Huecos de la criba de centros de primos gemelos y distribución de sus factores locales

Miguel Cerdá Bennassar

Octubre de 2026

Resumen

Estudiamos los residuos k para los que $6k-1$ y $6k+1$ evitan los divisores primos de un producto finito de primos mayores o iguales que 5. Calculamos el límite normalizado de la frecuencia de cada hueco fijo y, para huecos crecientes, el término de segundo orden de esa frecuencia, que da una estimación uniforme mientras el hueco es pequeño frente al inverso de la densidad de la criba. La uniformidad no puede extenderse a huecos mayores que un múltiplo $c > 1$ de ese inverso. Estudiamos también la distribución asintótica de los límites cuando varía el hueco. La ley límite se representa mediante un producto aleatorio de Euler; obtenemos sus momentos, la ausencia de átomos y el soporte completo. Todos los resultados se refieren a la criba, no a la distribución efectiva de los primos gemelos. No se reivindica novedad general frente a la teoría clásica de series singulares.

1. Introducción

Toda pareja de primos gemelos mayores que 3 es de la forma $(6k-1, 6k+1)$, y llamamos a k su centro. Este escrito estudia la versión de criba de ese problema: los residuos k módulo $Q_y = \prod_{5\le p\le y} p$ para los que ni $6k-1$ ni $6k+1$ tienen factores primos que dividan Q_y, y los huecos entre elementos consecutivos de este conjunto. Al tratarse de un objeto periódico definido por condiciones locales, sus propiedades pueden demostrarse sin hipótesis; a cambio, ninguna de ellas se traslada directamente a los primos gemelos.

Para la criba de Eratóstenes, Holt [7, 8] demostró que la frecuencia de cada hueco par, relativa a la del hueco 2, tiende al factor de la conjetura B de Hardy y Littlewood [6]. En la sección 2 se obtiene el resultado correspondiente para la presente criba: para cada h fijo, el cociente entre el número de huecos de longitud h y el de huecos de longitud 1 tiende a un producto L_h de factores locales (teorema 2.2).

La sección 3 estudia la uniformidad en h. Mediante el argumento de Gallagher [4] para la media de las series singulares se calcula el término de segundo orden, $-(h-1)L_h\delta_y$, donde δ_y es la densidad de la criba (teorema 3.2). De ahí resulta una estimación relativa uniforme mientras h sea pequeño frente a $1/\delta_y \asymp (\log y)^2$ (corolario 3.4). Un recuento del número total de huecos muestra, por otra parte, que la aproximación por L_h no puede ser uniforme hasta c/δ_y con $c > 1$ (proposición 5.3).

Las secciones 4 y 5 estudian la distribución de los valores L_h cuando h varía. Esos valores se representan mediante un producto aleatorio de Euler Z, en el marco del teorema de Erdős–Wintner [3] y de los promedios de productos de Euler de Kowalski [11]; se obtienen sus momentos, su soporte y la ausencia de átomos de su ley. La sección 6 relaciona L_h con la serie singular de las constelaciones correspondientes y sitúa los resultados respecto a otros trabajos, en particular respecto a la estadística de Poisson [5, 17]. La sección 7 recoge las cuestiones que quedan abiertas.

Ninguno de estos resultados demuestra la existencia de infinitas parejas de primos gemelos ni una ley para sus separaciones reales: se refieren exclusivamente a clases admisibles y a factores locales de criba.

2. La criba y los huecos

Sea $Q_y = \prod_{5 \le p \le y} p$, donde p recorre los primos, y sea

$$A_Q = \{k \pmod{Q} : \gcd(36k^2 - 1, Q) = 1\}.$$

Denotamos por $H_Q(h)$ el número de huecos circulares de longitud h entre elementos consecutivos de A_Q. Para un conjunto finito T de desplazamientos enteros, escribimos

$$N_Q(T) = \#\{k \pmod{Q} : k + t \in A_Q \text{ para todo } t \in T\}.$$

Por el teorema chino del resto,

$$N_Q(T) = \prod_{p \mid Q} (p - \nu_p(T)), \quad \nu_p(T) = \#\bigcup_{t \in T} \{-t - 6^{-1}, -t + 6^{-1}\} \pmod{p}.$$

En particular, $H_Q(1) = \prod_{p \mid Q}(p-4)$, pues las cuatro clases de $\{0, 1\}$ son distintas para $p \ge 5$.

Lema 2.1. *Para $h \ge 1$ y $p \ge 5$, el número de clases prohibidas para $T = \{0, h\}$ es*

$$\nu_p(\{0,h\}) = \begin{cases} 2, & p \mid h, \\ 3, & p \mid 9h^2 - 1, \\ 4, & \text{en otro caso.} \end{cases}$$

Demostración. Las clases son $\pm 6^{-1}$ y $-h \pm 6^{-1}$. Dos coinciden si $h \equiv 0$ o $h \equiv \pm 2 \cdot 6^{-1} = \pm 3^{-1}$ (mód p). Para $p \ge 5$ estas tres posibilidades son disjuntas. □

Más en general, si a y b son dos centros y $d = |b - a|$, una clase de a coincide con una de b si y solo si $b - a \equiv 0$ o $b - a \equiv \pm 3^{-1} \pmod{p}$, es decir, si y solo si $p \mid d(3d-1)(3d+1)$. Las dos clases de un mismo centro son distintas para $p \ge 5$. Esta observación se usa repetidamente en la sección 3.

Teorema 2.2. *Para cada entero fijo $h \ge 1$ existe el límite*

$$L_h = \lim_{y \to \infty} \frac{H_{Q_y}(h)}{H_{Q_y}(1)} = \prod_{\substack{p \mid h \\ p \ge 5}} \frac{p-2}{p-4} \prod_{\substack{p \mid 9h^2 - 1 \\ p \ge 5}} \frac{p-3}{p-4}.$$

Demostración. Para $Q_y > h$, la inclusión–exclusión da

$$H_{Q_y}(h) = \sum_{T \subseteq \{1, \ldots, h-1\}} (-1)^{|T|} N_{Q_y}(\{0, h\} \cup T).$$

El término con $T = \varnothing$, dividido por $H_{Q_y}(1)$, tiene el valor del producto del enunciado para y suficientemente grande, por el lema 2.1. Para un T no vacío, el conjunto de desplazamientos contiene al menos tres enteros distintos. Fuera de un conjunto finito de primos excepcionales, sus seis clases prohibidas son distintas. Si algún factor local es cero, el término se anula; en caso contrario su cociente con $H_{Q_y}(1)$ está acotado por una constante dependiente de T, h multiplicada por

$$\prod_{p_0 < p \le y} \frac{p-6}{p-4} = \prod_{p_0 < p \le y} \left(1 - \frac{2}{p-4}\right) \longrightarrow 0,$$

por la divergencia de $\sum_p 1/p$. Hay un número finito de subconjuntos T, de modo que desaparecen todos los términos interiores. □

El argumento anterior se refiere a cada h fijo; la sección siguiente da estimaciones uniformes.

3. Uniformidad y término de segundo orden

Abreviemos $N_y(T) = N_{Q_y}(T)$ y sea

$$\delta_y = \prod_{5\le p\le y}\left(1-\frac{2}{p}\right) = \frac{|A_{Q_y}|}{Q_y}$$

la densidad de la criba. Puesto que $1-2/p = \big(1-1/(p-1)^2\big)(1-1/p)^2$, el teorema de Mertens da

$$\delta_y \sim \frac{12\,C_2\,e^{-2\gamma}}{(\log y)^2}, \qquad C_2 = \prod_{p\ge 3}\left(1-\frac{1}{(p-1)^2}\right),$$

y en particular $\delta_y \asymp (\log y)^{-2}$ para $y \ge 5$. Definimos

$$\mathcal{S}_y(h) = \sum_{t=1}^{h-1}\frac{N_y(\{0,t,h\})}{H_{Q_y}(1)}, \qquad \mathcal{T}_y(h) = \sum_{1\le t<t'\le h-1}\frac{N_y(\{0,t,t',h\})}{H_{Q_y}(1)}.$$

Para $h < Q_y$, un elemento k de A_{Q_y} con $k+h \in A_{Q_y}$ inicia un hueco de longitud h si y solo si $k+t \notin A_{Q_y}$ para $0<t<h$. Las desigualdades de Bonferroni dan

$$\frac{N_y(\{0,h\})}{H_{Q_y}(1)} - \mathcal{S}_y(h) \le \frac{H_{Q_y}(h)}{H_{Q_y}(1)} \le \frac{N_y(\{0,h\})}{H_{Q_y}(1)} - \mathcal{S}_y(h) + \mathcal{T}_y(h). \tag{1}$$

Si $3h+1 \le y$, todos los primos que dividen $h(9h^2-1) = h(3h-1)(3h+1)$ son menores o iguales que y, y el lema 2.1 da $N_y(\{0,h\})/H_{Q_y}(1) = L_h$.

Lema 3.1. *Para $p \ge 5$ y $h \ge 1$,*

$$\sum_{t \bmod p}\big(p-\nu_p(\{0,t,h\})\big) = (p-2)\big(p-\nu_p(\{0,h\})\big).$$

Demostración. El primer miembro cuenta los pares (k,t) módulo p tales que k, $k+t$ y $k+h$ son clases admisibles. Hay $p-\nu_p(\{0,h\})$ valores de k con k y $k+h$ admisibles, y para cada uno de ellos hay exactamente $p-2$ valores de t con $k+t$ admisible. □

Este es el argumento de Gallagher [4] para la media de las series singulares.

Teorema 3.2 (Término de segundo orden)**.** *Existe una constante absoluta c_0 (sirve $c_0 = 30$) tal que, para $h \ge 2$ e $y \ge 3h+1$,*

$$\mathcal{S}_y(h) = (h-1)\,L_h\,\delta_y\left(1+O\left(\frac{(\log 3h)^{c_0}}{h}\right)\right),$$

con una constante absoluta en el símbolo O.

Demostración. Para $0<t<h$ y $p \ge 5$ escribimos

$$\rho_p(t) = \frac{p-\nu_p(\{0,t,h\})}{p-\nu_p(\{0,h\})},$$

que solo depende de t mód p. Como $3h+1 \le y$, se tiene

$$\frac{N_y(\{0,t,h\})}{H_{Q_y}(1)} = L_h\prod_{5\le p\le y}\rho_p(t).$$

Si $p > 3h+1$, no hay coincidencias entre las clases de los tres centros, porque p no divide ningún entero $d(3d \pm 1)$ con $1 \le d \le h$, y $\nu_p(\{0,h\}) = 4$; por tanto $\rho_p(t) = (p-6)/(p-4)$. Resulta

$$\mathcal{S}_y(h) = L_h \prod_{3h+1<p\le y} \left(1 - \frac{2}{p-4}\right) M(h), \qquad M(h) = \sum_{t=1}^{h-1} \prod_{5\le p\le 3h+1} \rho_p(t).$$

Por el lema 3.1, la media de ρ_p sobre las clases módulo p es exactamente $\bar a_p = 1 - 2/p$. Escribimos $\rho_p = \bar a_p(1+\eta_p)$, donde η_p tiene media nula módulo p, y

$$M(h) = \prod_{5\le p\le 3h+1} \bar a_p \sum_{d\mid P} \sum_{t=1}^{h-1} \prod_{p\mid d} \eta_p(t), \qquad P = \prod_{5\le p\le 3h+1} p.$$

El término $d=1$ vale $(h-1)\prod_{5\le p\le 3h+1}\bar a_p$. Para $d>1$, la función $\prod_{p\mid d}\eta_p$ es periódica módulo d con media nula, de modo que su suma sobre cualquier intervalo está acotada, por el teorema chino del resto, por

$$\sum_{t \text{ mód } d} \Big|\prod_{p\mid d} \eta_p(t)\Big| = \prod_{p\mid d} p\,\mathbb{E}_p|\eta_p|,$$

donde $\mathbb{E}_p$ denota la media sobre las clases módulo p.

Acotemos $p\,\mathbb{E}_p|\eta_p|$. Sea $\nu_0 = \nu_p(\{0,h\})$ y $g_p = (p-\nu_0-2)/(p-\nu_0)$. Como $\nu_0 \le \nu_p(\{0,t,h\}) \le \nu_0+2$, todos los valores de ρ_p están en $[g_p, 1]$, intervalo de longitud $2/(p-\nu_0) \le 2/(p-4)$; también $\bar a_p$ está en ese intervalo. Por la observación que sigue al lema 2.1, $\rho_p(t) = g_p$ salvo en las clases $t \equiv 0$, h, $\pm 3^{-1}$, $h \pm 3^{-1}$, que son a lo sumo seis. En ellas $|\rho_p - \bar a_p| \le 2/(p-4)$. En las demás,

$$|g_p - \bar a_p| = \frac{1}{p} \sum_{t \text{ excepc.}} (\rho_p(t) - g_p) \le \frac{12}{p(p-4)}.$$

Por tanto $\mathbb{E}_p|\eta_p| \le 24/(\bar a_p\, p(p-4))$ y, para $p \ge 11$, donde $\bar a_p \ge 9/11$,

$$p\,\mathbb{E}_p|\eta_p| \le \frac{88}{3(p-4)}.$$

Para $p = 5, 7$ basta observar que $0 \le \rho_p \le 1$, luego $p\,\mathbb{E}_p|\eta_p|$ está acotado por una constante absoluta. En consecuencia,

$$\Big|M(h) - (h-1)\prod_{5\le p\le 3h+1}\bar a_p\Big| \le \prod_{5\le p\le 3h+1}\bar a_p \prod_{5\le p\le 3h+1}\big(1 + p\,\mathbb{E}_p|\eta_p|\big) \ll \prod_{5\le p\le 3h+1}\bar a_p\ (\log 3h)^{c_0},$$

por Mertens, con $c_0 = 30$.

Por último,

$$\begin{aligned}\prod_{5\le p\le 3h+1}\left(1-\frac{2}{p}\right)\prod_{3h+1<p\le y}\left(1-\frac{2}{p-4}\right) &= \delta_y \prod_{3h+1<p\le y}\left(1-\frac{8}{(p-2)(p-4)}\right)\\ &= \delta_y\big(1+O(1/h)\big),\end{aligned}$$

y reuniendo las estimaciones se obtiene el enunciado. □

Lema 3.3. *Para $y \ge 11$ y $1 \le h < Q_y$,*

$$\mathcal{T}_y(h) \ll \frac{h^2\big(\log(3h+3)\big)^6}{(\log y)^4},$$

con una constante absoluta.

Demostración. Sean $0 < t < t' < h$. Por la observación que sigue al lema 2.1, las ocho clases prohibidas por los centros $0, t, t', h$ son distintas módulo $p \geq 5$ salvo si p divide algún $d(3d-1)(3d+1)$, con d una de las seis diferencias entre centros; todo primo excepcional satisface, pues, $p \leq 3h+1$. Para los primos no excepcionales el factor local es $p-8$, y para los excepcionales es, como máximo, $p-2$. Los factores de $p = 5, 7$, normalizados por los de $H_{Q_y}(1)$, están acotados por una constante absoluta. Por Mertens,

$$\frac{N_y(\{0,t,t',h\})}{H_{Q_y}(1)} \ll \prod_{11 \leq p \leq y} \left(1 - \frac{4}{p-4}\right) \prod_{11 \leq p \leq 3h+1} \left(1 + \frac{6}{p-8}\right) \ll \frac{\big(\log(3h+3)\big)^6}{(\log y)^4},$$

uniformemente en t, t'. Hay menos de $h^2/2$ pares (t, t'). □

Corolario 3.4 (Uniformidad cuantitativa)**.** *Sea $h \geq 2$ e $y \geq 3h+1$. Entonces:*

(a) Se tiene

$$\frac{H_{Q_y}(h)}{H_{Q_y}(1)} = L_h\big(1 - (h-1)\delta_y\big) + O\big(L_h\,\delta_y(\log 3h)^{c_0}\big) + O\left(\frac{h^2\big(\log(3h+3)\big)^6}{(\log y)^4}\right).$$

En particular, si $h \to \infty$ y $h(\log h)^6/(\log y)^2 \to 0$,

$$L_h - \frac{H_{Q_y}(h)}{H_{Q_y}(1)} \sim (h-1)\,L_h\,\delta_y.$$

(b) Existe una constante absoluta C tal que

$$0 \leq 1 - \frac{H_{Q_y}(h)}{L_h\,H_{Q_y}(1)} \leq C\,\frac{h}{(\log y)^2}.$$

(c) Si $X = X(y)$ satisface $X\big(\log\log(3X+3)\big)^2 = o\big((\log y)^2\big)$, entonces

$$\sup_{1 \leq h \leq X} \left| \frac{H_{Q_y}(h)}{H_{Q_y}(1)} - L_h \right| \longrightarrow 0.$$

En particular, hay convergencia uniforme para $1 \leq h \leq (\log y)^2/\psi(y)$ siempre que $\psi(y)/(\log\log\log y)^2 \to \infty$.

Demostración. La parte (a) resulta de (1), del teorema 3.2 y del lema 3.3, usando $\delta_y \asymp (\log y)^{-2}$ y $L_h \geq 1$. Para (b), (1) da $0 \leq L_h - H_{Q_y}(h)/H_{Q_y}(1) \leq \mathcal{S}_y(h)$, y por el teorema 3.2

$$\mathcal{S}_y(h) \ll L_h\,\delta_y\big(h + (\log 3h)^{c_0}\big) \ll h\,L_h\,\delta_y,$$

donde el último paso usa que $(\log 3h)^{c_0}/h$ está acotado por una constante absoluta para $h \geq 2$, pues tiende a cero cuando $h \to \infty$. Como $\delta_y \asymp (\log y)^{-2}$, dividiendo por L_h se obtiene (b). Para (c), basta combinar (b) con la cota

$$L_h \leq \prod_{\substack{p \mid h(9h^2-1) \\ p \geq 5}} \left(1 + \frac{2}{p-4}\right) \ll \big(\log\log(3h+3)\big)^2,$$

que se obtiene por el argumento usual: el producto es máximo cuando los primos que dividen $h(9h^2-1) \leq 9h^3$ son los menores posibles. La condición $3X+1 \leq y$ se cumple para y suficientemente grande. Para $h = 1$ no hay nada que probar, pues $L_1 = 1$. □

Observación 3.5. *El parámetro de uniformidad es y, el mayor primo empleado en la criba, y no el módulo Q_y, que crece exponencialmente en y. La parte (b) es una estimación relativa respecto a L_h. El término $-(h-1)L_h\delta_y$ de la parte (a) coincide con el primer término del desarrollo de $L_h e^{-h\delta_y}$, forma que sugiere la heurística habitual de los huecos; esa forma no se demuestra aquí. Como comprobación numérica, con $y = 10^6$ (un millón) el cociente $\mathcal{S}_y(h)/\big((h-1)L_h\delta_y\big)$ vale:*

h	10	30	100	300	1000
cociente	0,778	0,857	0,947	0,975	0,990

El error relativo observado es mucho menor que la cota $(\log 3h)^{c_0}/h$ del teorema 3.2.

4. Modelo probabilístico y distribución

Para cada primo $p \geq 5$ tomamos variables independientes Z_p con distribución

$$\mathbb{P}\left(Z_p = \frac{p-2}{p-4}\right) = \frac{1}{p}, \quad \mathbb{P}\left(Z_p = \frac{p-3}{p-4}\right) = \frac{2}{p}, \quad \mathbb{P}(Z_p = 1) = 1 - \frac{3}{p}.$$

Definimos $Z = \prod_{p\geq 5} Z_p$. Como $\log Z_p \geq 0$ y $\sum_p \mathbb{E}\log Z_p < \infty$ (cada sumando es $O(p^{-2})$), la serie $\sum_p \log Z_p$ converge casi seguramente, y el producto converge a un valor finito casi seguramente.

Teorema 4.1. *Para toda función continua y acotada $F : [1,\infty) \to \mathbb{R}$,*

$$\lim_{X\to\infty} \frac{1}{X} \sum_{h\leq X} F(L_h) = \mathbb{E}F(Z).$$

Demostración. Sea $L_h^{(y)}$ el producto de los factores locales hasta y. Por periodicidad módulo Q_y, la distribución de $L_h^{(y)}$ para h uniforme en un período coincide con la de $\prod_{p\leq y} Z_p$. Además,

$$0 \leq \log L_h - \log L_h^{(y)} \leq R_y(h) := \sum_{\substack{p>y \\ p\mid h(9h^2-1)}} \frac{2}{p-4}.$$

Para $h \leq X$, cada primo $p \geq 5$ contribuye en, a lo sumo, tres clases residuales, pues las congruencias $h \equiv 0, \pm 3^{-1} \pmod p$ son distintas. Asimismo, todo primo que aparece divide h, $3h-1$ o $3h+1$, y por ello $p \leq 3X+1$. Usando $\#\{h \leq X : h \equiv a \pmod p\} \leq X/p + 1$, obtenemos

$$\frac{1}{X}\sum_{h\leq X} R_y(h) \leq 6 \sum_{y<p\leq 3X+1} \frac{1}{p(p-4)} + \frac{6}{X} \sum_{y<p\leq 3X+1} \frac{1}{p-4}.$$

La segunda suma, dividida por X, tiende a cero: por ejemplo, se acota por $O(\log X/X)$ al sumar sobre todos los enteros. De ahí

$$\limsup_{X\to\infty} \frac{1}{X}\sum_{h\leq X} R_y(h) \leq 6 \sum_{p>y} \frac{1}{p(p-4)} \longrightarrow 0.$$

Sea G una función acotada y lipschitziana en $[0,\infty)$. La estimación anterior permite sustituir $G(\log L_h)$ por $G(\log L_h^{(y)})$ con un error medio que tiende a cero cuando $y \to \infty$, uniformemente en X grande. Como $\log \prod_{p\leq y} Z_p \to \log Z$ casi seguramente, se obtiene $X^{-1}\sum_{h\leq X} G(\log L_h) \to \mathbb{E}G(\log Z)$. Las funciones acotadas lipschitzianas caracterizan la convergencia débil, de modo que $\log L_h$ converge débilmente a $\log Z$, y la transformación exponencial da el enunciado. □

Este teorema pertenece al marco del teorema de Erdős–Wintner [3] para funciones aditivas, aquí evaluadas también en los valores del polinomio $9h^2-1$, y al de los promedios de productos de Euler y series singulares estudiados por Kowalski [11].

5. Momentos y soporte

Lema 5.1. *Para cada $r > 0$, existe K_r tal que $\sup_{X\geq 1} X^{-1}\sum_{h\leq X} L_h^r \leq K_r$.*

Demostración. Sea $G(n) = \prod_{p\mid n, p\geq 5}(1+2/(p-4))$. Se tiene $L_h \leq G(h)G(3h-1)G(3h+1)$. Para $s > 0$ escribimos $G(n)^s = \sum_{d\mid n} c_s(d)$, con c_s multiplicativa, soportada en libres de cuadrados cuyos factores primos son mayores o iguales que 5, y $c_s(p) = (1+2/(p-4))^s - 1 = O_s(1/p)$. Para cada uno de los pares $(a,b) \in \{(1,0),(3,-1),(3,1)\}$, a es invertible módulo todo primo $p \geq 5$. En consecuencia,

$$\frac{1}{X}\sum_{h\leq X} G(ah+b)^s \leq \sum_{d\leq 3X+1}\frac{c_s(d)}{d} + \frac{1}{X}\sum_{d\leq 3X+1} c_s(d).$$

La primera suma es acotada uniformemente, pues $\prod_p(1+c_s(p)/p)$ converge absolutamente, ya que $c_s(p)/p = O_s(p^{-2})$. La segunda es $O_s((\log(2X))^{C_s})$ para alguna constante C_s, por comparación con $\prod_{p\leq 3X+1}(1+C_s/p)$; dividida por X tiende a cero. Hölder con exponente 3 concluye la prueba. □

Teorema 5.2. *Para todo real $r > 0$,*

$$\lim_{X\to\infty}\frac{1}{X}\sum_{h\leq X} L_h^r = \prod_{p\geq 5}\left(1 - \frac{3}{p} + \frac{1}{p}\left(\frac{p-2}{p-4}\right)^r + \frac{2}{p}\left(\frac{p-3}{p-4}\right)^r\right).$$

En particular, para $r = 1$,

$$\mathbb{E}Z = \prod_{p\geq 5}\frac{(p-2)^2}{p(p-4)}.$$

Demostración. El lema 5.1 para algún $s > r$ da integrabilidad uniforme de las potencias r; se aplica el teorema 4.1 por truncación. El producto resulta de la independencia y convergencia monótona de los productos parciales. Para $r = 1$, el factor local es $1-3/p+(3p-8)/(p(p-4)) = (p-2)^2/(p(p-4))$. □

Proposición 5.3 (Escala de la uniformidad)**.** *Para todo $c > 1$,*

$$\liminf_{y\to\infty}\sup_{1\leq h\leq c/\delta_y}\left|\frac{H_{Q_y}(h)}{H_{Q_y}(1)} - L_h\right| \geq \left(1 - \frac{1}{c}\right)\mathbb{E}Z.$$

Demostración. Sea $X = \lfloor c/\delta_y\rfloor$ y ε_y el supremo del enunciado. Para y grande, $X < Q_y$. Cada elemento de A_{Q_y} inicia exactamente un hueco circular, luego

$$\sum_{h\leq X}\frac{H_{Q_y}(h)}{H_{Q_y}(1)} \leq \frac{|A_{Q_y}|}{H_{Q_y}(1)} = \prod_{5\leq p\leq y}\frac{p-2}{p-4} = \frac{1}{\delta_y}\prod_{5\leq p\leq y}\frac{(p-2)^2}{p(p-4)} \leq \frac{\mathbb{E}Z}{\delta_y},$$

pues los factores de $\mathbb{E}Z$ son mayores que 1. Por otra parte, por el teorema 5.2 con $r = 1$,

$$\sum_{h\leq X}\frac{H_{Q_y}(h)}{H_{Q_y}(1)} \geq \sum_{h\leq X} L_h - X\varepsilon_y = X\big(\mathbb{E}Z + o(1) - \varepsilon_y\big).$$

Como $X\delta_y \to c$, se obtiene $\mathbb{E}Z \geq c\big(\mathbb{E}Z - \varepsilon_y\big) + o(1)$, es decir, $\varepsilon_y \geq (1-1/c)\mathbb{E}Z + o(1)$. □

Observación 5.4 (Consecuencia condicional de una ley de Poisson)**.** *Supóngase que, para cada $c > 0$, la proporción de huecos circulares de A_{Q_y} de longitud menor o igual que c/δ_y tiende a $1 - e^{-c}$ cuando $y \to \infty$; es decir, que los huecos normalizados por la separación media $1/\delta_y$*

siguen en el límite la ley exponencial de media uno. Como el número total de huecos es $|A_{Q_y}|$ *y*

$$\frac{|A_{Q_y}|}{H_{Q_y}(1)} = \frac{C_y}{\delta_y}, \qquad C_y = \prod_{5\le p\le y} \frac{(p-2)^2}{p(p-4)} \longrightarrow \mathbb{E}Z,$$

se tendría, con $X = \lfloor c/\delta_y \rfloor$,

$$\sum_{h\le X} \frac{H_{Q_y}(h)}{H_{Q_y}(1)} = \frac{\mathbb{E}Z}{\delta_y}\big(1 - e^{-c} + o(1)\big),$$

y la comparación con $\sum_{h\le X} L_h = X(\mathbb{E}Z + o(1))$*, como en la demostración anterior, daría para todo* $c > 0$

$$\liminf_{y\to\infty} \sup_{1\le h\le c/\delta_y} \left| \frac{H_{Q_y}(h)}{H_{Q_y}(1)} - L_h \right| \ge \mathbb{E}Z \left(1 - \frac{1-e^{-c}}{c}\right).$$

Esta observación depende de la hipótesis indicada, que no se demuestra aquí (véase la sección 6); la proposición 5.3 no la utiliza.

Junto con el corolario 3.4(c), la proposición 5.3 muestra que la escala natural de la uniformidad es $h \asymp 1/\delta_y \asymp (\log y)^2$: hay convergencia uniforme cuando $h\,\delta_y$ tiende a cero algo más deprisa que $(\log\log\log y)^{-2}$, y no la hay para h hasta c/δ_y con $c > 1$.

Teorema 5.5. *La ley de* Z *tiene soporte* $[1, \infty)$.

Demostración. Escribamos $b_p = \log((p-3)/(p-4))$. Entonces $b_p \to 0$ y $\sum_p b_p = \infty$. Dado $x \ge 0$ y $\varepsilon > 0$, existe un conjunto finito de primos grandes $\mathcal{E}$ tal que $|\sum_{p\in\mathcal{E}} b_p - x| < \varepsilon/2$ (para $x = 0$, tomamos $\mathcal{E}$ vacío). Fijando un primo R mayor que todos los de $\mathcal{E}$, suficientemente grande, la esperanza de la cola $\sum_{p>R} \log Z_p$ es menor que $\varepsilon/4$. El suceso de que $\log Z_p = b_p$ para $p \in \mathcal{E}$ y $\log Z_p = 0$ para los restantes $p \le R$ tiene probabilidad positiva. Independientemente, la cola es menor que $\varepsilon/2$ con probabilidad positiva, por Markov. Así todo entorno de x tiene probabilidad positiva bajo la ley de $\log Z$. Exponenciando se obtiene el resultado. □

Proposición 5.6 (Ausencia de átomos; aplicación de un teorema clásico)**.** *La ley de* Z *carece de átomos.*

Demostración. Las variables discretas independientes $X_p = \log Z_p$ tienen masa puntual máxima $m_p = 1 - 3/p$ para todo $p \ge 5$ (para $p = 5$ coinciden $1 - 3/5$ y $2/5$). Sea $Y_y = \sum_{5\le p\le y} X_p$. Para una variable real W y $\lambda > 0$, sea $Q(W;\lambda) = \sup_x \mathbb{P}(W \in [x, x+\lambda])$ su función de concentración. La desigualdad de Kolmogórov–Rogozin en la formulación de Rogozin [18], aplicada a las sumas finitas independientes con el mismo ancho $\lambda > 0$, da

$$Q(Y_y;\lambda) \le C \left(\sum_{5\le p\le y} (1 - Q(X_p;\lambda)) \right)^{-1/2},$$

con una constante absoluta C independiente de y y λ. Para un y fijo, podemos hacer $\lambda \downarrow 0$; como cada X_p tiene soporte finito, $Q(X_p;\lambda) \to m_p$ y $Q(Y_y;\lambda) \to \sup_a \mathbb{P}(Y_y = a)$. En consecuencia,

$$\sup_a \mathbb{P}(Y_y = a) \le C \left(\sum_{5\le p\le y} (1 - m_p) \right)^{-1/2} = C \left(\sum_{5\le p\le y} \frac{3}{p} \right)^{-1/2} \longrightarrow 0.$$

La cola $\sum_{p>y} X_p$ es independiente de Y_y; condicionar por ella demuestra que $\mathbb{P}(\log Z = a) = 0$ para todo a. La desigualdad se aplica con ancho final y anchos individuales iguales. El límite $\lambda \downarrow 0$ se toma para cada suma finita, antes de hacer $y \to \infty$. □

El resultado es también un caso del teorema de Lévy [12]: una serie convergente de variables discretas independientes tiene ley continua si y solo si $\prod_p m_p = 0$, condición que aquí equivale a $\sum_p 3/p = \infty$. Su análogo aritmético es el criterio de continuidad de Erdős y Wintner [3].

Corolario 5.7. *La función de distribución de Z es continua y estrictamente creciente en $[1,\infty)$; en particular, para $1 < u < v$, la densidad natural de $\{h : u < L_h < v\}$ existe y es positiva.*

Observación 5.8 (Naturaleza de la distribución límite)**.** *Para p grande, los valores no triviales de $\log Z_p$ satisfacen*

$$\log\frac{p-2}{p-4} = \frac{2}{p} + O(p^{-2}), \qquad \log\frac{p-3}{p-4} = \frac{1}{p} + O(p^{-2}),$$

y se presentan, en conjunto, con probabilidad $3/p$. Las perturbaciones se hacen pequeñas, pero ocurren infinitas veces casi seguramente, porque $\sum_p 3/p = \infty$; a la vez, $\sum_p \mathbb{E}[\log Z_p] < \infty$ garantiza la convergencia casi segura de su suma. Esta combinación no determina por sí sola la regularidad de la ley límite.

Observación 5.9 (Continuidad absoluta o singularidad)**.** *La función característica de $Y = \log Z$ es*

$$\varphi_Y(u) = \prod_{p\geq 5}\left(1 - \frac{3}{p} + \frac{1}{p}\left(\frac{p-2}{p-4}\right)^{iu} + \frac{2}{p}\left(\frac{p-3}{p-4}\right)^{iu}\right).$$

La ausencia de átomos no implica que la ley tenga densidad. La propiedad de tipo puro de Jessen y Wintner [10], aplicada a esta convolución infinita, deja las alternativas de continuidad absoluta y singularidad continua. Decidir entre ellas exigiría información adicional, por ejemplo estimaciones globales de la función característica. Erdős [2] estudió esta alternativa para funciones aditivas y dio condiciones que implican la singularidad; por ese método se sabe que la distribución de $\varphi(n)/n$ es puramente singular. La función $\log L_h$ toma valores del orden de $1/p$ en los primos que dividen h o $9h^2-1$, como $\log(n/\varphi(n))$ en los que dividen n, por lo que cabe esperar que el método sea aplicable; no se ha comprobado aquí.

6. Serie singular y relación con otros trabajos

Para las cuatro formas lineales $6k-1$, $6k+1$, $6(k+h)-1$, $6(k+h)+1$, la serie singular local es

$$\mathfrak{S}_h = \prod_p \frac{1-\nu_p/p}{(1-1/p)^4} = 81 B_4 L_h, \qquad B_4 = \prod_{p\geq 5}\frac{1-4/p}{(1-1/p)^4},$$

donde $\nu_2 = \nu_3 = 0$ y $\nu_p = \nu_p(\{0,h\})$ para $p \geq 5$. En $p = 2, 3$ ninguna de las cuatro formas se anula y sus factores son 16 y $81/16$. Esta igualdad es una normalización en la variable k; no debe confundirse con una parametrización diferente de las mismas constelaciones.

Observación 6.1 (Media de las series singulares)**.** *Por el teorema 5.2 con $r = 1$,*

$$\lim_{X\to\infty}\frac{1}{X}\sum_{h\leq X}\mathfrak{S}_h = 81 B_4\,\mathbb{E}Z = 81\prod_{p\geq 5}\left(1-\frac{1}{(p-1)^2}\right)^2 = (12\,C_2)^2,$$

y $12\,C_2 = 9\prod_{p\geq 5}(1-2/p)/(1-1/p)^2$ es la serie singular del par $6k-1$, $6k+1$ en la variable k. Es la forma que adopta aquí el teorema de Gallagher [4] sobre la media de las series singulares, y sirve de comprobación de las normalizaciones anteriores.

Relación con otros trabajos. Para la criba de Eratóstenes, Holt [7, 8] demuestra que cada hueco par $2n$ aparece en el ciclo de huecos módulo el primorial y que el cociente entre su número de apariciones y el del hueco 2 tiende a $\prod_{q>2,\,q\mid n}(q-1)/(q-2)$, el factor de la conjetura B de Hardy y Littlewood [6]; el teorema 2.2 es el resultado correspondiente para las formas $6k \pm 1$. En el sistema dinámico de Holt y Rudd [8], los autovalores rigen la velocidad con que esos cocientes convergen a su límite. La identidad del lema 3.1 es el argumento de Gallagher [4]. Las sumas de series singulares con sus términos secundarios se estudian en [14], y los momentos de los huecos entre residuos reducidos, uniformemente en el módulo, en [15]. Banks, Ford y Tao [1] observan que las versiones uniformes conocidas de esas sumas solo valen en un rango limitado del tamaño de los conjuntos de desplazamientos, y emplean un modelo de criba aleatoria. En la heurística de Lemke Oliver y Soundararajan [13] sobre primos consecutivos en clases residuales, el producto correspondiente se desarrolla por inclusión–exclusión y se trunca como en la criba de Brun, lo que da términos secundarios; las desigualdades (1) desempeñan aquí un papel análogo dentro de la criba. La heurística de los huecos entre primos consecutivos, con el factor exponencial correspondiente, se expone en [16]. El régimen en que $h\,\delta_y$ permanece acotado, donde cabría esperar $H_{Q_y}(h)/H_{Q_y}(1) \approx L_h e^{-h\delta_y}$, no se trata aquí.

Relación con la estadística de Poisson. Granville y Kurlberg [5] dan condiciones suficientes para que las separaciones de conjuntos de residuos construidos mediante el teorema chino del resto sigan, en el límite, una estadística de Poisson; de ellas se deduce el resultado de Hooley [9] para los residuos coprimos con el módulo. Su teorema 1 exige que, para cada k,

$$N_k(\mathbf{h}, \Omega_p) = r_p^k\, p\,\big(1 + O_k((1-r_p)p^{-\epsilon})\big)$$

siempre que $0, h_1, \ldots, h_{k-1}$ sean distintos módulo p, donde Ω_p es el conjunto local y $r_p = |\Omega_p|/p$. Para la presente criba, $r_p = 1 - 2/p$ si $p \geq 5$. Si las clases prohibidas por los k centros son todas distintas, $N_k = p - 2k$ y el cociente $N_k/(r_p^k p)$ es $1 + O_k(p^{-2})$, de modo que la condición se cumple. Pero cuando dos centros distintos difieren en $\pm 3^{-1}$ módulo p hay coincidencias entre sus clases prohibidas, y el cociente difiere de 1 en una cantidad del orden de $1/p$; por ejemplo, para $k = 2$ y $h_1 \equiv \pm 3^{-1}$, vale $(p-3)/(p-4+4/p)$. Esas mismas coincidencias producen los factores $(p-3)/(p-4)$ de L_h. La hipótesis local uniforme no se verifica, pues, literalmente, y el teorema 1 de [5] no proporciona por sí solo una ley de Poisson para los conjuntos A_{Q_y}; los autores señalan también la dificultad de sustituirla por una condición promediada. Esto no impide que tal ley pueda demostrarse por otro procedimiento. Portela [17] ha estudiado las separaciones de los residuos que evitan las clases $\pm\lfloor (p+1)/6 \rfloor$, que coinciden con $\pm 6^{-1}$, módulo $\prod_{5 \leq p \leq 6n+1} p$, con el objetivo de obtener un límite de Poisson a la escala del hueco medio. Ese problema es distinto del estudio de las frecuencias de huecos concretos desarrollado aquí. Ninguna de nuestras demostraciones utiliza una hipótesis poissoniana; la observación 5.4 es la única afirmación condicional del texto.

7. Cuestiones abiertas

La primera cuestión es el comportamiento de $H_{Q_y}(h)/H_{Q_y}(1)$ cuando $h\,\delta_y$ permanece acotado. El corolario 3.4 da convergencia uniforme hacia L_h cuando $h\,\delta_y$ tiende a cero algo más deprisa que $(\log\log\log y)^{-2}$, y la proposición 5.3 excluye la uniformidad hasta c/δ_y con $c > 1$; entre ambos regímenes, la forma esperada $L_h e^{-h\delta_y}$ no se ha demostrado. Un primer paso sería la asintótica de $\mathcal{T}_y(h)$, que cabe esperar igual a $\binom{h-1}{2} L_h \delta_y^2$ por el mismo argumento de Gallagher; el desarrollo completo exigiría controlar las configuraciones de muchos centros con uniformidad en su número, como en [14, 1].

La segunda es la ley de Poisson para las separaciones de A_{Q_y} a la escala de la separación media. Como se ha explicado en la sección 6, el teorema 1 de [5] no se aplica literalmente a esta criba. Una demostración daría la conclusión de la observación 5.4 para todo $c > 0$.

La tercera se refiere al tamaño del error en el teorema 3.2. La cota $(\log 3h)^{c_0}/h$, con $c_0 = 30$, procede de una acotación de cada factor local por separado; la tabla de la observación 3.5 sugiere un error relativo mucho menor, cuyo orden exacto no se ha determinado.

Por último, queda sin decidir si la ley de Z es absolutamente continua o singular continua (observación 5.9). Es natural examinar si el método de Erdős [2] para $\varphi(n)/n$ se aplica a este caso.

Referencias

[1] W. Banks, K. Ford y T. Tao, "Large prime gaps and probabilistic models", *Invent. Math.* **233** (2023), 1471–1518, doi:10.1007/s00222-023-01199-0.

[2] P. Erdős, "On the smoothness of the asymptotic distribution of additive arithmetical functions", *Amer. J. Math.* **61** (1939), 722–725, doi:10.2307/2371327.

[3] P. Erdős y A. Wintner, "Additive arithmetical functions and statistical independence", *Amer. J. Math.* **61** (1939), 713–721.

[4] P. X. Gallagher, "On the distribution of primes in short intervals", *Mathematika* **23** (1976), 4–9.

[5] A. Granville y P. Kurlberg, "Poisson statistics via the Chinese remainder theorem", *Adv. Math.* **218** (2008), 2013–2042, doi:10.1016/j.aim.2008.04.001.

[6] G. H. Hardy y J. E. Littlewood, "Some problems of 'Partitio numerorum'; III: On the expression of a number as a sum of primes", *Acta Math.* **44** (1923), 1–70.

[7] F. B. Holt, "Combinatorics of the gaps between primes", arXiv:1510.00743.

[8] F. B. Holt y H. Rudd, "Eratosthenes sieve and the gaps between primes", arXiv:1408.6002.

[9] C. Hooley, "On the difference between consecutive numbers prime to n. II", *Publ. Math. Debrecen* **12** (1965), 39–49; "III", *Math. Z.* **90** (1965), 355–364.

[10] B. Jessen y A. Wintner, "Distribution functions and the Riemann zeta function", *Trans. Amer. Math. Soc.* **38** (1935), 48–88, doi:10.1090/S0002-9947-1935-1501802-5.

[11] E. Kowalski, "Averages of Euler products, distribution of singular series and the ubiquity of Poisson distribution", *Acta Arith.* **148** (2011), 153–187, doi:10.4064/aa148-2-4.

[12] P. Lévy, "Sur les séries dont les termes sont des variables éventuelles indépendantes", *Studia Math.* **3** (1931), 119–155.

[13] R. J. Lemke Oliver y K. Soundararajan, "Unexpected biases in the distribution of consecutive primes", *Proc. Natl. Acad. Sci. USA* **113** (2016), E4446–E4454, doi:10.1073/pnas.1605366113.

[14] H. L. Montgomery y K. Soundararajan, "Primes in short intervals", *Comm. Math. Phys.* **252** (2004), 589–617.

[15] H. L. Montgomery y R. C. Vaughan, "On the distribution of reduced residues", *Ann. of Math. (2)* **123** (1986), 311–333.

[16] A. Odlyzko, M. Rubinstein y M. Wolf, "Jumping champions", *Experiment. Math.* **8** (1999), 107–118.

[17] F. Portela, "A Poisson limit model for twin primes via multi-residue sieve spacings", preprint, Zenodo (2026), doi:10.5281/zenodo.18758253.

[18] B. A. Rogozin, "An Estimate for Concentration Functions", *Theory Probab. Appl.* **6** (1961), 94–97, doi:10.1137/1106009.