante.

5. Los arcos menores

¿Cómo acotamos $|S_\eta(\alpha, x)|$ cuando α no está cerca de ningún racional a/q de denominador pequeño? Que esto sea posible fue el gran logro de Vinogradov. El progreso desde entonces ha sido gradual. Doy mi propio enfoque al asunto en mi artículo «Minor arcs...» [11]. Déjenme comentar algunas de las ideas detrás de los avances allí contenidos.

La demostración de Vinogradov fue simplificada sustancialmente en los 70 (del siglo XX) por Vaughan, quien introdujo la identidad que ahora lleva su nombre [28]. Básicamente, la identidad de Vaughan es un gambito: otorga una gran flexibilidad, pero a un precio —aquí, un precio de dos logaritmos, en lugar de, digamos, dos peones—. El problema es que, si queremos alcanzar nuestro objetivo, no podemos permitirnos el lujo de perder logaritmos. La única opción es recuperar esos logaritmos encontrando cancelaciones en las diferentes sumas que surgen de la identidad de Vaughan. Esto se tiene que hacer, por cierto, sin usar funciones L, puesto que ya no podemos asumir que q sea pequeño.

He aquí otro aspecto de esta parte de la prueba. Todo α tiene una aproximación de la forma $\alpha = a/q + \delta/x$; el hecho de que α esté en los arcos menores nos dice que q no es muy pequeño. Estamos buscando cotas que decrezcan a medida que q crece; la cota que yo obtengo es proporcional a $(\log q)/\sqrt{\phi(q)}$. ¿Cuál es el efecto de δ?

Algo de lo que me di cuenta pronto fue que, si δ no es muy pequeño, puede de hecho ser utilizado en nuestro beneficio. Una razón es que hay términos de la forma $\widehat{\eta}(\delta)$, y la tranformada de Fourier de funciones suaves decae conforme el argumento crece. Hay otras razones, empero. Algo que podemos usar es lo siguiente: por un resultado básico de aproximación diofántica, todo α tiene muy buenas aproximaciones por racionales con denominador no demasiado grande. Si δ no es muy pequeño, entonces la aproximación $\alpha = a/q + \delta/x$ es buena, pero no muy buena; por lo tanto, debe haber otra mejor aproximación $\alpha \sim a'/q'$ con q' no demasiado grande (lo que significa «considerablemente más pequeño que x»). Podemos ir y volver entre las aproximaciones a/q y a'/q', dependiendo de cuál sea más útil en cada momento. Ello resulta ser mejor que usar una sola aproximación a/q, por muy buena que esta sea.

Otra manera en la que se consigue sacar provecho de un δ grande es esparciendo las entradas en una *gran criba*. La gran criba puede ser vista como una forma apro-

ximada de la identidad de Plancherel, reformulada como una desigualdad. Mientras la identidad de Plancherel nos dice que la norma $|\widehat{f}|_2$ (norma ℓ_2) de la transformada de Fourier $\widehat{f} : \mathbb{R}/\mathbb{Z} \to \mathbb{Z}$ de una función f definida en los enteros (también es cierto para los reales u otros grupos) es igual a la norma $|f|_2$ de la misma función f, la gran criba nos dice que el total de $|\widehat{f}(\alpha_i)|^2$ para una muestra bien espaciada de puntos $\alpha_i \in \mathbb{R}/\mathbb{Z}$ está acotada por (un múltiplo de) $|f|^2$. Ahora bien, en nuestro caso, los puntos α_i son múltiplos de nuestro ángulo α. Si $\alpha = a/q$, el espacio entre los puntos α_i es $1/q$, lo cual es bueno —pero puede ser que tengamos que aplicar la gran criba varias veces, ya que tenemos que aplicarla de nuevo para cada tanda de q puntos—. Sin embargo, si $\alpha = a/q + \delta/x$ y δ no es demasiado pequeño, podemos rodear el círculo varias veces y confiar en δ/x en vez de en $1/q$ para darnos el espacio. Sí, δ/x (e incluso $\delta q/x$) es más pequeño que $1/q$, pero el efecto de esto está más que compensado por el hecho de que tenemos que recurrir a la gran criba muchas menos veces (quizás solamente una vez).

Lo que es más interesante, esta manera de esparcir los ángulos puede ser combinada con otra manera más tradicional de esparcirlos (lema de Montgomery; ver [18, (3.9)], o la exposición en [15, §7.4]) con el fin de aprovechar el hecho de que estamos tratando con sumas donde la variable recorre los primos p.

6. Conclusión

Hemos estado hablando acerca de acotar $S_\eta(\alpha, x)$ para α dentro de los arcos menores, pero lo que queremos hacer realmente es acotar la integral $\int_m |S_\eta(\alpha, x)|^3 \, d\alpha$. Una forma fácil —y tradicional— de hacer esto consiste en usar la desigualdad trivial

$$\int_m |S_\eta(\alpha, x)|^3 \, d\alpha \leq \max_{\alpha \in m} |S_\eta(\alpha, x)| \cdot \int_m |S_\eta(\alpha, x)|^2 \, d\alpha.$$

Desgraciadamente, así perderíamos un factor de un logaritmo.

Como nuestras cotas para $S_\eta(\alpha, x)$, $\alpha \sim aq$, están dadas en términos de q, tiene sentido combinarlas con estimaciones para integrales del tipo $\int_{m_r} |S_\eta(\alpha, x)|^2 \, d\alpha$, donde m_r es una unión de arcos alrededor de racionales con denominador más grande que una constante pero menor que r. ¿Cómo estimamos estas integrales? Esta pregunta está muy relacionada con otra que entra dentro del marco de la gran criba: ¿qué cotas se pueden conseguir para $\alpha_i = a/q$, $q \leq r$, donde r es de tamaño moderado, si es que estamos trabajando con una sucesión con soporte en los primos?

Había una respuesta en la literatura (basada en el lema de Montgomery; el enlace con el método del círculo ya fue observado por Heath-Brown) pero era peor que la cota óptima por un factor de al menos e^γ (o de hecho más); es este el resultado utilizado en [25, § 4] y la primera version de [11, § 6]. También había una estimación más reciente para la gran criba debida a Ramaré ([22, Thm. 2.1]; ver también [22, Thm. 5.2]), pero no se había hecho totalmente explícita. Tuve que hacerla explícita, y luego adapté el nuevo resultado sobre la gran criba a la tarea de estimar la integral sobre m_r. Como era de esperar, el factor e^γ (o realmente un poco más) desapareció.

Queda por comparar el término principal con el error. Resulta que tenemos cierto margen para elegir lo que será el término principal, ya que depende de los pesos η que utilicemos. El término principal es proporcional a

$$\int_0^\infty \int_0^\infty \eta_+(t_1)\eta_+(t_2)\eta_*(N/x-(t_1+t_2))\,dt_1\,dt_2,$$

donde η_+ y η_* son los dos pesos con los que escogemos trabajar, N es el número impar que queremos expresar como suma de tres primos y x es de nuevo un parámetro de nuestra elección. En comparación, el error es proporcional a $|\eta_+|^2|\eta_*|_1$. Así, tenemos un problema de optimización («maximizar el tamaño de la doble integral dividida por $|\eta_+|^2|\eta_*|_1$»). Lo mejor es elegir un peso η_+ simétrico o cercano a ser simétrico ($\eta_+(t) \sim \eta_+(2-t)$), asegurándonos, por otra parte, que $\eta_+(t) \sim 0$ para $t \geq 2$. Esto no es demasiado difícil de conseguir aun bajo la restricción de que η_+ sea de la forma $\eta(t) = h(t)e^{-t^2/2}$, donde h es de banda limitada.

¿Qué pasa con η_*? La solución del problema de optimización nos dice que debe ser de soporte pequeño, o por lo menos concentrado cerca del origen. Aparte de eso, hay, por decirlo así, un problema político: η_*, a diferencia de η_+, se usa tanto en los arcos mayores como en los menores; los arcos mayores quieren de verdad que sea de la forma $e^{-t^2/2}$ o $t^k e^{-t^2/2}$, mientras los arcos menores preferirían algo más simple, como $\eta_{[0,1]}$ o como $\eta_2 = (2\eta_{[1/2,1]}) *_M (2\eta_{[1/2,1]})$, donde $f *_M g$ es la *convolución multiplicativa* (o *convolución de Mellin*):

$$(f *_M g)(x) = \int_0^\infty f(y)g\left(\frac{x}{y}\right)\frac{dy}{y}.$$

(Aquí η_2 es precisamente el peso usado en el artículo de Tao sobre los cinco primos, o en mi propio artículo sobre los arcos menores.)

La solución es simple: definamos $\eta_*(t) = (f *_M g)(\kappa t)$, donde κ es una constante grande, $f(t) = \eta_2(t)$ y $g(t) = t^2 e^{-t^2/2}$. Para f y g esencialmente arbitrarias, si se sabe cómo calcular (o estimar) $S_f(\alpha, x)$ para algunos α, y se sabe estimar $S_g(\alpha, x)$ para todos los otros α, entonces se sabe cómo estimar $S_{f*_M g}(\alpha, x)$ para todo α. (La prueba sale en un par de líneas; se escribe qué es $S_{f*_M g}$ en detalle y se cambia el orden de la suma y la integral. En el proceso también se aclara que ayuda si $f(t)$ y $g(t)$ son pequeños para t cercano a 0.)

La moraleja de esta historia es que diferentes problemas, y diferentes partes del mismo problema, exigen diferentes pesos η. Al menos en el contexto de sumas exponenciales, resulta haber un simple truco para combinarlas, como acabamos de ver.

7. Algunos comentarios finales sobre computación

Una demostración analítica normalmente da una prueba válida para todo n mayor que una constante C. La razón es simple: digamos que queremos mostrar que una cantidad es positiva. Generalmente, después de bastante trabajo analítico, se llega

a probar que la cantidad es de la forma 1 + error, donde el valor absoluto de este error es menor que, digamos, C/n (para dar un ejemplo simple). Esto ciertamente muestra que la cantidad es positiva, a condición de que $n \geq C$. La tarea, entonces, es refinar la demostración hasta que la constante C sea suficientemente pequeña para que todos los casos en los que $n \leq C$ puedan ser comprobados a mano (literalmente a mano o con un ordenador). Esto fue, en gran parte, mi trabajo: comprobar la conjetura hasta $C = 10^{29}$ (y de hecho hasta $8.8 \cdot 10^{30}$) fue, en comparación, una tarea secundaria —como veremos, no era siquiera el principal esfuerzo computacional.

Primero, permítanme decir algunas palabras más en general sobre resultados analíticos. Hay resultados del tipo «la proposición es cierta para todo n mayor que una constante C, pero esta demostración no nos dice nada sobre C aparte de que existe». A esto se le llama una «estimación *inefectiva*»; muchas demostraciones de los resultados de Vinogradov en libros de texto son de este tipo. (La razón detrás de esto es la posible existencia de los así llamados «ceros de Siegel».) Un resultado puede decir también «la sentencia es cierta para todo $n > C$, y en principio se debería poder determinar algún valor de C usando las ideas de la prueba, pero el autor preferiría irse a tomar un café». Esta es una proposición efectiva no explícita; la versión definitiva de Vinogradov de su propio resultado fue de este tipo (como lo son muchos otros resultados en matemáticas, incluyendo algunos de mi propio pasado). Si se da un valor explícito de C, entonces el resultado se denomina (¡sorpresa!) «explícito». Queda la cuarta etapa: conseguir que C sea razonable, esto es, suficientemente pequeña como para que el caso $n \leq C$ pueda ser comprobado a mano. Estuvo claro desde el principio que, en el caso de la conjetura ternaria de Goldbach, «razonable» significaba aproximadamente $C \sim 10^{30}$ (aunque las verificaciones existentes llegaban a bastante menos que 10^{30}).

Dije antes que D. Platt y yo comprobamos la conjetura para todos los impares hasta $8.8 \cdot 10^{30}$. He aquí como procedimos. Ya se sabía (gracias a un esfuerzo de gran envergadura de parte de Oliveira e Silva, Herzog y Pardi [19]) que la conjetura binaria de Goldbach es cierta hasta $4 \cdot 10^{18}$ —esto es, todo número par hasta $4 \cdot 10^{18}$ es suma de dos números primos—. Sabiendo esto, todo lo que teníamos que hacer era construir una *escalera de primos*, esto es, una lista de primos desde 2 hasta $8.8 \cdot 10^{30}$ tal que la diferencia entre cualesquiera dos primos consecutivos de la lista fuera a lo más $4 \cdot 10^{18}$. Por tanto, si alguien le da a uno un entero impar n hasta $8.8 \cdot 10^{30}$, se sabe que hay un primo p en la lista tal que $n - p$ es positivo y a lo más $4 \cdot 10^{18}$. Por hipótesis, podemos escribir $n - p = p_1 + p_2$ para algunos primos p_1, p_2, y, por tanto, $n = p + p_1 + p_2$.

Construir esta escalera no nos llevó mucho tiempo. (De hecho, conseguir una escalera hasta 10^{29} es probablemente algo que el lector pueda hacer en su ordenador personal en unas pocas semanas —aunque almacenarla es otro asunto—.) La tarea se hace en «aritmética entera», y comprobamos la primalidad de los números en la escalera de manera determinista (restringiéndonos a primos para los cuales hay un algoritmo rápido de comprobación de primalidad), así que no hay que preocuparse.

El cálculo computacional más grande ha consistido en verificar que, para toda función L de conductor q hasta sobre 15 000 (o dos veces esto para q par), todos los ceros de la función L con parte imaginaria acotada por $10^8/q$ yacen sobre la

línea crítica. Esto fue por completo obra de Platt [20]; mi única contribución fue ir a buscar tiempo de ordenador por muchas partes (ver la sección de agradecimientos del artículo «Major arcs...» [12]). De hecho, Platt llegó hasta conductor 200 000 (o dos veces esto para q par); ya había llegado hasta el conductor 100 000 en su tesis. La verificación llevó, en total, unas 400 000 horas de núcleo (esto es, el número total de núcleos (*cores*) de procesador usados por el número de horas que corrieron es igual a 400 000; hoy en día, un procesador de primera línea —como los de la máquina en *MesoPSL*— normalmente tiene ocho núcleos). Al final, como decía, usé solamente $q \leq 150\,000$ (o el doble de esto para q par), por lo que el número de horas necesarias fue de hecho unas 160 000; como me hubiera bastado con aproximadamente $q \leq 120\,000$, podría decir que, en retrospectiva, se necesitaban sólo unas 80 000 horas de núcleo. Los ordenadores y yo fuimos cavando por lados opuestos de la montaña, y nos encontramos en el centro. El hecho de que los ordenadores trabajaran más de lo necesario no es nada que lamentar: el cálculo hecho es de uso general, y por tanto es mucho mejor que no esté «hecho a la medida» de mis necesidades. Por otra parte, con demostraciones de esta longitud, lo mejor es «construir como un romano», es decir, calcular de más por si uno (¡no el ordenador!) ha cometido algún pequeño error en algún sitio. (¿Por qué creen que esas paredes eran tan gruesas?)

Comprobar los ceros de la función L computacionalmente es algo tan viejo como Riemann (quien lo hizo a mano); es también una de las cosas que se intentaron en computadoras electrónicas ya en sus primeros días (Turing tenía un artículo sobre eso). Una de las principales cuestiones con las que hay que tener cuidado surge cuando se quieren manipular números reales: hablando honestamente, un ordenador no puede almacenar π; más aún, si bien un ordenador puede manejar números racionales, realmente se siente cómodo sólo cuando maneja aquellos racionales cuyos denominadores son potencias de dos. Por tanto, en realidad no se puede decir: «ordenador, dame el seno de este número» y esperar un resultado preciso. Lo que se debería hacer, si realmente se quiere *probar* algo (¡como en este caso!) es decir: «ordenador, te estoy dando un intervalo $I = [a/2^k, b/2^k]$; dame un intervalo $I' = [c/2^l, d/2^l]$, preferiblemente muy pequeño, tal que $\operatorname{sen}(I) \subset I'$». Esto se llama «aritmética de intervalos»; es realmente la forma más sencilla de hacer cálculos en coma flotante de manera rigurosa.

Ahora, los procesadores no hacen esto de forma nativa, y si se hace puramente con software se retrasan las cosas en un factor de más o menos 100. Afortunadamente, hay maneras de hacer esto a medias con hardware y a medias con software. Platt tiene su propia biblioteca de rutinas, pero hay otras *online* (por ejemplo, *PROFIL/BIAS* [16]).

(Oh, a propósito, no usen la función seno en un procesador Intel si quieren que el resultado sea correcto hasta el último bit. ¿En qué habrán estado pensando? Usen la biblioteca *CRlibm* [4] en su lugar.)

Por último, hubo varios cálculos bastante menores que hice yo mismo; los encontrarán mencionados en mis artículos. Un cálculo habitual fue una versión rigurosa de una «demostración por gráfica» («el máximo de una función f es claramente menor que 4 porque *gnuplot* me lo dijo»). El lector encontrará algoritmos para esto en cualquier libro de texto sobre «computación validada» —básicamente, es suficiente

combinar el método de la bisección con aritmética de intervalos.

Finalmente, déjenme indicar que hay una desigualdad elemental en el artículo «Minor arcs...» (viz., (4.24) en la demostración del lema 4.2) que fue probada en parte por un humano (yo) y en parte por un programa de eliminación de cuantificadores. En otras palabras, ya existen programas de ordenador (en este caso, *QEPCAD* [14]) que pueden probar cosas útiles. Ahora bien, no tengo dudas de que la misma desigualdad puede ser probada puramente mediante el uso de seres humanos, pero es bonito saber que nuestros amigos los ordenadores pueden (pretender) hacer algo más que masticar números...

Nota. Este artículo está basado en un texto del autor publicado en su blog [13]. Se deben gracias a M. A. Morales, por una primera traducción, y a J. Cilleruelo y M. Helfgott, por muchos comentarios, así como a F. Chamizo, por la gráfica de la figura 4.

Nota añadida en la imprenta. R. Vaughan me indica que Descartes mencionó que «todo entero es igual a [la suma de] uno, dos o tres primos» en un manuscrito publicado de manera póstuma en 1901 [6, p. 298]. Dickson alude a esto en su historia [8, p. 421], pero de manera un tanto oscurecida por una traducción dudosa: la palabra latina «par» fue traducida como «even» (en castellano, par) cuando, dice Vaughan, debió haber sido traducida como «equal» (igual). En resumen, Descartes planteó la conjetura de Goldbach un siglo antes de éste. Se trata, claro está, de un enunciado empírico (no publicado) y no de algo que Descartes supiera probar; ni siquiera parece haberlo planteado de manera explícita como un problema para ser resuelto.

Referencias

[1] J. R. Chen, On the representation of a larger even integer as the sum of a prime and the product of at most two primes, *Sci. Sinica* **16** (1973), 157–176.

[2] J. R. Chen y T. Z. Wang, On the Goldbach problem, *Acta Math. Sinica* **32** (1989), 702–718.

[3] J. R. Chen y T. Z. Wang, The Goldbach problem for odd numbers, *Acta Math. Sinica (Chin. Ser.)* **39** (1996), 169–174.

[4] C. Daramy-Loirat, F. de Dinechin, D. Defour, M. Gallet, N. Gast y Ch. Lauter, *CRlibm – Correctly Rounded mathematical library*, mayo 2010, version 1.0beta4.

[5] H. Davenport, *Multiplicative number theory*, Graduate Texts in Mathematics 74, Springer-Verlag, New York, tercera edición, 2000.

[6] R. Descartes, *Oeuvres* (Ch. Adam y P. Tannery, eds.), vol. 10, L. Cerf, Paris, 1901.

[7] J.-M. Deshouillers, Sur la constante de Šnirel'man, *Séminaire Delange-Pisot-Poitou*, 17e année: (1975/76), Théorie des nombres: Fac. 2, Exp. No. G16, Secrétariat Math., Paris, 1977.

[8] L. E. Dickson, *History of the theory of numbers. Vol. I: Divisibility and primality*, Chelsea Publishing Co., New York, 1966.

[9] T. Estermann, On Goldbach's Problem: Proof that almost all even positive integers are sums of two primes, *Proc. London Math. Soc. Ser. 2* **44** (1937), no. 4, 307–314.

[10] G. H. Hardy y J. E. Littlewood, Contributions to the theory of the Riemann zeta-function and the theory of the distribution of primes, *Acta Math.* **41** (1916), 119–196.

[11] H. A. Helfgott, Minor arcs for Goldbach's problem. Prepublicación, `http://arxiv.org/abs/1205.5252`.

[12] H. A. Helfgott, Major arcs for Goldbach's problem. Prepublicación, `http://arxiv.org/abs/1305.2897`.

[13] H. A. Helfgott, The ternary Goldbach conjecture, `http://valuevar.wordpress.com/2013/07/02/the-ternary-goldbach-conjecture/`, 2013.

[14] H. Hong y Ch. W. Brown, *QEPCAD B – Quantifier elimination by partial cylindrical algebraic decomposition*, mayo 2011, version 1.62.

[15] H. Iwaniec y E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, 53, Amer. Math. Soc., Providence, RI, 2004.

[16] O. Knüppel, PROFIL/BIAS, febrero 1999, versión 2.

[17] M.-Ch. Liu y T. Wang, On the Vinogradov bound in the three primes Goldbach conjecture, *Acta Arith.* **105** (2002), 133–175.

[18] H. L. Montgomery, *Topics in multiplicative number theory*, Lecture Notes in Mathematics, 227, Springer-Verlag, Berlin, 1971.

[19] T. Oliveira e Silva, S. Herzog y S. Pardi, Empirical verification of the even Goldbach conjecture, and computation of prime gaps, up to $4 \cdot 10^{18}$, *Math. Comp.*, aceptado.

[20] D. Platt, Numerical computations concerning GRH. Prepublicación.

[21] O. Ramaré, On Šnirel'man's constant, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* **22** (1995), 645–706.

[22] O. Ramaré, *Arithmetical aspects of the large sieve inequality*, with the collaboration of D. S. Ramana, Harish-Chandra Research Institute Lecture Notes, vol. 1, Hindustan Book Agency, New Delhi, 2009.

[23] H. Riesel y R. C. Vaughan, On sums of primes, *Ark. Mat.* **2** (1983), 46–74.

[24] L. Schnirelmann, Über additive Eigenschaften von Zahlen, *Math. Ann.* **107** (1933), 649–690.

[25] T. Tao, Every odd number greater than 1 is the sum of at most five primes, *Math. Comp.*, aceptado.

[26] T. Tao, Heuristic limitations of the circle method, `http://terrytao.wordpress.com/2012/05/20/heuristic-limitations-of-the-circle-method/`, 2012.

[27] R. C. Vaughan, On the estimation of Schnirelman's constant, *J. Reine Angew. Math.* **290** (1977), 93–108.

[28] R. C. Vaughan, Sommes trigonométriques sur les nombres premiers, *C. R. Acad. Sci. Paris Sér. A-B* **285** (1977), A981–A983.

[29] I. M. Vinogradov, Representation of an odd number as a sum of three primes, *Dokl. Akad. Nauk. SSR* **15** (1937), 291–294.

Harald Andrés Helfgott, Ecole Normale Supérieure, Département de Mathématiques, 45 rue d'Ulm, F-75230 Paris, France

Correo electrónico: harald.helfgott@ens.fr

Página web: http://www.math.ens.fr/~helfgott/