

Distribuciones aritméticas: errores y carreras

por

Joan-C. Lario

RESUMEN. En este artículo pretendemos ilustrar un fenómeno que se presenta de modo recurrente en Teoría de Números. Cuando se quieren contar cantidades donde los números primos juegan un papel destacado, se plantea el problema de analizar una estimación principal y controlar el error cometido. Con este fin, se suelen usar funciones L apropiadas para cada caso: sus polos, su no anulación en el punto central y sus ceros no triviales en la recta crítica describen el comportamiento del error. Los ejemplos escogidos abarcan diversos frentes aritméticos: distribuciones y carreras de primos en sucesiones aritméticas, representaciones de enteros positivos como suma de cuadrados, repartición de automorfismos de Frobenius en extensiones de Galois sobre cuerpos de números, número de puntos en las reducciones en característica positiva de curvas elípticas, así como la conjetura de Sato-Tate sobre variedades abelianas. El artículo se basa en la exposición sobre el mismo tema presentada en las Quintas Jornadas de Teoría de Números celebradas en la Universidad de Sevilla en 2013.

INTRODUCCIÓN

Desde Eratóstenes, en el oficio de la Teoría de Números, tenemos la costumbre de tratar problemas de contar cantidades donde están involucrados los números primos. Son problemas del tipo:

- (I) Dado un número primo p , calcular el número $\mathcal{N}(p)$ de ciertos objetos que dependen del primo p en cuestión.
- (II) Dado un número real x , calcular el número $\Pi(x)$ de números primos $p \leq x$ que cumplen cierta propiedad.

Habitualmente, en ambos casos, no disponemos de expresiones exactas para $\mathcal{N}(p)$ y $\Pi(x)$, o bien éstas son fórmulas explícitas bastante complicadas. Entonces debemos conformarnos con estimaciones; se persigue hallar un término principal (o dominante) y un término de error:

$$\begin{aligned}\mathcal{N}(p) &= \text{Principal}(p) + \text{Error}(p), \\ \Pi(x) &= L(x) + O(f(x)).\end{aligned}$$

Pongamos algunos ejemplos para ser más precisos:

$$\begin{aligned} \mathcal{N}(p) &= \# \tilde{E}(\mathbb{F}_p), \\ \mathcal{N}(p) &= \# \{(x_1, \dots, x_k) \in \mathbb{Z}^k : p = x_1^2 + \dots + x_k^2\}, \\ \Pi(x) &= \pi(x; a, N) = \# \{p \leq x : p \equiv a \pmod{N}\}, \\ \Pi(x) &= \pi(x; \mathcal{C}, K/\mathbb{Q}) = \# \{p \leq x : \text{Frob}_p = \mathcal{C}\}. \end{aligned}$$

En el primer ejemplo, $\# \tilde{E}(\mathbb{F}_p)$ denota el número de puntos de la reducción módulo p de una curva elíptica E/\mathbb{Q} . El segundo consiste en contar de cuántas maneras un número primo p puede expresarse como suma de k cuadrados. En el tercer ejemplo, queremos estudiar cómo se distribuyen los números primos en progresiones aritméticas módulo N . En el cuarto, K/\mathbb{Q} denota una extensión finita de Galois, \mathcal{C} una clase de conjugación del grupo de Galois $\text{Gal}(K/\mathbb{Q})$ y Frob_p es la clase del automorfismo de Frobenius asociado al primo p .

Estos ejemplos, y algunos más, serán tratados en este artículo. Como hemos indicado, siempre el negocio consiste en hallar un candidato óptimo a término principal, de modo que tengamos control del término de error y éste sea razonablemente pequeño. Se considera que la aproximación es buena siempre que el error sea, como mucho, del orden de la raíz cuadrada del tamaño del dato de entrada. A este problema de errores le añadiremos otro más *competitivo*: se trata de las carreras de primos, donde se estudia a qué velocidad los primos acuden a su distribución asintótica dependiendo del problema de conteo que estemos analizando. En las secciones siguientes abordaremos distintas distribuciones aritméticas, errores y carreras, agrupadas en diferentes apartados:

- Números primos
- Progresiones aritméticas
- Teoría de Galois
- Curvas elípticas
- Formas modulares
- La conjetura de Sato-Tate
- Variedades abelianas

El lector debe ser advertido de que el andamiaje del presente artículo proviene de una charla impartida en las Quintas Jornadas de Teoría de Números, celebradas en la Universidad de Sevilla (2013). Se ha pretendido seguir el estilo de la exposición, primando resaltar el hilo conductor y las ideas subyacentes por encima de los detalles, aunque sin faltar al debido rigor. También quiero agradecer la lectura atenta de una versión previa por parte de mis colegas Francesc Fité, Eugenio Jesús Gómez Ayala, Mariángeles Gómez Molleda y Jordi Quer.

1. NÚMEROS PRIMOS

Como es costumbre, denotamos por $\pi(x)$ la función que cuenta el número de primos inferiores a un número real x dado:

$$\pi(x) = \#\{p \leq x: p \text{ primo}\}.$$

Claramente se trata de una función creciente. Pero, ¿cómo crece la función $\pi(x)$? Euclides prueba que la función $\pi(x)$ es no acotada superiormente: hay números primos tan grandes como se quiera. El mayor número primo conocido hasta la fecha es $2^{82\,589\,933} - 1$; tiene 24 862 048 dígitos decimales y fue obtenido en diciembre de 2018 por el proyecto comunitario GIMPS, *Great Internet Mersenne Prime Search* (ver [24]).

La tabla de logaritmos que usaba Gauss en su adolescencia contaba con un reverso donde figuraba una lista ordenada de números primos. A sus 14 años de edad, Gauss deja anotado en su cuaderno el comportamiento asintótico de $\pi(x)$ en función del logaritmo neperiano:

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

Es decir,

$$\lim_{x \rightarrow \infty} \frac{\pi(x)}{x/\log x} = 1,$$

o, si se prefiere, la probabilidad de que un entero positivo n sea primo es aproximadamente $1/\log(n)$. Así, Gauss proporciona (conjeturalmente) un primer candidato a término principal para la función $\pi(x)$. Su obsesión por listar y contar números primos le persiguió de tal modo a lo largo de su vida que, posteriormente, afina (conjeturalmente) el término principal:

$$\pi(x) \sim \text{Li}(x) = \int_2^x \frac{dt}{\log t}.$$

Chebyshev (1848 y 1850) trató de demostrar esta ley asintótica de la distribución de los números primos mediante la función $\zeta(s)$ —prolongación meromorfa de la serie $\sum_{n=1}^{\infty} 1/n^s$ a todo el plano complejo—, al uso de Euler (1737), precediendo la célebre memoria de Riemann (1859). Chebyshev logra probar que, si el límite existe, entonces es necesariamente igual a uno. También es capaz de demostrar que el cociente $\pi(x)/\text{Li}(x)$ está acotado superior e inferiormente por constantes explícitas cercanas a 1, para todo x suficientemente grande. Sus resultados parciales le valen para demostrar el postulado de Bertrand, que afirma la existencia de un número primo entre n y $2n$ para cualquier entero $n \geq 2$.

La demostración de la conjetura de Gauss, hoy conocida como teorema de los números primos, tuvo que esperar hasta el año 1896, cuando confluyeron dos demostraciones obtenidas de forma independiente por Hadamard y de la Vallée-Poussin. Ambas demostraciones se basaron en ideas de Riemann (ver [13]).

Precisamente fue Riemann quien se encargó de estudiar el término de error al aproximar $\pi(x)$ por $\text{Li}(x)$. Concretamente, Riemann obtiene en [14] la *fórmula explícita*

$$\begin{aligned} \pi(x) &= \text{Li}(x) - \text{Error}(x) \\ &= \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \left(\text{Li}\left(\frac{x}{n}\right) - \sum_{\rho} \text{Li}\left(x \frac{e^{-\rho}}{n}\right) + \int_{x^{1/n}}^{\infty} \frac{du}{u(u^2-1)\log u} \right), \end{aligned}$$

donde μ denota la función de Moebius y ρ recorre los ceros no triviales de la función $\zeta(s)$ de Riemann. Por ceros no triviales entendemos los números complejos ρ con $0 < \text{Re}(\rho) < 1$ tales que $\zeta(\rho) = 0$.

La famosa hipótesis de Riemann predice que los ceros no triviales deben hallarse todos en la recta crítica $\text{Re}(s) = 1/2$. La justificación y virtud de la hipótesis de Riemann reside en la optimalidad de la aproximación de $\pi(x)$ por $\text{Li}(x)$, pues proporciona una cota deseable para el error:

$$\pi(x) = \text{Li}(x) + O(x^{1/2+\varepsilon}),$$

para cualquier $\varepsilon > 0$.

La tabla siguiente muestra cómo la función logaritmo integral $\text{Li}(x)$ aproxima a $\pi(x)$ para distintos valores de x :

x	$\pi(x)$	$\text{Li}(x) - \pi(x)$
10^8	5761455	753
10^9	50847534	1700
10^{10}	455052511	3103
10^{11}	4118054813	11587
10^{12}	37607912018	38262
10^{13}	346065536839	108970
10^{14}	3204941750802	314889
10^{15}	29844570422669	1052618
10^{16}	279238341033925	3214631

Gauss, sin disponer de una base de datos extensa, debió preguntarse si $\text{Li}(x)$ siempre se aproxima a $\pi(x)$ por exceso. Littlewood (1912) muestra que no siempre debemos guiarnos por la intuición, pues prueba que las gráficas de $\text{Li}(x)$ y $\pi(x)$ se cruzan infinitas veces cuando x tiende a infinito. Sin embargo, el primer valor en que esto ocurre no ha sido encontrado todavía. El discípulo de Littlewood, Skewes (1933), prueba que existe un cambio de tendencia para algún valor $x < 10^{10^{34}}$ si asumimos la hipótesis de Riemann (y una cota superior, también explícita, en caso contrario). Esta cota se obtiene a partir de la expresión del error para el cálculo de $\pi(x)$ obtenida por Riemann y ha sido progresivamente mejorada al disponer de más capacidad de cálculo para localizar ceros no triviales de la función $\zeta(s)$ en la recta crítica.

En resumen, la función $\zeta(s)$ de Riemann parece estar muy ligada a la función $\pi(x)$: su polo en $s = 1$ muestra que $\pi(x)$ no está acotada superiormente (Euclides),

su no anulación en $\operatorname{Re}(s) = 1$ implica el teorema de los números primos (conjetura de Gauss, probada por Hadamard y de la Vallée-Poussin) y, finalmente, la localización de sus ceros no triviales explica el ajuste deseado del término de error (hipótesis de Riemann, problema abierto todavía).

En todas las secciones que siguen, se reproduce el mismo esquema: a cierto problema de conteo con intervención de los números primos le corresponden una o más funciones de variable compleja. El comportamiento de estas funciones (prolongación analítica o meromorfa según el caso, su no anulación en cierto punto y sus ceros no triviales en la recta crítica) se traduce en aproximaciones óptimas y errores admisibles en el problema de conteo.

2. PROGRESIONES ARITMÉTICAS

Sean N y a enteros ≥ 1 y primos entre sí. Denotamos por

$$\pi(x; a, N) = \#\{p \leq x : p \text{ primo y } p \equiv a \pmod{N}\}$$

la función que cuenta cuántos números primos menores que el número real x hallamos en la progresión aritmética $\{a + \lambda N : \lambda \in \mathbb{Z}\}$.

Dirichlet (1837) prueba que $\pi(x; a, N)$ no está acotada superiormente. Hadamard y de la Vallée-Poussin (1896) establecen la estimación asintótica

$$\pi(x; a, N) \sim \frac{1}{\varphi(N)} \operatorname{Li}(x),$$

que generaliza el teorema de los números primos. En particular, los primos se equidistribuyen uniformemente en las distintas clases de congruencia invertibles módulo N . En el límite, todos los primos se reparten con la misma proporción $1/\varphi(N)$ en cada una de las clases. Sin embargo, como veremos, algunas clases se rellenan de primos con más rapidez que otras de forma general —aunque con infinitas excepciones—, fenómeno conocido como sesgo de Chebyshev y que, en cierto modo, generaliza las oscilaciones del signo en $\pi(x) - \operatorname{Li}(x)$ estudiadas por Littlewood.

Para probar la generalización del teorema de los números primos para el caso de progresiones aritméticas, deben considerarse varias (todavía en número finito para esta sección) funciones L que generalizan de manera natural la función ζ de Riemann. Concretamente, para cada carácter $\chi: (\mathbb{Z}/N\mathbb{Z})^* \rightarrow \mathbb{C}^*$ se define

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

que, cuando $\chi = \chi_0$ es el carácter trivial, se trata esencialmente de la función ζ de Riemann (salvo un factor que es una función entera). La demostración de la

distribución asintótica se basa en el cálculo

$$\begin{aligned} \sum_{p \equiv a \pmod N} \frac{1}{p^s} &= \sum_{\chi} \frac{\chi(a^{-1})}{\varphi(N)} \sum_p \frac{\chi(p)}{p^s} = \sum_{\chi} \frac{\chi(a^{-1})}{\varphi(N)} \log L(\chi, s) + O(1) \\ &= \frac{1}{\varphi(N)} \left(\log L(\chi_0, s) + \sum_{\chi \neq \chi_0} \chi(a^{-1}) \log L(\chi, s) \right) + O(1) \end{aligned}$$

y el punto clave, parecido al caso de $\pi(x)$, está en probar que, si $\chi \neq \chi_0$, entonces $L(\chi, s)$ es holomorfa en $\operatorname{Re}(s) \geq 1$ y se cumple la no anulación:

$$L(\chi, 1) \neq 0.$$

En este caso, la fórmula explícita que proporciona el error se debe a Rubinstein-Sarnak (1994). Si escribimos

$$\pi(x; a, N) = \frac{\operatorname{Li}(x)}{\varphi(N)} - \operatorname{Error}(x),$$

y suponemos ciertas (ver [16]) la Hipótesis de Riemann Generalizada (GRH) y la Gran Hipótesis de Simplicidad (GSH), se tiene

$$\operatorname{Error}(x) = \frac{\sqrt{x}}{\log x} \left(\frac{c(a, N)}{\varphi(N)} + \sum_{\chi \neq \chi_0} \bar{\chi}(a) \sum_{\gamma_{\chi}} \frac{\sin(\gamma_{\chi} \log x)}{\gamma_{\chi}} \right),$$

donde

$$c(a, N) = -1 + \sum_{\substack{b^2 \equiv a \pmod N \\ 0 \leq b \leq N-1}} 1$$

y γ_{χ} cumple $L(1/2 + \gamma_{\chi}i, \chi) = 0$.

Observemos que los ceros no triviales de las funciones $L(s, \chi)$ en la recta crítica vuelven a jugar un papel determinante en el término de error. Notemos también que, a más soluciones de la congruencia $x^2 \equiv a \pmod N$, más resistencia a la velocidad para el crecimiento de $\pi(x; a, N)$.

La discusión anterior nos conduce a considerar el problema de las carreras en progresiones aritméticas, como en [6]. Dados los dorsales a_1, \dots, a_r en $(\mathbb{Z}/N\mathbb{Z})^*$, queremos estudiar la evolución de carrera entre $\pi(x; a_1, N), \dots, \pi(x; a_r, N)$. Sea

$$\mathcal{P}_{a_1, \dots, a_r}^N = \{x \in \mathbb{R}_{\geq 2} : \pi(x; a_1, N) > \dots > \pi(x; a_r, N)\}.$$

Rubinstein y Sarnak prueban, bajo las ya mencionadas hipótesis GRH y GSH, que el conjunto $\mathcal{P}_{a_1, \dots, a_r}^N$ tiene densidad logarítmica positiva. Es decir, existe el límite

$$\delta(\mathcal{P}_{a_1, \dots, a_r}^N) = \lim_{x \rightarrow \infty} \frac{1}{\log x} \sum_{\substack{n \leq x \\ n \in \mathcal{P}_{a_1, \dots, a_r}^N}} \frac{1}{n}$$

y es no nulo [16]. Por ejemplo, $\delta(\mathcal{P}_{3,1}^4) = 0.9959\dots$, de modo que normalmente la cantidad de primos $p \equiv 3 \pmod{4}$ aventaja a la cantidad de primos $p \equiv 1 \pmod{4}$, aunque en infinitas ocasiones estos últimos avanzarán y tomarán la delantera.

La tabla siguiente muestra algunas de las primeras imágenes de la carrera de primos en las progresiones aritméticas módulo $N = 5$. Cada entrada se corresponde con el valor $\pi(x; a, 5)$:

x	$a = 1$	$a = 2$	$a = 3$	$a = 4$
10^2	5	7	7	5
10^3	40	47	42	38
10^4	306	309	310	303
10^5	2387	2412	2402	2390
10^6	19617	19622	19665	19593
10^7	166104	166212	166230	166032
10^8	1440298	1440496	1440474	1440186
10^9	12711386	12712315	12712499	12711333

Los dorsales no cuadrados $a = 2$ y $a = 3$ módulo 5 encabezan la carrera y éste será el patrón estándar a lo largo del campeonato; pero sabemos que en algún momento el dorsal $a = 1$ tomará el relevo y liderará la competición. En realidad, cualquier disposición de dorsales se producirá en infinitas ocasiones a lo largo de la carrera.

3. TEORÍA DE GALOIS

Sea K/\mathbb{Q} una extensión de Galois finita. Recordemos que una tal extensión viene determinada por el cuerpo $K = \mathbb{Q}(\alpha_1, \dots, \alpha_n)$, donde α_i son las raíces de un polinomio irreducible $f(X) \in \mathbb{Q}[X]$ de grado n . El grupo de Galois $\text{Gal}(K/\mathbb{Q})$ es el subgrupo del grupo simétrico \mathcal{S}_n formado por todas aquellas permutaciones que respetan las relaciones algebraicas entre las raíces $\alpha_1, \dots, \alpha_n$.

Entonces, para cada número primo p (para ser precisos, hay que excluir los primos ramificados en K/\mathbb{Q} , que son un número finito) se puede considerar la clase de conjugación

$$\text{Frob}_p \subseteq \text{Gal}(K/\mathbb{Q}),$$

que consiste en todos los elementos del grupo de Galois $\text{Gal}(K/\mathbb{Q})$ tales que su comportamiento residual es el automorfismo de Frobenius $x \mapsto x^p$.

El grupo de Galois $\text{Gal}(K/\mathbb{Q})$ es finito (de orden que divide a $n!$) y, por lo tanto, tiene un número finito de clases de conjugación. Tiene sentido preguntarse cómo se distribuyen los Frob_p en las distintas clases de conjugación de $\text{Gal}(K/\mathbb{Q})$ al variar el primo p .

Pongamos un ejemplo. Consideramos la extensión de Galois K/\mathbb{Q} obtenida al adjuntar las tres raíces del polinomio $f(X) = X^3 - 2$. Es decir, $K = \mathbb{Q}(\alpha_1, \alpha_2, \alpha_3)$ con $\alpha_1 = \sqrt[3]{2}$, $\alpha_2 = \alpha_1\omega$, $\alpha_3 = \alpha_1\omega^2$, siendo $\omega = \exp(2\pi i/3)$. En este caso, el grupo

de Galois $\text{Gal}(K/\mathbb{Q})$ es isomorfo a \mathcal{S}_3 , de orden 6 y con tres clases de conjugación:

$$\begin{aligned} [(1)] &= \{\text{id}\}, \\ [(12)] &= \{(12), (13), (23)\}, \\ [(123)] &= \{(123), (132)\}. \end{aligned}$$

Para saber a qué clase de conjugación corresponde cada primo p , basta con factorizar el polinomio $f(X)$ módulo p . Las primeras 30 clases de Frobenius se distribuyen según

Frob _p	p
[(1)]	31, 43, 109
[(12)]	5, 11, 17, 23, 29, 41, 47, 53, 59, 71, 83, 89, 101, 107, 113
[(123)]	7, 13, 19, 37, 61, 67, 73, 79, 97, 103

La mayoría de primos pertenecen a la clase [(12)] y parecen ser del orden de la mitad; los de la clase [(123)] son una tercera parte; y los de la clase trivial [(1)] quedan en franca minoría. Si consideramos los 1000 primeros números primos, los resultados que se obtienen son: 508 en la clase [(12)], que tiene tres elementos; 334 en la clase [(123)], con dos elementos; y 156 en la clase [(1)], de un solo elemento. Las frecuencias relativas parecen indicar una distribución con pesos 3/6, 2/6, 1/6, reflejando el tamaño de cada clase de conjugación.

En el caso general de una extensión de Galois finita K/\mathbb{Q} cualquiera, consideramos para cada clase de conjugación \mathcal{C} de $\text{Gal}(K/\mathbb{Q})$ la función

$$\pi(x; \mathcal{C}, K/\mathbb{Q}) = \#\{p \leq x : p \text{ primo y } \text{Frob}_p = \mathcal{C}\}.$$

El teorema de Chebotarev (1922, ver [20]) afirma que, como sugería nuestro ejemplo,

$$\lim_{x \rightarrow \infty} \frac{\pi(x; \mathcal{C}, K/\mathbb{Q})}{\pi(x)} = \frac{|\mathcal{C}|}{|\text{Gal}(K/\mathbb{Q})|}.$$

En este caso, el término de error en

$$\pi(x; \mathcal{C}, K/\mathbb{Q}) = \frac{|\mathcal{C}|}{|\text{Gal}(K/\mathbb{Q})|} \pi(x) - \text{Error}(x)$$

depende, entre otros, de los ceros de las funciones $L(s, \rho)$ en la recta crítica $\text{Re}(s) = 1/2$ para cada una de las representaciones lineales irreducibles ρ del grupo $\text{Gal}(K/\mathbb{Q})$.

Lagarias-Montgomery-Odlyzko [8], Serre [17] y Oesterlé [12] demuestran una versión efectiva del teorema de Chebotarev: asumiendo la hipótesis de Riemann generalizada, se cumple

$$|\text{Error}(x)| \leq 2 \frac{|\mathcal{C}|}{|\text{Gal}(K/\mathbb{Q})|} \sqrt{x} \log(\Delta_{K/\mathbb{Q}} x^{[K:\mathbb{Q}]}),$$

donde $\Delta_{K/\mathbb{Q}}$ denota el discriminante de la extensión K/\mathbb{Q} .

En este contexto también podemos plantear el problema de las carreras (o sesgo de Chebyshev): dadas unas clases de conjugación $\mathcal{C}_1, \dots, \mathcal{C}_r$ en $\text{Gal}(K/\mathbb{Q})$, existe la densidad logarítmica para la competición

$$\pi(x; \mathcal{C}_1, K/\mathbb{Q}) > \dots > \pi(x; \mathcal{C}_r, K/\mathbb{Q})$$

y es no nula. También en este caso intervienen los cuadrados en \mathcal{C} y el posible cero de $L(s, \rho)$ en $s = 1/2$ (ejemplos de Armitage y Serre, ver [11]).

4. CURVAS ELÍPTICAS

Consideremos la curva elíptica dada por un modelo afín $E/\mathbb{Q}: y^2 = x^3 + ax + b$ con $a, b \in \mathbb{Z}$. Su discriminante $\Delta = -16(4a^3 + 27b^2)$ es no nulo y, para cada primo p con $(p, \Delta) = 1$, podemos considerar la curva elíptica reducida módulo p :

$$\tilde{E}/\mathbb{F}_p: y^2 \equiv x^3 + ax + b \pmod{p}.$$

El conjunto $\tilde{E}(\mathbb{F}_p)$ de los puntos de \tilde{E} con coordenadas en \mathbb{F}_p (incluido el punto proyectivo del infinito) forma un grupo, que es de gran interés en asuntos cripto-gráficos, entre otros. Queremos calcular $\mathcal{N}(p) = \#\tilde{E}(\mathbb{F}_p)$. Para las curvas elípticas con multiplicación compleja, el número $\mathcal{N}(p)$ puede ser calculado cómodamente a partir de caracteres de Hecke, pero en el caso general debemos conformarnos con la estimación del número de puntos \mathbb{F}_p -racionales mediante

$$\mathcal{N}(p) = \#\mathbb{P}^1(\mathbb{F}_p) - \text{Error}(p) = (p + 1) - a_p.$$

Es decir, apostamos por $p + 1$ como término principal y denotamos por a_p el error cometido. Artin (1924) conjetura que el error cumple

$$|a_p| \leq 2\sqrt{p},$$

y Hasse (1933) prueba la conjetura [19]. El interés sobre el error a_p fue creciendo con el paso de los años, configurándose la conjetura de Shimura-Taniyama-Weil (hoy teorema de Wiles et al.) que organiza todos los errores a_p como coeficientes de Fourier de una forma modular en un espacio vectorial de dimensión finita. Entre otras virtudes, ello permite poner a nuestra disposición la prolongación analítica de la función L de Hasse-Weil,

$$L(E, s) = \prod_p \frac{1}{1 - a_p p^{-s} + p^{1-2s}} = \sum_{n=1}^{\infty} \frac{a_n}{n^s},$$

que, a priori, sólo está definida en el semiplano $\text{Re}(s) > 3/2$. Quedan aún muchas cuestiones abiertas sobre la relación de esta función L con las propiedades aritméticas de la curva elíptica (tales como la conjetura de Birch y Swinnerton-Dyer [23]), pero volvamos a nuestra discusión sobre errores y carreras.

Gracias a la cota del error de Hasse, podemos escribir

$$\frac{a_p}{2\sqrt{p}} = \cos \theta_p, \quad 0 < \theta_p < \pi.$$

Sato y Tate (1963), de forma independiente, conjeturan que, dado un intervalo $[\alpha, \beta] \subseteq [0, \pi]$, si la curva elíptica E no tiene multiplicación compleja, entonces los primos se distribuyen según la ley

$$\lim_{x \rightarrow \infty} \frac{\#\{p \leq x: \theta_p \in [\alpha, \beta]\}}{\#\{p \leq x\}} = \frac{2}{\pi} \int_{\alpha}^{\beta} \sin^2 \theta \, d\theta = \frac{\beta - \alpha}{\pi} - \frac{1}{2\pi} (\sin 2\beta - \sin 2\alpha).$$

Mientras que Sato partió de la experimentación numérica, gracias a las facilidades de las primeras computadoras de gran capacidad de cálculo, Tate, por su parte, elaboró la conjetura desde un posicionamiento teórico, llegando ambos a la misma conclusión. Como veremos más adelante, las ideas de Tate son las que han permitido a Serre (2011) generalizar la conjetura de Sato-Tate, encuadrándola en el marco más general posible [18]. Volveremos a la conjetura de Sato-Tate en la sección 6, pero antes trataremos otro de los ejemplos anunciados en la introducción.

5. FORMAS MODULARES

Fijemos k un entero positivo con $k \equiv 0 \pmod{4}$. Para cada número primo p , denotemos

$$\mathcal{N}(p) = \#\{(x_1, \dots, x_k) \in \mathbb{Z}^k : p = x_1^2 + \dots + x_k^2\}.$$

De nuevo nos planteamos la ecuación $\mathcal{N}(p) = \text{Principal}(p) - \text{Error}(p)$. La forma modular θ de nivel 4 y peso fraccionario $1/2$,

$$\theta(q) = \sum_{n \in \mathbb{Z}} q^{n^2} \in M_{1/2}(\Gamma_0(4)), \quad q = \exp(2\pi iz),$$

nos brinda el camino a la solución.

Se procede del modo siguiente. Elevamos $\theta(q)$ a la potencia k -ésima, y así obtenemos una forma modular de nivel 4 y peso $k/2$:

$$\theta^k(q) = \sum_{n \geq 0} \mathcal{N}(n) q^n \in M_{k/2}(\Gamma_0(4)).$$

Puesto que el espacio de formas modulares se descompone como suma directa del subespacio generado por las series de Eisenstein y el subespacio de las formas modulares parabólicas

$$M_{k/2}(\Gamma_0(4)) = \mathcal{E}_{k/2}(\Gamma_0(4)) \oplus S_{k/2}(\Gamma_0(4)),$$

podemos escribir de manera única

$$\theta^k(q) = E(q) + F(q),$$

con $E(q) = \sum_{n=0}^{\infty} A_n q^n \in \mathcal{E}_{k/2}(\Gamma_0(4))$ y $F(q) = \sum_{n=1}^{\infty} a_n q^n \in S_{k/2}(\Gamma_0(4))$. Entonces, el término principal y el error vienen dados por los coeficientes de Fourier respectivos de ambas formas modulares

$$N(p) = A_p + a_p.$$

En el ejemplo clásico de Ramanujan [10], donde $k = 24$, se cumple

$$\begin{aligned} \text{Principal}(p) &= \frac{16}{691}(p^{11} + 1), \\ |\text{Error}(p)| &\leq \frac{66304}{691} \sqrt{p^{11}}. \end{aligned}$$

El p -ésimo coeficiente de Fourier de la forma modular $\theta^k(q)$ proporciona directamente el valor de $N(p)$, pero si se pretende una estimación rápida resulta muy conveniente la expresión anterior; más si tenemos en cuenta que podemos acotar los coeficientes de las formas parabólicas gracias a una conjetura de Ramanujan. Concretamente, a partir del estudio de la función τ , que interviene en la forma modular discriminante

$$\Delta(q) = q \prod_{n=1}^{\infty} (1 - q^n)^{24} = \sum_{n=1}^{\infty} \tau(n) q^n,$$

Ramanujan (1916) conjetura que $|\tau(p)| \leq 2p^{11/2}$. Más generalmente, si $f(q) = \sum_{n=1}^{\infty} a_n q^n$ es una forma modular nueva normalizada ($a_1 = 1$) vector propio del álgebra de Hecke del espacio vectorial de dimensión finita $S_{2k}(\Gamma_1(N))$, se cumple $|a_p| \leq 2p^{k/2}$. Este resultado fue probado por Deligne (1974) como consecuencia de su demostración de las conjeturas de Weil.

6. LA CONJETURA DE SATO-TATE

Esta sección podría haberse titulado «Doble ST», una ST por Sato-Tate y la otra por Serre-Taylor (ver figura 1). Empecemos por la primera pareja.

Durante la primavera de 1963, Sato, junto con algunos de sus estudiantes y colaboradores, Nagashima, Namba, Yamamoto y Kuga, se aprovecha del aumento de la capacidad de cálculo para experimentar con la distribución de los coeficientes a_p/\sqrt{p} en el intervalo $[-2, 2]$ o, mejor, de los correspondientes ángulos θ_p en $[0, 2\pi]$. En las figuras 2 y 3 pueden verse parte de las cartas que se cruzaron en abril de 1963 desde Tokyo y Osaka. La figura 4 nos muestra una Hitachi HIPAC103, una de las computadoras con mayor potencia de cálculo en su época, que fueron creadas para la compañía de suministros eléctricos Kansai Electric Power y que luego pasaron a formar parte del parque informático de la Universidad de Tokyo (ver [15]).

Por su parte, Tate [21] sustenta la conjetura en el siguiente argumento teórico basado en la visión cohomológica de Weil. Fijando un número primo ℓ , el grupo de Galois absoluto actúa linealmente sobre el módulo de Tate de la curva elíptica

$$\rho_\ell: \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \rightarrow \text{Aut}(V_\ell(E)) \simeq \text{GL}(2, \mathbb{Q}_\ell),$$

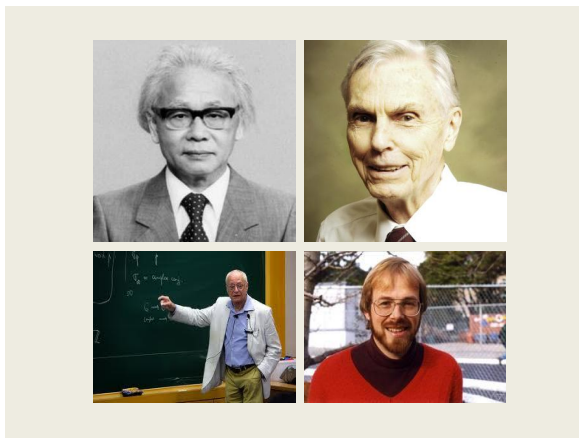


Figura 1: La doble ST: Sato-Tate y Serre-Taylor.

y para cada número primo $p \nmid \ell N$ tenemos

$$\det(\text{Id} - x\rho_\ell(\text{Frob}_p)) = x^2 - a_p x + p = (x - \sqrt{p} e^{i\theta_p})(x - \sqrt{p} e^{-i\theta_p}).$$

La asignación

$$p \rightsquigarrow x^2 - \frac{a_p}{\sqrt{p}}x + 1$$

asocia a cada primo el polinomio característico de una matriz en la clase de conjugación en el grupo especial unitario

$$G = \text{SU}(2) = \{A \in \mathcal{M}_2(\mathbb{C}) : \bar{A} = (A^{-1})^t \text{ y } \det A = 1\}.$$

En efecto, sus clases de conjugación vienen dadas por

$$X = \text{clases de conjugación de } G = \left\{ G^{-1} \begin{pmatrix} e^{i\theta} & \\ & e^{-i\theta} \end{pmatrix} G \right\}.$$

Serre traza en [18] una estrategia general. Sean X un espacio topológico compacto, μ_* una medida de Radon (es decir, una forma lineal del espacio de las funciones reales continuas $C(X)$). Se dice que una sucesión $\{x_n\}_{n \geq 1}$ de X está μ_* -equidistribuida si para cada función $f \in C(X)$ se tiene

$$\mu_*(f) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n f(x_i).$$

En las aplicaciones, se parte de un grupo compacto G y se considera el espacio $X = G/\sim$ formado por sus clases de conjugación. Entonces μ_* es la imagen directa de la medida de Haar normalizada de G . En tal caso, se cumple:



Figura 4: Hitachi HIPAC103 (1962).

PROPOSICIÓN 6.1. *Una sucesión $\{x_n\}_{n \geq 1}$ de X está μ_* -equidistribuida si y sólo si para cada carácter irreducible no trivial χ de G se cumple*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \chi(x_i) = 0.$$

En nuestro caso consideraremos sucesiones indexadas $\{x_p\}_{p \in \Sigma}$, donde Σ es un cierto conjunto infinito de números primos y, para cada representación lineal irreducible ρ de G , consideramos la función L asociada

$$L(s, \rho) = \prod_{p \in \Sigma} \det(1 - \rho(x_p)p^{-s})^{-1}, \quad \operatorname{Re}(s) > 1.$$

TEOREMA 6.2 (Serre). *Si para cada representación irreducible ρ no-trivial, la función $L(s, \rho)$ admite prolongación holomorfa y no se anula en $\operatorname{Re}(s) \geq 1$, entonces la sucesión $\{x_p\}_{p \in \Sigma}$ está μ_* -equidistribuida.*

La demostración consiste en seguir los pasos de Hadamard y de la Vallée-Poussin en la obtención del teorema de los números primos. Si χ es el carácter de ρ , debemos comprobar que

$$\lim_{N \rightarrow \infty} \frac{1}{\#\{p \leq N\}} \sum_{p \leq N} \chi(x_p) = 0 = \mu_*(\chi).$$

Podemos escribir

$$L = L(s, \rho) = \prod_p \prod_{i=1}^n \frac{1}{1 - p^{-s} \lambda_{i,p}}$$

y calcular su derivada logarítmica

$$\begin{aligned}
 -\frac{L'}{L} &= \sum_p \sum_i \frac{p^{-s} \log p}{1 - p^{-s} \lambda_{i,p}} \\
 &= \sum_p \sum_i \sum_{m \geq 1} p^{-ms} \lambda_{i,p}^m \log p \\
 &= \sum_p \sum_{m \geq 1} p^{-ms} \chi(x_p^m) \log p \\
 &= \sum_p p^{-s} \chi(x_p) \log p + \text{función holomorfa en } \operatorname{Re}(s) > 1/2.
 \end{aligned}$$

El teorema de Wiener-Ikehara implica que

$$\sum_{p \leq N} \chi(x_p) \log p = o(N).$$

Podemos entonces aplicar el criterio de sumación de Abel, y llegamos a

$$\sum_{p \leq N} \chi(x_p) = o\left(\frac{N}{\log N}\right).$$

Puesto que $\#\{p \leq N\} \sim \frac{N}{\log N}$, se sigue la conclusión del teorema.

Richard Taylor, conjuntamente con un equipo de colaboradores, consiguen seguir el camino trazado por Serre para probar la conjetura de Sato-Tate en el caso de curvas elípticas sin multiplicación compleja [7]. Como hemos indicado anteriormente, debe considerarse el grupo especial unitario $SU(2)$, cuyas representaciones irreducibles no triviales son la estándar 2-dimensional y sus potencias simétricas. Entonces, debemos considerar las funciones L siguientes, para $m \geq 1$:

$$L_m(s) := L(s, \operatorname{Sym}^m E) = \prod_p \prod_{i=0}^m (1 - \beta_p^{-m+2i} p^{-s})^{-1},$$

donde $\beta_p = e^{i\theta_p}$ (recordamos la relación $a_p/\sqrt{p} = 2 \cos \theta_p$). Finalmente, Taylor et al. logran acercarse al programa de Langlands:

CONJETURA 6.3 (Langlands). *Para cada $m \geq 1$, la representación $\rho_m = \operatorname{Sym}^m E$ es isomorfa a una representación cuspidal automorfa π_m de $GL_{m+1}(\mathbb{A})$.*

De ser cierta esta conjetura de Langlands, se sabe que $L(s, \pi_m) = L(s, \rho_m)$ se comporta bien en $s = 1$, de modo que se puede aplicar el teorema 6.2 (de Serre), de donde se sigue la conjetura de Sato-Tate para la curva elíptica E sin multiplicación compleja. Pues bien, en realidad el caso $m = 1$ no es más que el célebre teorema de Wiles y Taylor-Wiles sobre la modularidad de las curvas elípticas (del que se deduce el último teorema de Fermat). Más concretamente, se tiene $L(s, \rho_1) = L(s + \frac{1}{2}, f)$ donde f es la forma modular asociada a la curva elíptica E .

Para $m \geq 2$, Taylor et al. muestran que $L_m(s)$ tiene automorfía potencial. Ello significa automorfía tras un cambio de base, resultado parcial camino de la conjetura de Langlands, pero la hazaña resulta ser suficiente para poder aplicar igualmente el teorema de Serre y, por lo tanto, deducen la ley de Sato-Tate.

7. VARIEDADES ABELIANAS

Aunque la generalización de la conjetura de Sato-Tate propuesta por Serre tiene un radio de alcance mucho mayor, vamos a esbozarla aquí solamente para el caso de las variedades abelianas.

Sea A una variedad abeliana de dimensión g definida sobre un cuerpo de números K . Para cada número primo ℓ , consideramos la representación ℓ -ádica

$$\rho_{A,\ell}: \text{Gal}(\overline{\mathbb{Q}}/K) \rightarrow \text{Aut}(T_\ell(A)) \simeq \text{GL}(2g, \mathbb{Q}_\ell),$$

que describe la acción del grupo de Galois absoluto sobre el módulo de Tate $T_\ell(A) = \varprojlim A[\ell^n]$. Exceptuando un número finito de ideales primos \mathfrak{P} de K donde la representación ramifica, queremos estudiar la distribución de los polinomios característicos de las matrices

$$\frac{1}{N(\mathfrak{P})^{1/2}} \rho_{A,\ell}(\text{Frob}_{\mathfrak{P}}).$$

Debido a las conjeturas de Weil (teorema de Deligne), el factor local normalizado

$$L_{\mathfrak{P}}(T, A) = \det \left(1 - \frac{\rho_{A,\ell}(\text{Frob}_{\mathfrak{P}})}{N(\mathfrak{P})^{1/2}} T \right) = \prod_{j=1}^{2g} (T - e^{i\theta_j})$$

es el polinomio característico de una matriz en el grupo unitario simpléctico $\text{USp}(2g)$.

Se trata de acomodar todos estos polinomios $L_{\mathfrak{P}}(T, A)$ como polinomios característicos de las clases de conjugación en un subgrupo adecuado dentro del grupo $\text{USp}(2g)$, conforme al espacio que ocupe la imagen de la representación $\rho_{A,\ell}$, con el fin de que la medida de Haar del subgrupo en cuestión explique la equidistribución de los polinomios. Este grupo se denomina grupo de Sato-Tate asociado a la variedad abeliana A y se denota por $\text{ST}(A)$.

Más concretamente, el grupo de Sato-Tate $\text{ST}(A)$ se define como un subgrupo compacto maximal en

$$\text{ST}(A) \subseteq \rho_{A,\ell}(\ker \chi_\ell)^{\text{Zar}} \otimes_{\iota} \mathbb{C},$$

donde Zar denota la clausura de Zariski, χ_ℓ el ℓ -ésimo carácter ciclotómico y $\iota: \overline{\mathbb{Q}}_\ell \hookrightarrow \mathbb{C}$ es una inmersión escogida. Un tal subgrupo compacto maximal no es único, pero todos ellos son conjugados entre sí. Se conjetura la existencia de un subgrupo algebraico de $\text{GL}_{2g}/\mathbb{Q}$ cuya realización proporciona $\text{ST}(A)$ independientemente de ℓ y $\iota: \overline{\mathbb{Q}}_\ell \hookrightarrow \mathbb{C}$.

En cualquier caso, $\text{ST}(A)$ es un subgrupo de $\text{GL}_{2g}(\mathbb{C})$, que en cierto modo imita la definición del grupo de Galois asociado a un polinomio, así como la conjetura de Sato-Tate generaliza el teorema de densidad de Chebotarev tratado en la sección 3. Tenemos ahora la sucesión de matrices

$$x_{\mathfrak{P}} := \frac{1}{N(\mathfrak{P})^{1/2}} \rho_{A,\ell}(\text{Frob}_{\mathfrak{P}}) \in \text{ST}(A).$$

Si para cada representación irreducible $\rho: \text{ST}(A) \rightarrow \text{GL}_m(\mathbb{C})$ no trivial se cumple que la función

$$L(s, \rho, \{x_{\mathfrak{P}}\}) = \prod_{\mathfrak{P}} \det(1 - \rho(x_{\mathfrak{P}}) N(\mathfrak{P})^{-s})^{-1}$$

es holomorfa y no nula en $\text{Re}(s) \geq 1$, entonces los polinomios característicos de las matrices $x_{\mathfrak{P}}$ estarán equidistribuidos como los de las matrices en las clases de conjugación de $\text{ST}(A)$ según su medida de Haar. Ésta es la propuesta de Serre para seguir el camino trazado por Hadamard, de la Vallée-Poussin, Dirichlet y Chebotarev en el caso de las variedades abelianas.

Dada una variedad abeliana A definida sobre un cuerpo de números, la estrategia para verificar que A satisface la conjetura de Sato-Tate generalizada siempre consiste en tres pasos:

- (1) identificar el grupo de Sato-Tate $\text{ST}(A)$ y sus representaciones lineales irreducibles;
- (2) tratar de obtener la holomorfia y no anulación de $L(s, \rho, \{x_{\mathfrak{P}}\})$ para todas las representaciones irreducibles de $\text{ST}(A)$ excepto la trivial;
- (3) calcular los momentos estadísticos de la distribución asociada a la medida de Haar de $\text{ST}(A)$.

Como ya hemos comentado, el primer paso recuerda el cálculo del grupo de Galois de un polinomio; se debe abordar específicamente en cada caso, resultando la mayor parte de las veces un arte poder lograr su identificación. El segundo paso es más complicado, si cabe, y la mayoría de las veces uno debe apelar a la filosofía de Langlands para garantizar las propiedades deseadas de las funciones L involucradas. El tercer paso no es necesario, pero a menudo resulta una comprobación gratificante ver que los momentos teóricos obtenidos a partir de la medida de Haar del grupo de Sato-Tate concuerdan con los momentos estadísticos obtenidos experimentalmente en el laboratorio (calculando las frecuencias de los polinomios característicos normalizados para primos p hasta cierta cota). A veces es posible llevar a cabo alguno de los tres pasos (o incluso todos ellos):

- Curvas elípticas (Taylor con Clozel, Harris y Shepherd-Barron [2, 7]).
- Formas modulares (Taylor con Barnet-Lamb, Geraghty y Harris [22, 1]).
- Superficies abelianas (Fité-Kedlaya-Rotger-Sutherland [4]).
- Curvas de género 3 (Fité-Sutherland [5], Lario-Somoza [9]).
- Jacobianas de cocientes de las curvas de Fermat (Fité-González-Lario [3]).

REFERENCIAS

- [1] T. BARNET-LAMB, D. GERAGHTY, M. HARRIS Y R. TAYLOR, A family of Calabi-Yau varieties and potential automorphy II, *Publ. Res. Inst. Math. Sci.* **47** (2011), núm. 1, 29–98.

- [2] L. CLOZEL, M. HARRIS Y R. TAYLOR, Automorphy for some ℓ -adic lifts of automorphic mod ℓ Galois representations, *Publ. Math. Inst. Hautes Études Sci.* **108** (2008), 1–181.
- [3] F. FITÉ, J. GONZÁLEZ Y J.-C. LARIO, Frobenius distribution for quotients of Fermat curves of prime exponent, *Canad. J. Math.* **68** (2016), núm. 2, 361–394.
- [4] F. FITÉ, K. S. KEDLAYA, V. ROTGER Y A. V. SUTHERLAND, Sato-Tate distributions and Galois endomorphism modules in genus 2, *Compositio Math.* **148** (2012), núm. 5, 1390–1442.
- [5] F. FITÉ Y A. V. SUTHERLAND, Sato-Tate groups of $y^2 = x^8 + c$ and $y^2 = x^7 - cx$, *Contemporary Mathematics* **663** (2016), 103–126.
- [6] A. GRANVILLE Y G. MARTIN, Carreras de números primos, *Gac. R. Soc. Mat. Esp.* **8** (2005), núm. 1, 197–240. Publicado también en inglés: Prime number races, *Amer. Math. Monthly* **113** (2006), 1–33.
- [7] M. HARRIS, N. SHEPHERD-BARRON Y R. TAYLOR, A family of Calabi-Yau varieties and potential automorphy, *Ann. Math.* **171** (2010), núm. 2, 779–813.
- [8] J. C. LAGARIAS, H. L. MONTGOMERY Y A. M. ODLYZKO, A bound for the least prime ideal in the Chebotarev density theorem, *Invent. Math.* **54** (1979), 271–296.
- [9] J. C. LARIO Y A. SOMOZA, The Sato-Tate conjecture for a Picard curve with complex multiplication, with an appendix by F. Fité, por aparecer en *Contemporary Math.*
- [10] B. MAZUR, Finding meaning in error terms, *Bull. Amer. Math. Soc. (N. S.)* **45** (2008), núm. 2, 185–228.
- [11] N. C. NG, *Limiting Distributions and zeros of Artin L-Functions*, Ph. D. Dissertation, University of British Columbia, Canadá, 2000.
- [12] J. OESTERLÉ, Versions effectives du théorème de Chebotarev sous l’hypothèse de Riemann généralisée, *Astérisque* **61** (1979), 165–167.
- [13] J. QUER, La funció ζ de Riemann, *Butll. Soc. Cat. Mat.* **22** (2007), núm. 2, 197–228.
- [14] G. F. B. RIEMANN, Über die Anzahl der Primzahlen unter einer gegebenen Größe, *Monatsberichte der Berliner Akademie* (1859), 671–680.
- [15] J. ROSEN Y R. SCHMIDT, The Sato-Tate conjecture, <http://www2.math.ou.edu/~rschmidt/satotate/page5.html>
- [16] M. RUBINSTEIN Y P. SARNAK, Chebyshev’s bias, *Experiment. Math.* **3** (1994), núm. 3, 173–197.
- [17] J.-P. SERRE, Quelques applications du théorème de densité de Chebotarev, *Publ. Math. IHES* **54** (1981), 323–401.
- [18] J.-P. SERRE, *Lectures on $N_X(p)$* , A K Peters/CRC Press, 2011.
- [19] J. H. SILVERMAN, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics **106**, 2.^a ed., Springer-Verlag, 2009.
- [20] P. STEVENHAGEN Y H. W. LENSTRA JR., Chebotarëv and his density theorem, *Math. Intelligencer* **18** (1996), núm. 2, 26–37.

- [21] J. TATE, Algebraic cycles and poles of zeta functions, *Arithmetical Algebraic Geometry* (Proc. Conf. Purdue Univ., 1963), Harper & Row, New York, 1965.
- [22] R. TAYLOR, Automorphy for some ℓ -adic lifts of automorphic mod ℓ Galois representations II, *Pub. Math. IHES* **108** (2008), 183–239.
- [23] A. WILES, The Birch and Swinnerton-Dyer Conjecture, <http://www.claymath.org/prizeproblems/birchsd.htm>
- [24] G. WOLTMAN ET AL., GIMPS, the Great Internet Mersenne Prime Search, <https://www.mersenne.org>

JOAN CARLES LARIO, DEPARTAMENT DE MATEMÀTIQUES, UNIVERSITAT POLITÈCNICA DE CATALUNYA, BARCELONA

Correo electrónico: joan.carles.lario@upc.edu

Página web: <https://mat-web.upc.edu/people/joan.carles.lario/>