Las Medallas Fields

Sección a cargo de

Ana Jeremías López y Leovigildo Alonso Tarrío

Atle Selberg: 1917–2007

por

Fernando Chamizo Lorente

1. Introducción

Atle Selberg nació el 14 de junio de 1917 en Langesund, una ciudad costera al sur de Noruega, y murió el día 6 del pasado mes de agosto en Princeton (Estados Unidos). Se doctoró en la Universidad de Oslo en 1943 y, después del aislamiento que supuso la Segunda Guerra Mundial, se trasladó a Estados Unidos, donde su

A. Selberg celebrando su 90 cumpleaños en Princeton (junio de 2007). Foto cortesía de C.J. Mozzochi. © C. J. Mozzochi, Princeton, N.J.

vida académica estuvo ligada al IAS (Institute for Advanced Studies) del que fue miembro permanente desde 1949. Falleció siendo profesor emérito, dos meses después de asistir a la celebración de su 90 cumpleaños en el IAS.

La obra de Selberg se centra en la teoría analítica de números, pero su influencia excede este marco. Tras ocuparse inicialmente de temas relacionados con formas modulares (en [23] introduce lo que hoy conocemos como *convolución de Rankin-Selberg*) y con interpolación de funciones complejas, llegaron sus notorios trabajos [24], [25] y [26], sobre los ceros de la función ζ de Riemann. Finalizada la guerra mundial, ven la luz sus espectaculares avances en los métodos de criba y la demostración elemental del teorema de los números primos, después de lo cual recibe la medalla Fields en 1950. En los años siguientes dirige su atención a las formas automorfas no holomorfas que había introducido H. Maass en 1949 (en el artículo [17]) y desarrolla la teoría espectral. Dentro de ella obtiene el resultado que le ha dado más fama: *la fórmula de traza de Selberg*, que aparece incluida en [29] con meras explicaciones esquemáticas en un contexto más general. De su trabajo posterior más innovador cabe destacar su artículo [30] de 1962 sobre el tamaño de los coeficientes de formas modulares, basado en investigaciones no publicadas anteriores, y la prueba de la *rigidez de los retículos* [31]. Posiblemente la última de sus publicaciones tardías que ha marcado una línea de investigación viva y activa es [32], que ha dado lugar a la llamada *clase de Selberg* de series de Dirichlet.

Según se recoge en [14], Selberg ya septuagenario dijo acerca de su obra: *"las ideas básicas fueron siempre sencillas y podrían ser explicadas en términos bastante simples"* y *"Probablemente tenga una mente simplista de modo que éste es el tipo de ideas con el que puedo trabajar. No creo que otras personas hayan tenido serias dificultades para comprender mi trabajo"*. Naturalmente, estas afirmaciones son opinables, y no sería extraño que más de un matemático viera retorcida ironía en las últimas palabras, dados los pocos esfuerzos por parte de Selberg para difundir adecuadamente su obra y dotarla de referencias que permitieran ubicarla mejor. Pero es un hecho que habitualmente el corazón de sus pruebas es una idea elemental, una especie de truco maravilloso. E. Bombieri lo resume sentenciosamente cuando, después de citar en [2] un resultado de criba, afirma: *"Esto es típico del estilo de Selberg: simplicidad y elegancia del método, resultados potentes"*.

En este artículo trataremos de aprovechar la oportunidad que ofrece la sencillez interna que el propio Selberg atribuye a su obra, y su brevedad, para divulgar los temas que la rodean. Preferimos alterar el orden cronológico en favor de la claridad y la conexión. Organizaremos el material en grandes bloques, comenzando por los más elementales, y reservaremos un apartado final como glosario de algunos términos que no aparecen en ellos. Nuestro propósito general es dirigir la mirada más a los temas que al enunciado preciso de los resultados particulares.

2. El teorema de los números primos

2.1. La distribución de los primos

La prehistoria de la distribución de los números primos se remonta a Euler, de quien parten los primeros tímidos albores de la teoría analítica de números. A partir de la fórmula producto

$$\zeta(s) = \prod_p (1 - p^{-s})^{-1}, \qquad \text{donde} \qquad \zeta(s) = \sum_{n=1}^{\infty} n^{-s},$$

(p recorre los primos) empleó la singularidad en $s = 1$ para deducir, en su famoso artículo [9] de 1737, afirmaciones imprecisas como: "Los primos son infinitamente más numerosos que los cuadrados" o "Los primos son infinitamente menos numerosos que los naturales". También probó que la suma de los inversos de los primos diverge.

A finales del siglo XVIII, Gauss y Legendre obtuvieron numéricamente sendas fórmulas aproximadas para $\pi(x)$, la cantidad de primos menores o iguales que x. La aproximación de Gauss fue más acertada y se basaba en el hecho experimental de que la densidad de los primos de tamaño comparable a t (grande) está cercana a $1/\log t$. Por ello, $\pi(x)$ debería parecerse a $\mathrm{li}(x) = \int_2^x dt/\log t$. El *teorema de los números primos* afirma que el error relativo en esta aproximación tiende a cero. Teniendo en cuenta que[1] $\mathrm{li}(x) \sim x/\log x$ se puede enunciar este resultado como

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

Dicho sea de paso, aunque el enunciado habitual del teorema sea éste, se puede probar que $\mathrm{li}(x)$ da aproximaciones numéricas más precisas. Por ejemplo $\pi(10^8)/\mathrm{li}(10^8) = 0{,}99987\ldots$, mientras que $\pi(10^8)(\log 10^8)/10^8 = 1{,}061\ldots$

El primer avance significativo hacia (1) se debe a P. Chebyshev a mediados del siglo XIX, quien probó con ingeniosos argumentos elementales que, si $\pi(x)(\log x)/x$ tiene límite, entonces éste debe ser 1. En su exitoso empeño de establecer el *postulado de Bertrand* (si $n > 1$, existe p primo tal que $n \le p < 2n$) aproximó ciertas cantidades relacionadas con los primos. Por ejemplo, siempre con argumentos elementales (pero muy ingeniosos), F. Mertens dedujo del trabajo de Chebyshev (en una forma más precisa):

$$\sum_{p \le x} \frac{\log p}{p} = \log x + O(1). \tag{2}$$

Parece, tras estos dos resultados, que (1) debería poder deducirse con algunos malabarismos técnicos. Sin embargo no es así (aunque después de Selberg uno podría reconsiderarlo).

El trabajo crucial sobre la distribución de los números primos es la celebérrima memoria [22] de B. Riemann de 1859. En ella se expresa $\pi(x)$ en términos de los ceros de la extensión meromorfa a $\mathbb{C}$ de la función ζ introducida al comienzo de esta

[1] Con $f \sim g$ indicamos $\lim f/g = 1$, es decir, que ambas funciones son asintóticamente iguales.

sección (véase [20]). A dicha extensión, que tiene un único polo simple en $s = 1$, se le llama *función ζ de Riemann.*

En 1896, J. Hadamard y C.J. de la Vallée-Poussin lograron independientemente completar casi todas las lagunas en el esquema esbozado por Riemann en su breve memoria, y obtener con técnicas de variable compleja información suficiente sobre los ceros de ζ como para deducir finalmente el teorema de los números primos.

2.2. UNA PRUEBA ELEMENTAL

No es difícil imaginar el argumento que flotaba en la comunidad matemática hasta la mitad del siglo XX: se sabía que el tamaño de $\pi(x) - \mathrm{li}(x)$ depende de la distribución de los ceros de la función ζ de Riemann. Por otro lado esta distribución parece un tema intrínsecamente de variable compleja, lo que limita a primera vista las posibilidades de pruebas elementales. En palabras de G.H. Hardy de 1921 (véase [11]):

> Una prueba de tal teorema que no dependa fundamentalmente de la teoría de funciones me parece extraordinariamente improbable. Es erróneo afirmar que un teorema matemático *no puede* ser probado de cierto modo pero hay una cosa que está clara: Tenemos cierta visión acerca de la lógica de la teoría, pensamos que ciertos teoremas 'son profundos' y otros más próximos a la superficie. Si alguien crea una prueba elemental del teorema de los números primos mostrará que esta visión está equivocada, que el tema no encaja de la forma que habíamos supuesto y que es hora de dejar a un lado los libros y reescribir la teoría.

Como veremos más adelante, la incidencia en la teoría no fue tal como Hardy pensaba. Lo cierto es que las pruebas elementales del teorema de los números primos que dieron P. Erdős [8] y Selberg [27] en 1949 causaron una gran sorpresa. Es bien conocido que el resultado estuvo empañado por una amarga disputa de prioridad (o más propia y sutilmente, de colaboración) entre ambos matemáticos. No nos detendremos aquí en este punto y remitimos al lector curioso al artículo de D. Goldfeld [11] que tiene una amplia documentación no publicada anteriormente. Ambas pruebas se basan en la *fórmula de Selberg*:

$$\theta(x)\log x + \sum_{p\le x}\theta(x/p)\log p = 2x\log x + O(x), \qquad \text{donde } \ \theta(x) = \sum_{p\le x}\log p,$$

a la que, según [2], llegó Selberg inicialmente con métodos de criba. Esta fórmula admite una prueba ingeniosa pero elemental y breve debida al propio Selberg [27] (en [16] hay un simplificación). Es bien conocido (y elemental y sencillo) que el teorema de los números primos es equivalente a $\theta(x) \sim x$. Empleando (2) está claro que esta asintótica es compatible con la fórmula de Selberg pero no está claro cómo "despejar" $\theta(x)$ a partir de ella para deducir la asintótica. Esto requiere un argumento tauberiano que se puede llevar a cabo de forma elemental en cada uno de sus pasos pero que es laborioso en su conjunto.

2.3. El resultado en perspectiva

Las demostraciones elementales fueron acogidas con mucho entusiasmo porque abrían la puerta a una manera novedosa de entender la distribución de los primos, e indirectamente la distribución de los ceros de la función ζ. Sin embargo, es indudable que el tema ha perdido fuerza y no se han cumplido las esperanzas más optimistas. Posiblemente, de los grandes resultados de Selberg, la demostración elemental del teorema de los números primos sea el que peor ha envejecido.

Goldfeld revela en [11] un fragmento de una carta de Selberg que ilustra su punto de vista 50 años después. En ella se refiere a la siguiente opinión mantenida por E.G. Strauss en un manuscrito no fechado, casi antitético a la previsión de Hardy:

> La prueba elemental no ha producido hasta ahora las innovaciones en teoría de números que muchos de nosotros esperábamos. Puede ser que el paso del tiempo pruebe que lo que presenciamos en 1948 fuera un avance brillante pero hasta cierto punto secundario, sin la importancia histórica que parecía tener.

La reacción de Selberg en la mencionada carta a Goldfeld datada en 1998 es:

> Estoy de acuerdo con esta última cita de Strauss (realmente yo nunca esperé ninguna revolución a partir de esto). Sin embargo la idea de la criba local ha producido muchas cosas que no se han logrado con otros métodos.

A pesar de las décadas transcurridas y de que varios autores han dado versiones más precisas de pruebas elementales (véanse [1], [6] y [33]), hasta ahora todas ellas son menos poderosas que la obtenida con variable compleja y conceptualmente más artificiosas. Otra razón que ha hecho perder fuerza a las pruebas elementales es la simplificación técnica que han conseguido algunos autores en la exposición de la prueba natural con variable compleja. Mencionamos aquí una demostración minimalista de D. Newman [19] que ha alcanzado cierta difusión. También destacamos una prueba sencilla y menos conocida debida a H. Iwaniec [16] que sólo utiliza la función ζ en $\Re s > 1$, sin apelar a su extensión ni a propiedades analíticas.

3. Métodos de criba

3.1. Ideas básicas

Dado un subconjunto finito de enteros positivos $\mathcal{A}$, uno de los problemas básicos que tratan los métodos de criba es la estimación del número de elementos de $\mathcal{A}$ que no tienen factores primos menores que z, en símbolos:

$$S(\mathcal{A},z) = \#\{a \in \mathcal{A} \;:\; \mathrm{mcd}(a, P(z)) = 1\}, \qquad \text{donde} \;\; P(z) = \prod_{p<z} p.$$

El objetivo real que se pretende es, al igual que en el bien conocido proceso de Eratóstenes, extraer primos grandes eliminando múltiplos de primos pequeños. Por ejemplo, si $\mathcal{A} = \{n \le N^2\}$ entonces $S(\mathcal{A}, N+1)$ cuenta, aparte del uno, los primos en $(N, N^2]$.

Otras veces se buscan números con una cantidad pequeña de factores primos. Por ejemplo, si $\mathcal{A} = \{n(n+2) \ : \ n < N-2\}$ los números de $\mathcal{A}$ sin factores $p < N^{1/2}$ cumplen que n y $n+2$ son primos; entonces $S(\mathcal{A}, N^{1/2})$ cuenta primos gemelos.

La información de la que se parte es la densidad aproximada $\rho(d)$ de los múltiplos de d en $\mathcal{A}$, es decir,

$$|\mathcal{A}_d| = |\mathcal{A}|\rho(d) + r_d \tag{3}$$

donde $\mathcal{A}_d = \{a \in \mathcal{A} \ : \ d|a\}$, $0 \le \rho(d) < 1$ y se espera que r_d sea pequeño. Pedimos además que $\rho(\cdot)$ sea una función multiplicativa, lo cual manifiesta la condición natural de que ser divisible por d_1 y ser divisible por d_2 son sucesos independientes si d_1 y d_2 no tienen factores comunes. Por ejemplo, si $\mathcal{A} = \{n \le N\}$ entonces $|\mathcal{A}_d| = [N/d]$ y tomaríamos $\rho(d) = 1/d$, con lo que se cumple $|r_d| < 1$.

Por el principio de inclusión-exclusión,

$$S(\mathcal{A}, z) = |\mathcal{A}| - |\mathcal{A}_2| - |\mathcal{A}_3| + |\mathcal{A}_6| - |\mathcal{A}_5| \cdots = \sum_{d|P(z)} \mu(d)|\mathcal{A}_d|$$

donde $\mu(\cdot)$ es la función de Möbius definida como $\mu(p_1p_2 \dots p_r) = (-1)^r$ si p_i son primos distintos, $\mu(1) = 1$, y $\mu(n) = 0$ en otro caso. Usando (3) se tiene

$$S(\mathcal{A}, z) = |\mathcal{A}| \sum_{d|P(z)} \mu(d)\rho(d) + O\Big(\sum_{d|P(z)} |r_d|\Big) = |\mathcal{A}| \prod_{p<z} \big(1 - \rho(p)\big) + O\Big(\sum_{d|P(z)} |r_d|\Big).$$

El problema es que el número de divisores de $P(z)$ crece exponencialmente con z y eso limita mucho las aplicaciones ya que en prácticamente todos los ejemplos de interés esperamos que $S(\mathcal{A}, z)$ crezca, como mucho, como un polinomio en z.

V. Brun en 1915 marcó el comienzo de los métodos de criba modernos cuando tuvo la idea de limitar el número de sumandos al aplicar el principio de inclusión-exclusión. En su versión más sencilla, el punto de partida es:

$$\sum_{\substack{p|P(z) \\ \nu(d)<2k}} \mu(d)|\mathcal{A}_d| \le S(\mathcal{A}, z) \le \sum_{\substack{p|P(z) \\ \nu(d)<2k+1}} \mu(d)|\mathcal{A}_d|$$

donde $\nu(d)$ indica el número de factores primos de d, y k es un número natural arbitrario. Por un lado no debemos escoger k muy grande para que no se acumulen muchos términos de error, y por otra parte no debemos escogerlo muy pequeño para que la aproximación sea precisa. Jugando con este parámetro k se puede probar por ejemplo, como hizo Brun, que si hay infinitos primos gemelos crecen al menos como $N(\log N/ \log\log N)^2$ y, de ahí, que la suma de los inversos de los primos gemelos converge (véase [21]).

Se llama *criba combinatoria* a la técnica basada en suprimir términos en el principio de inclusión-exclusión conservando desigualdades. En contra de lo que pueda parecer, es un tema realmente complicado arropado por una teoría muy profunda.

3.2. La criba de Selberg

Selberg introdujo un método de criba intrínsecamente sencillo. Aconsejamos al lector interesado el artículo expositorio [28], donde Selberg describe magistralmente, en apenas siete páginas sin ningún formalismo, varias ideas cruciales. Su criba no es combinatoria, en vez de omitirse términos en el principio de inclusión-exclusión se reemplazan los coeficientes $\mu(d)$ por otros números. Esto no parece tener sentido porque las propiedades de la función μ son esenciales al proceso de criba.

El núcleo del argumento de Selberg es la siguiente línea, donde $\lambda_1 = 1$ y el resto de los λ_n son arbitrarios y con soporte en $n < z$,

$$S(\mathcal{A}, z) \le \sum_{a\in\mathcal{A}} \Big(\sum_{d|a} \lambda_d\Big)^2 = \sum_{a\in\mathcal{A}} \sum_{d_1|a} \sum_{d_2|a} \lambda_{d_1}\lambda_{d_2} = \sum_{d_1,d_2} \lambda_{d_1}\lambda_{d_2} |\mathcal{A}_{\mathrm{mcm}(d_1,d_2)}|.$$

Las posibilidades para los λ_n dan lugar a una familia muy general de cribas; lo que hizo Selberg a continuación fue escoger “la mejor” de todas ellas. Teniendo en cuenta (3), se busca minimizar la forma cuadrática

$$Q(\vec{\lambda}) = \sum_{d_1,d_2} \rho(\mathrm{mcm}(d_1, d_2))\lambda_{d_1}\lambda_{d_2}$$

entre todos los posibles $\vec{\lambda} = (\lambda_n)_{n<z}$ como antes (por razones técnicas se impone también $\lambda_n = 0$ si n no es libre de cuadrados). En principio esto es un problema de álgebra lineal, o más bien bilineal, con una apariencia sospechosamente complicada por la naturaleza aritmética del mínimo común múltiplo. Selberg logró obtener la fórmula exacta

$$\min_{\vec{\lambda}} Q(\vec{\lambda}) = \Big({\sum_{n<z}}' h(n)\Big)^{-1}, \qquad \text{con} \quad h(n) = \prod_{p|n} \frac{\rho(p)}{1-\rho(p)},$$

donde la suma es sobre los libres de cuadrados.

Una ventaja añadida de la criba de Selberg es la forma bilineal del término de error que se ajusta al tipo de técnicas empleadas en teoría analítica de números. Por otro lado, hay variantes que permiten dar cotas inferiores para $S(\mathcal{A}, z)$ (véase [2] y [28]), pero no tienen la fuerza y la simplicidad de la cota superior: éste es un punto débil de la criba de Selberg.

No queremos cerrar esta sección sin mencionar el *problema de la paridad*. Un importante fenómeno teórico descubierto por Selberg que impide, en situaciones generales, bajo las hipótesis de los métodos de criba, distinguir un primo de un producto de dos primos. Éste es el límite al que se ha llegado con técnicas de criba en algunos problemas clásicos, por ejemplo J.-R. Chen probó, en [4], la conjetura de Goldbach para números pares suficientemente grandes si se admite que uno de los sumandos tenga a lo más dos factores primos. Por otro lado, en [10], un hito en los métodos de criba, se ha podido superar el problema de la paridad en un caso particular utilizando información adicional.

4. La función ζ de Riemann

4.1. Ceros y primos

Ya hemos mencionado que Riemann relacionó los ceros de la función ζ con la función $\pi(x)$ que cuenta los primos hasta x. Desde el punto de vista actual es más conveniente tratar con la siguiente función:

$$\psi(x) = \sum_{p^n \le x} \log p,$$

que se relaciona con $\pi(x)$ mediante

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

de modo que aproximar bien $\pi(x)$ por $\mathrm{li}(x)$ es consecuencia de aproximar bien $\psi(x)$ por x. Se tiene para $x \in [1, \infty) - \mathbb{Z}^+$ la *fórmula explícita*

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

donde ρ recorre los ceros de ζ en la *banda crítica* $0 < \Re s < 1$. Si damos por supuesto la convergencia (lo cual es un tema espinoso y realmente la fórmula es bastante inútil si no se trunca), inferimos que el tamaño de $\pi(x) - \mathrm{li}(x)$ depende de lo lejos que estén los ceros ρ de la recta $\Re s = 1$.

La función ζ goza de cierta simetría regida por la *ecuación funcional*

$$\pi^{-s/2}\Gamma(s/2)\zeta(s) = \pi^{-(1-s)/2}\Gamma\big((1-s)/2\big)\zeta(1-s), \tag{4}$$

de la que se deduce que si ρ es un cero como antes, entonces $1 - \rho$ también lo es. Por ello el menor error posible en el teorema de los números primos ocurriría si todos los ceros estuvieran en la *línea crítica* $\Re s = 1/2$. Ésta es la famosa *hipótesis de Riemann*, uno de los problemas abiertos más conocidos. Si se cumpliera, se tendría $\pi(x) - \mathrm{li}(x) = O(x^{1/2}\log x)$ y, en caso contrario, $\pi(x) - \mathrm{li}(x)$ sería de orden mayor.

Nuestro conocimiento acerca de la hipótesis de Riemann es débil, y la mejor región libre de ceros probada se aproxima asintóticamente a $\Re s = 1$; ni siquiera permite obtener $\pi(x) - \mathrm{li}(x) = O(x^\alpha)$ para ningún $\alpha < 1$.

4.2. La distribución de los ceros

La distribución vertical de los ceros de la función ζ en la banda crítica viene dada por

$$N(T) = \#\{\rho \,:\, \zeta(\rho) = 0,\ 0 < \Re\rho < 1,\ 0 < \Im\rho < T\}.$$

La ecuación funcional (4) permite que, al aplicar convenientemente el principio del argumento, se obtenga un término principal para $N(T)$:

$$N(T) = \frac{T}{2\pi}\log\frac{T}{2\pi e} + \frac{7}{8} + S(T) + O(T^{-1}),$$

donde $S(T) = O(\log T)$ proviene de la variación del argumento de ζ.

Selberg dio varios resultados en [25] y [26] acerca de $S(T)$. Por una parte demostró que para infinitos valores de T era mayor que cierta potencia de $\log T$ y, por otro lado, dedujo el comportamiento asintótico de sus momentos. Siguiendo [18], aunque no fuera el lenguaje empleado por Selberg, se pueden parafrasear parte de sus resultados diciendo que $\pi\sqrt{2}S(T)/\sqrt{\log\log T}$ tiende a seguir una distribución normal $N(0,1)$ cuando T crece.

Los avances más espectaculares los hizo Selberg en relación con los ceros en la línea crítica o cerca de ella. Demostró que una proporción positiva de los ceros contados en $N(T)$ están en la línea crítica y que con probabilidad nula dejan de estar extremadamente cerca de ella. En términos matemáticos, si se define $N(\sigma, T)$ como $N(T)$ pero con la condición añadida $\Re\rho \geq \sigma$, Selberg obtuvo, en [26],

$$N(\sigma, T) = O\big(T^{1-\frac{1}{4}(\sigma-\frac{1}{2})}\log T\big), \qquad \text{para } \ \sigma \geq 1/2.$$

Con otros métodos, 32 años después de [24], N.A. Levinson logró probar que más de un tercio de los ceros están en la línea crítica y B. Conrey ha extendido este valor a más del 40 %.

No resistimos la tentación de mencionar aquí brevemente algunas ideas en la prueba de Selberg.

Empleando (4) se deduce que

$$Z(t) = \frac{\Gamma(1/4 + it/2)}{\pi^{it/2}|\Gamma(1/4 + it/2)|}\,\zeta(1/2 + it)$$

es una función real. Obviamente, si $\rho = 1/2 + it_0$ es un cero en la línea crítica, $Z(t_0) = 0$, y viceversa. Si se cumpliera, para ciertos T y H,

$$\Big|\int_T^{T+H} Z(t)\,dt\Big| < \int_T^{T+H} |Z(t)|\,dt,$$

entonces existiría al menos un cero ρ en la línea crítica con $T < \Im\rho < T + H$. Esta estrategia para detectar ceros se debe a G.H. Hardy y a J.E. Littlewood (véase [13]). El inconveniente que muestra es que se pierde información si $\zeta(1/2+it)$, y por tanto $Z(t)$, presenta grandes variaciones. Para remediarlo, Selberg sustituyó $Z(t)$ por $Z(t)|\phi(1/2 + it)|^2$ donde $\phi(s)$ es una versión suavizada de los términos iniciales al desarrollar $\prod(1 - p^{-s})^{1/2}$. La idea es que formalmente $|\phi(1/2 + it)|^2$ se parece a $|\zeta(1/2 + it)|^{-1}$, y por tanto $Z(t)|\phi(1/2 + it)|^2$ debería imitar de algún modo a una función que fuera de los ceros sólo toma valores ± 1. Las dificultades para llevar a cabo este programa son grandes, y en el transcurso de su realización se observa el germen de lo que más adelante sería la criba de Selberg.

5. La teoría espectral de formas automorfas

5.1. Teoría espectral y fórmula de traza

El grupo de transformaciones

$$\mathrm{PSL}_2(\mathbb{Z}) = \Big\{\gamma(z) = \frac{az+b}{cz+d} \;:\; a,b,c,d \in \mathbb{Z},\ ad-bc=1\Big\},$$

actúa sobre los puntos del semiplano superior $\mathbb{H} = \{x+iy \;:\; x \in \mathbb{R},\ y \in \mathbb{R}^+\}$. Consideremos el conjunto de órbitas correspondiente que denotamos con $\mathrm{PSL}_2(\mathbb{Z})\backslash\mathbb{H}$. Se puede probar que cada órbita tiene un elemento en el *dominio fundamental*

$$D = \{z \in \mathbb{H} \;:\; |\Re z| \le 1/2,\ |z| \ge 1\}.$$

De hecho tiene exactamente uno si en D identificamos pares de puntos en la frontera relacionados por una simetría de eje $\Re z = 0$. Con ello $\mathrm{PSL}_2(\mathbb{Z})\backslash\mathbb{H}$ adquiere una estructura de variedad compleja unidimensional, es decir, de superficie de Riemann. Topológicamente, con la identificación indicada, es como una esfera con un tentáculo que se dirige al infinito. Empleando en $\mathbb{H}$ la métrica de Poincaré $ds^2 = y^{-2}(dx^2+dy^2)$, este tentáculo se va estrechando cada vez más y el área de la superficie es finita.

Incluso si consideramos Γ un subgrupo de índice finito de $\mathrm{PSL}_2(\mathbb{Z})$ o, más en general, un grupo fuchsiano de primera especie, $S = \Gamma\backslash\mathbb{H}$ es una superficie de Riemann de área finita.

Las funciones y formas diferenciales meromorfas en S constituyen las funciones y formas modulares clásicas. Pensemos ahora en términos analíticos: para desarrollar todas las funciones de $L^2(S)$, no necesariamente holomorfas, podemos utilizar las autofunciones de un operador elíptico de segundo orden. El más natural es el operador de Laplace-Beltrami $-\Delta$ en $\mathbb{H}$. Sus autofunciones normalizadas en $L^2(S)$ son las llamadas *formas de Maass*. Si S es no compacta hay un espectro continuo que se superpone con el espectro discreto, y para analizar funciones necesitamos también considerar integrales de unas familias de funciones asociadas a los puntos del infinito llamadas *series de Eisenstein espectrales*.

El teorema espectral correspondiente en S no es sencillo en absoluto. La continuación analítica de las series de Eisenstein espectrales resulta ser un punto crucial, y especialmente complicado si S no tiene cierta estructura aritmética. Selberg supo tratarlo dando hasta tres pruebas.

La fórmula de traza de Selberg es una relación explícita entre el espectro del operador de Laplace-Beltrami (análisis) y las longitudes de las geodésicas (geometría). Aunque sea explícita, tiene un aspecto suficientemente complicado como para no incluir el enunciado exacto aquí (véase [15]). Preferimos dar una versión incompleta pero ilustrativa:

$$\sum_j h(t_j) = \frac{|S|}{4\pi}\int_{-\infty}^{\infty} th(t)\tanh(\pi t)\,dt + \frac{1}{4\pi}\sum_l \frac{l^*}{\operatorname{senh}(l/2)}\,\widehat{h}(l) + \cdots,$$

donde h es una función par que satisface ciertas condiciones de regularidad, $\widehat{h}(\xi) = \int_{-\infty}^{\infty} h(t)e^{-i\xi t}\,dt$, $|S|$ es el área de S, los t_j son tales que $1/4+t_j^2$ es un autovalor,

l son las longitudes de las geodésicas cerradas recorridas un número arbitrario de veces y finalmente, para cada l, l^* indica la longitud cuando la geodésica se recorre una sola vez (es decir $l = kl^*$ con $k \in \mathbb{Z}^+$). Los términos omitidos en la fórmula anterior tienen que ver con los puntos del infinito y con otros puntos especiales de S.

El impacto de la fórmula de traza ha sido notorio en varias áreas que van de la teoría de números a la física (M.C. Gutzwiller afirma en [12]: *"La fórmula de la traza de Selberg responde de la manera más clara a una pregunta que Einstein hizo en 1917"*). Anticipó importantes resultados más generales sobre fórmulas de sumación de Poisson en variedades (véase [3] y [7]), pero a diferencia de ellos es perfectamente explícita. La teoría espectral, en general, se ha revelado un arma poderosa en el terreno aritmético. Según P. Sarnak, refiriéndose a [29]: *"Éste es uno de los artículos matemáticos con mayor influencia del siglo XX"*.

Selberg se ocupó también de análogos en dimensiones superiores de las superficies de Riemann que hemos tratado aquí, y probó un teorema totalmente inesperado. Intuitivamente afirma que, para dimensiones mayores, las únicas posibilidades tienen un sentido aritmético, o más propiamente algebraico. Concretamente, si $n > 2$ los subgrupos Γ de $\mathrm{SL}_n(\mathbb{R})$ con $\Gamma\backslash\mathrm{SL}_2(\mathbb{R})$ compacto son conjugados de grupos de matrices que tienen números algebraicos como elementos. En particular estos objetos son *rígidos*, no puede haber familias continuas de ellos que no se reduzcan a la conjugación. Además probó, bajo ciertas condiciones, que esta propiedad también se aplicaba al caso no compacto.

5.2. La idea tras la fórmula de traza

Muchas veces se dice que la fórmula de traza de Selberg es la fórmula de sumación de Poisson en superficies de Riemann. Para explicar esta afirmación veremos primero una demostración extraña de la fórmula de sumación de Poisson clásica:

$$\sum_{n=-\infty}^{\infty} f(n) = \sum_{n=-\infty}^{\infty} \widehat{f}(2\pi n), \qquad \text{donde } \ \widehat{f}(\xi) = \int_{-\infty}^{\infty} f(t)e^{-it\xi}\, dt.$$

Consideramos en $\mathbb{R}$ el grupo de traslaciones enteras $\Gamma = \{\gamma \ : \ \gamma(x) = x + n,\ n \in \mathbb{Z}\}$. Evidentemente el conjunto de órbitas $\Gamma\backslash\mathbb{R}$ se puede identificar con el toro unidimensional usual $\mathbb{T}$. Tomemos una función regular $f : \mathbb{R} \longrightarrow \mathbb{C}$ que decaiga bien y por razones técnicas supongamos que f es par (la fórmula de sumación de Poisson es trivial para funciones impares). Entonces

$$F(x, y) = \sum_{\gamma\in\Gamma} f(d(x, \gamma y)),$$

con $d(x, y) = |x - y|$ la distancia usual en $\mathbb{R}$, está bien definida y es invariante por cualquier elemento $\gamma \in \Gamma$ en cada una de sus variables, $F(\gamma x, y) = F(x, y) = F(x, \gamma y)$. Podemos entonces considerar F definida en $\mathbb{T} \times \mathbb{T}$ y desarrollar en serie de Fourier, es decir, con las autofunciones $\phi_n(t) = e^{2\pi int}$ del operador de Laplace-Beltrami $-\Delta$ en $\mathbb{T}$ que tienen autovalores $\lambda_n = 4\pi^2 n^2$. Con ello se deduce

$$\sum_{\gamma\in\Gamma} f(d(x, \gamma y)) = \sum_{n} \widehat{f}\big(\sqrt{\lambda_n}\big)\phi_n(x)\overline{\phi_n(y)}.$$

Si calculamos la traza, es decir, si integramos sobre la diagonal $x = y$ en $\mathbb{T}$, obtenemos

$$\sum_{\gamma\in\Gamma}\int_0^1 f(d(x,\gamma x))\,dx = \sum_n \widehat{f}\big(\sqrt{\lambda_n}\big).$$

Nótese que $[0,1]$ es un "dominio fundamental" de $\mathbb{T}$, y que podemos pensar que en el segundo miembro hemos usado que las autofunciones forman un sistema ortonormal completo, aunque en este ejemplo de juguete todo es más sencillo.

A cada elemento de Γ distinto de la identidad le podemos asociar biyectivamente una geodésica cerrada en $\mathbb{T} = \mathbb{R}/\Gamma$, simplemente asignando a $\gamma \in \Gamma - \{\mathrm{Id}\}$ la proyección sobre $\mathbb{T}$ del segmento que une 0 y $\gamma(0)$ en $\mathbb{R}$ (recorrido en sentido contrario si $\gamma(0) < 0$). La longitud l_γ de la geódesica que corresponde a $\gamma(x) = x + n$ es $|n|$. Nótese que ocurre la casualidad de que $d(x,\gamma x)$ es independiente de x, de hecho esta cantidad es l_γ. Entonces

$$f(0) + \sum_{\gamma\in\Gamma-\{\mathrm{Id}\}} f(l_\gamma) = \sum_n \widehat{f}\big(\sqrt{\lambda_n}\big). \tag{5}$$

Esto no es más que una forma enrevesada de escribir la fórmula de sumación de Poisson en $\mathbb{R}$, pero que ahora se lee como una relación explícita entre las longitudes de las geodésicas y el espectro.

Veamos ahora cuáles son las dificultades para extender el esquema anterior a las superficies de Riemann $S = \Gamma\backslash\mathbb{H}$.

El punto de partida es similar, se analiza

$$F(z,w) = \sum_{\gamma\in\Gamma} f(\rho(z,\gamma w)),$$

donde ρ es la distancia correspondiente a la métrica de Poincaré en $\mathbb{H}$. Como esta función es Γ invariante en ambas variables, en el desarrollo espectral apareceran productos de formas de Maass y productos de series de Eisenstein espectrales, con coeficientes dados por una cierta transformada de Fourier no euclídea.

En primer lugar está el problema del espectro continuo. Las series de Eisenstein espectrales no pertenecen a $L^2(S)$, por ello el cálculo de la traza llevará a una cantidad infinita. La solución es trabajar con aproximaciones L^2 dependiendo de un parámetro. En ambos miembros de la fórmula de traza se obtendrán cantidades que tienden a infinito con dicho parámetro pero cuyos términos principales se cancelan, resultando a la postre una contribución finita.

La segunda dificultad es que no hay nada análogo a que $d(x,\gamma x)$ sea independiente de x, éste es un fenómeno puramente abeliano. Selberg agrupó la suma sobre $\gamma \in \Gamma$ en sumas sobre cada una de las diferentes clases de conjugación, y logró dotar a cada una de ellas de un significado geométrico (al deducir (5) pudimos tratar cada γ por separado porque las clases de conjugación son triviales en el caso abeliano). Dado $\gamma \in \Gamma$, sean C y Z su clase de conjugación y su centralizador:

$$C = \{\alpha^{-1}\gamma\alpha \;:\; \alpha \in \Gamma\} \qquad \text{y} \qquad Z = \{\alpha \in \Gamma \;:\; \alpha^{-1}\gamma\alpha = \gamma\}.$$

La parte de la traza que corresponde a los elementos de C es

$$\sum_{\beta\in C}\int_D f(\rho(z,\beta z))\,d\mu = \sum_{\alpha\in Z\setminus C}\int_D f(\rho(z,\alpha^{-1}\gamma\alpha z))\,d\mu,$$

donde D es un dominio fundamental y $d\mu$ es el elemento de área. La invariancia implica $\rho(z,\alpha^{-1}\gamma\alpha z)=\rho(\alpha z,\gamma\alpha z)$, y con un cambio de variable $\alpha z\mapsto z$ se deduce (usando la invariancia de la medida)

$$\sum_{\beta\in C}\int_D f(\rho(z,\beta z))\,d\mu = \int_{Z\setminus\mathbb{H}} f(\rho(z,\gamma z))\,d\mu.$$

Con ello hemos conseguido expresar la contribución de cada clase de conjugación en términos de una sola integral. La razón que subyace a que se pueda calcular explícitamente en términos geométricos es que ρ y $d\mu$ son de hecho invariantes por todos los elementos de $\mathrm{PSL}_2(\mathbb{R})$, definido como $\mathrm{PSL}_2(\mathbb{Z})$ pero con $a,b,c,d\in\mathbb{R}$. Y las clases de conjugación en $\mathrm{PSL}_2(\mathbb{R})$ son sencillas, aparte de la identidad sólo hay tres: elementos que son como dilataciones (hiperbólicos), como traslaciones (parabólicos) o como giros (elípticos); fácilmente reconocibles por el valor de $a+c$.

6. Breve glosario adicional

A pesar de la brevedad de la obra de Selberg, hay temas sin tratar en las secciones anteriores. Intentamos cubrir esta deficiencia incluyendo como suplemento comentarios muy sucintos sobre algunos términos matemáticos que han quedado asociados al nombre de Selberg.

Convolución de Rankin-Selberg [23]. Es una función L introducida independientemente por R.A. Rankin y Selberg, cuyos coeficientes son el producto de los coeficientes de formas modulares, salvo conjugación. Sus propiedades analíticas, notablemente su ecuación funcional, permiten extraer información acerca de dichos coeficientes.

Conjetura de Selberg [30]. Si se considera $\Gamma_0(q)$ definido como los elementos de $\mathrm{PSL}_2(\mathbb{Z})$ para los cuales c es divisible por q, entonces la conjetura de Selberg afirma que, para la superficie de Riemann $\Gamma_0(q)\setminus\mathbb{H}$, los autovalores ($\lambda\neq 0$) del operador de Laplace-Beltrami satisfacen $\lambda\geq 1/4$. El propio Selberg probó $\lambda\geq 3/16$. Nótese que $\lambda_j=1/4+t_j^2<1/4$ implica $t_j\notin\mathbb{R}$, y esto produce términos peculiares en la fórmula de traza.

Función zeta de Selberg [29]. Es una función asociada a las geodésicas que, en cierto modo, imita a la función ζ de Riemann asociada a los primos. Satisface una ecuación funcional y, sorprendentemente, una consecuencia de que el operador de Laplace-Beltrami sea autoadjunto es un análogo de la hipótesis de Riemann. Esto se acerca al llamado *sueño de Hilbert y Pólya*: una supuesta interpretación espectral de los ceros de ζ que resolviera la hipótesis de Riemann.

Fórmula de Chowla-Selberg [5]. Consiste en una representación de la serie de Dirichlet correspondiente a una forma cuadrática como una serie rápidamente convergente. Con ella S. Chowla y Selberg dedujeron fórmulas exactas para valores especiales de funciones elípticas. Como curiosidad, [5] es el único trabajo de Selberg escrito en colaboración. Los resultados allí contenidos fueron anunciados en una breve nota sin pruebas en 1949, pero [5] no apareció hasta 1967 (hay quien dice que Chowla estuvo rogando a Selberg casi 20 años).

Clase de Selberg [32]. No deja de ser un misterio hoy por hoy que la función ζ, las funciones L de Dirichlet, las funciones L globales de curvas elípticas y tantas otras, satisfagan conjeturalmente una hipótesis de Riemann. La clase de Selberg abstrae axiomáticamente las propiedades comunes de todos estos ejemplos y conjetura ciertas consecuencias. En [32], después de repasar sus investigaciones, Selberg señala: "he llegado a formular conjeturas algunas de las cuales son quizá de más interés que cualquiera de los resultados que pude probar".

Función de Beurling-Selberg. Inicialmente descubierta por A. Beurling y redescubierta por Selberg en relación con métodos de criba, es una función de variable compleja con propiedades extremales. En cierto sentido es la función entera que mejor aproxima a la función signo en el eje real y es mayor que ella.

Integral de Selberg. En su único artículo de investigación escrito en su lengua natal (noruego), Selberg evaluó una complicada familia de integrales múltiples. Este resultado aparentemente anecdótico (publicado en una revista dedicada en parte a la docencia) fue rescatado años después por su interés en la teoría de matrices aleatorias.

Agradecimientos: Quisiera expresar mi gratitud a J. Jiménez y a A. Ubis por su lectura del manuscrito. También deseo agradecer especialmente el apoyo de E. Valenti.

Referencias

[1] E. Bombieri, Sulle formule di A. Selberg generalizzate per classi di funzioni aritmetiche e le applicazioni al problema del resto nel "Primzahlsatz", *Riv. Mat. Univ. Parma (2)* **3** (1962), 393–440

[2] E. Bombieri, Selberg's sieve and its applications, en *Number theory, trace formulas and discrete groups (Oslo, 1987)*, 29–48, Academic Press, Boston, MA, 1989.

[3] J. Chazarain, Formule de Poisson pour les variétés riemanniennes, *Invent. Math.* **24** (1974), 65–82.

[4] 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.

[5] S. Chowla y A. Selberg, On Epstein's zeta-function, *J. Reine Angew. Math.* **227** (1967), 86–110.

[6] H.G. Diamond y J. Steinig, An elementary proof of the prime number theorem with a remainder term, *Invent. Math.* **11** (1970), 199–258.

[7] J.J. DUISTERMAAT Y V.W. GUILLEMIN, The spectrum of positive elliptic operators and periodic bicharacteristics, *Invent. Math.* **29** (1975), no. 1, 39–79.

[8] P. ERDŐS, On a new method in elementary number theory which leads to an elementary proof of the prime number theorem, *Proc. Nat. Acad. Sci. U.S.A.* **35** (1949), 374–384.

[9] L. EULER, Variae observationes circa series infinitas. (Hay una traducción inglesa por P. Viader y L. Bibiloni. Véase *The Euler Archive*, `http://www.math.dartmouth.edu/~euler/`).

[10] J. FRIEDLANDER Y H. IWANIEC, The polynomial $X^2 + Y^4$ captures its primes, *Ann. of Math. (2)* **148** (1998), no. 3, 945–1040.

[11] D. GOLDFELD, The elementary proof of the prime number theorem: an historical perspective, en *Number theory (New York, 2003)*, 179–192, Springer, New York, 2004.

[12] M.C. GUTZWILLER, Physics and Selberg's trace formula, en *The Selberg trace formula and related topics*, Contemp. Math., 53, Amer. Math. Soc., Providence, RI, 1986, 215–251.

[13] G.H. HARDY Y J.E. LITTLEWOOD, The zeros of Riemann's zeta-function on the critical line, *Math. Z.* **10** (1921), no. 3-4, 283–317.

[14] Institute for Advanced Study: Press Releases, `http://www.ias.edu/newsroom/announcements/view/1186683853.html`

[15] H. IWANIEC, *Spectral methods of automorphic forms* (second edition), Graduate Studies in Mathematics, 53, American Mathematical Society, Providence, RI; Revista Matemática Iberoamericana, Madrid, 2002.

[16] H. IWANIEC Y E. KOWALSKI, *Analytic number theory*, American Mathematical Society Colloquium Publications, 53, American Mathematical Society, Providence, RI, 2004.

[17] H. MAASS, Über eine neue Art von nichtanalytischen automorphen Funktionen und die Bestimmung Dirichletscher Reihen durch Funktionalgleichungen, *Math. Ann.* **121** (1949), 141–183.

[18] H.L. MONTGOMERY, Selberg's work on the zeta-function, en *Number theory, trace formulas and discrete groups (Oslo, 1987)*, 157–168, Academic Press, Boston, MA, 1989.

[19] D.J. NEWMAN, *Analytic number theory*, Graduate Texts in Mathematics, 177, Springer-Verlag, New York, 1998.

[20] J. QUER, La funció ζ de Riemann. Disponible provisionalmente en la web `http://www-ma2.upc.edu/~quer/docencia/Article.pdf`. Aparecerá en el Volum V, Curs Riemann 2007–2008, Facultat de Matemàtiques i Estadística, Universitat Politècnica de Catalunya.

[21] H. RADEMACHER, *Lectures on elementary number theory*, Robert E. Krieger Publishing Co., Huntington, N.Y., 1977.

[22] B. RIEMANN, *Riemanniana Selecta* (edición, estudio introductorio y notas de José Ferreirós), Colección Clásicos del Pensamiento, Madrid, CSIC, 2000.

[23] *A. Selberg*, Bemerkungen über eine Dirichletsche Reihe, die mit der Theorie der Modulformen nahe verbunden ist, *Arch. Math. Naturvid.* **43** (1940), 47–50.

[24] A. SELBERG, On the zeros of Riemann's zeta-function, *Skr. Norske Vid. Akad. Oslo* I (1942), no. 10, 5 pp.

[25] A. SELBERG, On the remainder in the formula for $N(T)$, the number of zeros of $\zeta(s)$ in the strip $0 < t < T$, *Avh. Norske Vid. Akad. Oslo* I (1944), no. 1, 27 pp.

[26] A. SELBERG, Contributions to the theory of the Riemann zeta-function, *Arch. Math. Naturvid.* **48** (1946) no. 5, 89–155.

[27] A. SELBERG, An elementary proof of the prime-number theorem, *Ann. of Math. (2)* **50** (1949), 305–313.

[28] A. SELBERG, The general sieve-method and its place in prime number theory, en *Proceedings of the International Congress of Mathematicians (Cambridge, Mass., 1950)*, vol. 1, pp. 286–292, Amer. Math. Soc., Providence, R.I., 1952.

[29] A. SELBERG, Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series, *J. Indian Math. Soc.* (N.S.) **20** (1956), 47–87.

[30] A. SELBERG, On the estimation of Fourier coefficients of modular forms, en *Proc. Sympos. Pure Math.*, Vol. VIII, pp. 1–15, Amer. Math. Soc., Providence, R.I., 1965.

[31] A. SELBERG, On discontinuous groups in higher-dimensional symmetric spaces, en *Contributions to function theory (Internat. Colloq. Function Theory, Bombay, 1960)*, pp. 147–164, Tata Institute of Fundamental Research, Bombay, 1960.

[32] A. SELBERG, Old and new conjectures and results about a class of Dirichlet series, en *Proceedings of the Amalfi Conference on Analytic Number Theory (Maiori, 1989)*, 367–385, Univ. Salerno, Salerno, 1992.

[33] E. WIRSING, Elementare Beweise des Primzahlsatzes mit Restglied. II, *J. Reine Angew. Math.* **214/215** (1964), 1–18.

FERNANDO CHAMIZO LORENTE, DPTO. DE MATEMÁTICAS, FACULTAD DE CIENCIAS, UNIVERSIDAD AUTÓNOMA DE MADRID, 28049-CANTOBLANCO (MADRID)

Correo electrónico: `fernando.chamizo@uam.es`

Página web: `http://www.uam.es/fernando.chamizo`