hexagon:2610.00661v2

CROSS-ENERGY OF DIRICHLET PARTIAL SUMS. PART 11: A TWO-SIDED CRITERION FOR THE RIEMANN HYPOTHESIS FROM WEIGHTED PRIME SUMS, VERIFIED UP TO $2 \cdot 10^{29}$

PEDRO CACERES

ABSTRACT. For $N > 1$ let $C(N) = 2\sum_{n<N} \Lambda(n) n^{-1/2} \log(N/n)$ and let $M(N) = R(N) - C(N)$, where $R(N) = 8\sqrt{N} - 5.3721834 \log N - 7.4013354 + O(N^{-5/2})$ is explicit. By a theorem of Suzuki, the Riemann hypothesis (RH) holds if and only if $M(N) \geq 0$ for all large N, and then $0 \leq M(N) \leq 4(\log \xi)''(\frac{1}{2}) = 0.18484$. We prove, with computer assistance, that this two-sided inequality holds for every real $1 < N \leq 2.15 \cdot 10^{29}$: up to 10^{10} from the primes with correctly rounded arithmetic, and beyond from the first 4520 zeros, certified with ball arithmetic, together with the verification of RH to $3 \cdot 10^{12}$ and an explicit zero-density estimate. The same ingredients give, for every N and without any hypothesis, $-E(N) \leq M(N) \leq 0.18484 + E(N)$ with $E(N)$ explicit and $E(10^{20}) < 3 \cdot 10^{-5}$. We describe the local minima of M exactly, through the slope $\alpha(x) = \sum_{n \leq x} \Lambda(n) n^{-1/2} - 2\sqrt{x} + \frac{c_0}{2}$. Under RH the margin is the variogram of α, and RH is equivalent to $\int_1^X \alpha^2 \, dx/x = X^{o(1)}$. Without any hypothesis we prove an energy identity, which expresses the energy of the slope as a C-transform of the margin against the prime measure, and a lower bound for the mean energy that is attained exactly under RH with simple zeros. We then show what such arguments cannot do. Positive weights within $q^{\delta - 1/2}$ of $\Lambda(q) q^{-1/2}$ satisfy the energy identity, the prime number theorem and Chebyshev's identity up to $x^{1/2+\delta}$, yet carry an off-line zero. So the structural inputs listed do not suffice, and a proof along these lines has to use $\Lambda = \mu * \log$ exactly. Carrying that identity through the slope, α is numerically, to within 1% in the range tested, a Möbius-weighted sawtooth, and the exact argument reduces to a type I sum of μ against the fractional parts of x/d, which is RH again.

1. INTRODUCTION

For $N > 1$ put

$$C(N) = 2 \sum_{n<N} \frac{\Lambda(n)}{\sqrt{n}} \log \frac{N}{n},$$

the Cesàro (Riesz) mean of order one of $\sum \Lambda(n) n^{-1/2}$, and

$$R(N) = 8\sqrt{N} - \Big(\log \pi + \gamma_E + \frac{\pi}{2} + 3 \log 2 \Big) \log N - 16 + \frac{\pi^2}{2} + 4G - 2 \sum_{k \geq 1} \frac{N^{-(2k+1/2)}}{(2k + \frac{1}{2})^2}, \tag{1}$$

with γ_E Euler's constant and G Catalan's constant. The three pieces of R are the pole of ζ, the conductor with the Γ factor at $s = \frac{1}{2}$, and the trivial zeros. The object of this paper is the *margin* $M(N) = R(N) - C(N)$. Both R and C are of size $8\sqrt{N}$. By Suzuki's theorem [Su1, Su2] (Theorem 3), RH holds if and only if $M(N) \geq 0$ for all large N, and then $0 \leq M(N) \leq 0.18484$ for every $N > 1$. So RH is equivalent to confining a sum of size $8\sqrt{N}$ over the prime powers below N to a window of fixed width, from both sides. This is as hard as RH.

Date: Version 3, 3 November 2026 (new Section 9: the variance of the margin and the area of the quadruples). November 2026 (v2); renumbered 2 November 2026 (formerly a standalone note). Part 11 of the book-series *Cross-Energy of Dirichlet Partial Sums* (forty-seven parts, Parts 0–46; block III: The margin and the constant $\widehat{B}$). Part 0 [CE0] is the guide. A public version of this part is the Hexagon preprint hexagon:2610.00661v1.

2020 *Mathematics Subject Classification.* Primary 11M26; Secondary 11N05, 11M06, 42A75.

Key words and phrases. Riemann hypothesis, explicit formula, screw function, Riesz means, almost periodic functions, Möbius function, Beurling primes.

What the paper does. It studies this one-variable inequality from three sides and records exactly where each side stops.

(1) *The margin* (Section 2). The margin is a Fejér sum over the zeros (Proposition 1). The window width is $4(\log\xi)''(\frac{1}{2})$, computable without the zeros (Corollary 4). The two-sided inequality holds for all real $1 < N \le 10^{10}$ by a computation from the primes with correctly rounded arithmetic, up to 10^{12} by a double-precision computation with a priori error bounds, and, from the certified first 4520 zeros, up to $2.15 \cdot 10^{29}$; without any hypothesis it holds up to an explicit error $E(N)$ (Proposition 8, Theorem 6, Corollary 7).

(2) *The minima and the slope* (Section 3). M is convex between prime powers, so its local minima form an explicit countable set, and RH is equivalent to their nonnegativity (Corollary 10). Under RH they accumulate at 0 (Proposition 11). The slope of the margin is the fluctuation α of $\sum\Lambda(n)n^{-1/2}$ (Proposition 12). The minima are the downward zeros of α, and they follow from one another by a recurrence (Proposition 14).

(3) *Energy* (Section 4). Under RH, M is the variogram of α, and the window is the inequality $0 \le \langle|\Delta f|^2\rangle \le 4\langle|f|^2\rangle$ (Proposition 15). RH is equivalent to finite energy of α (Proposition 16), and to a one-sided bound for the L-average of M (Proposition 21). Without hypothesis we prove an energy identity (Theorem 18) and a lower bound for the energy (Theorem 20). Block energies to 10^{10} agree with the prediction.

(4) *Variance and area* (Section 9, version 3). The four zeros $\frac{1}{2} \pm \delta \pm i\gamma$ of an off-line quadruple form a rectangle of area $4\delta\gamma = 2|\operatorname{Im}\lambda_\rho|$. The holomorphic sum $\sum_\rho t_\rho^{-4} = -\frac{1}{6}(\log\xi)^{(4)}(\frac{1}{2})$ and the modulus sum $\sum_\rho |t_\rho|^{-4}$ differ by twice a sum of squared areas (Proposition 30). Under RH the logarithmic variance of M is $8\sum m_\gamma^2\gamma^{-4} \ge -\frac{2}{3}(\log\xi)^{(4)}(\frac{1}{2}) = 2.9738 \cdot 10^{-4}$, with equality iff the zeros are simple, without LI (Proposition 31); the primes below $2 \cdot 10^8$ and the zeros agree on the measured variance to three to four digits.

(5) *Limits* (Section 5). Theorem 23 plants an off-line zero in positive weights that satisfy every structural input of the energy arguments, so these arguments cannot prove RH. Proposition 25 carries the identity $\Lambda = \mu * \log$, which excludes the planted weights, through the slope exactly. The result is a type I sum of μ against the sawtooth (Remark 28). Section 6 summarizes the companion results on Weil's form in windows [CE13, CE14, CE21].

What is due to others. Weil's explicit formula and criterion [We52, Bo, IK]; the explicit formula for Riesz means [In]; the one-variable criterion, its bounds and its Laplace transform are Suzuki's [Su1, Su2] (we found his papers after our proofs were written); Landau's theorem on Laplace transforms [Wi, In]; the evaluation of $\sum_\rho(\rho - \frac{1}{2})^{-2}$ [Vo]; the limiting distributions of prime error terms [RuSa, ANS]; Beurling systems with large oscillation [DMV]; the distribution of the Mertens function [Ng]. What we believe is new: the verification to 10^{12}, the exact description of the minima and their recurrence, the variogram reading of the margin, the energy identity and the energy lower bound, the planted-zero theorem, and the Möbius representation of the slope.

Controls. A statement that holds for the Davenport–Heilbronn function [DH, T], which has the functional equation of ζ but no Euler product and zeros off the line, cannot be the reason for RH. We use it, a Beurling system [DMV] and planted weights (Section 5.1) as controls throughout.

Notation. A nontrivial zero on the line is written $\rho = \frac{1}{2} + i\gamma_\rho$ with γ_ρ real, and a zero off the line $\rho = \beta + i\gamma$ with $\gamma > 0$ and $\delta = \beta - \frac{1}{2}$ (its conjugate is $\bar\rho$). For every nontrivial zero we put $t_\rho := (\rho - \frac{1}{2})/i$, so that $t_\rho = \gamma_\rho$ on the line and $t_\rho = \gamma - i\delta$ off the line; m_γ is the multiplicity. We write $T_0 = 3 \cdot 10^{12}$ for the height to which RH is verified [PT]. For a test function g, $\widehat{g}(z) = \int g(x)e^{izx}dx$. The explicit formula, for G even, continuous, of bounded variation and compact support [Bo, IK],

reads

$$\sum_{\rho} \widehat{G}(t_\rho) = \widehat{G}(\tfrac{i}{2}) + \widehat{G}(-\tfrac{i}{2}) - G(0)\log\pi + \frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{G}(r)\ \mathrm{Re}\,\psi\big(\tfrac{1}{4}+\tfrac{ir}{2}\big)dr - \sum_{n\geq 2}\frac{\Lambda(n)}{\sqrt{n}}\big(G(\log n)+G(-\log n)\big), \tag{2}$$

the sum over the zeros taken symmetrically. Weil's criterion says that RH holds if and only if the right side is nonnegative for every $G = g * \tilde{g}$, where $\tilde{g}(x) = \overline{g(-x)}$. The right side is *Weil's quadratic form* $Q(g)$.

Place in the book. This is Part 11 of the series, the first chapter of block III (The margin and the constant $\widehat{B}$). It builds on Part 13 [CE13] (the Euler cross-energy and the one-variable form of RH), Part 14 [CE14] (the margin written as $W - E$), Part 21 [CE21] (the minimizer of Weil's form in a window), Part 10 [CE10] (the Caceres prime kernel) and Part 37 [CE37] (the $6k \pm 1$ margin). It is used by Part 12 [CE12] (the constant $\widehat{B}$ seen from three sides), by Parts 23 and 29 [CE23, CE29] (window minimizers; Weil windows as density operators), by Part 36 [CE36] (the Möbius side of the same margin) and by Parts 10 and 37. It is also used by Parts 24, 27 and 30 [CE24, CE27, CE30].

2. The margin

The margin is the explicit formula for one test function, the autocorrelation of the indicator of $[0, \log N]$.

Proposition 1 (the margin over the zeros)**.** *Let $N > 1$ and $L = \log N$. Without any hypothesis,*

$$M(N) = \sum_{\rho} F_L(t_\rho), \qquad F_L(z) = \frac{4\sin^2(Lz/2)}{z^2} = \Big|\int_0^L e^{izu}du\Big|^2 \ (z \in \mathbb{R}),$$

the sum over all nontrivial zeros with multiplicity, absolutely convergent. More generally, for a window $\varphi \geq 0$ on $[0, L]$ with autocorrelation $A(u) = \int_0^{L-u}\varphi(v)\varphi(v+u)dv$ and $t \in \mathbb{R}$, let $C_N^\varphi(t) = 2\sum_{n<N}\Lambda(n)n^{-1/2}A(\log n)\cos(t\log n)$ and let $R_N^\varphi(t)$ be the right side of (2) without the prime term for $G(u) = A(|u|)\cos(tu)$. Then $R_N^\varphi(t) - C_N^\varphi(t) = \sum_\rho \frac{1}{2}\big(\widehat{\varphi}(t_\rho - t)\overline{\widehat{\varphi}(\bar{t}_\rho - t)} + \widehat{\varphi}(t_\rho + t)\overline{\widehat{\varphi}(\bar{t}_\rho + t)}\big)$, a two-parameter family [CE13, Proposition 10]*; the case $\varphi = 1$, $t = 0$ is the first statement.*

Proof. Take $G(u) = (L - |u|)_+$, which is $g * \tilde{g}$ for $g = 1_{[0,L]}$, so $\widehat{G} = |\widehat{g}|^2 = F_L$ on $\mathbb{R}$; G is even, continuous, of bounded variation, supported in $[-L, L]$, and $\widehat{G}(z) = O(|z|^{-2})$ on horizontal lines, so (2) applies. Its prime side is $\sum_n \Lambda(n)n^{-1/2}\cdot 2(L - \log n)_+ = C(N)$. The pole terms are $\int G(u)(e^{u/2} + e^{-u/2})du = 8(\sqrt{N} + N^{-1/2} - 2)$. For the archimedean term, Gauss's integral $\psi(z) = \int_0^\infty (\frac{e^{-x}}{x} - \frac{e^{-zx}}{1-e^{-x}})dx$ and $\frac{1}{2\pi}\int \widehat{G}(r)e^{-irx/2}dr = G(\frac{x}{2})$ give, after $x = 2u$,

$$-L\log\pi + \int_0^\infty \Big[\frac{Le^{-2u}}{u} - \frac{2e^{-u/2}G(u)}{1-e^{-2u}}\Big]du,$$

and Proposition 2 evaluates the total as $R(N)$. The general window is the same computation with $G(u) = A(|u|)\cos(tu) = \frac{1}{2}\sum_\pm (k_\pm * \tilde{k}_\pm)(u)$, $k_\pm(u) = \varphi(u)e^{\pm itu}$. □

Proposition 2 (the remainder in closed form)**.** *The right side of (2) for $G(u) = (L - |u|)_+$, without the prime term, is $R(N)$ given by (1).*

Proof. The pole term is $8(\sqrt{N} + N^{-1/2} - 2)$. Extending the archimedean integral to $(0, \infty)$,

$$L\int_0^\infty \Big[\frac{e^{-2u}}{u} - \frac{2e^{-u/2}}{1-e^{-2u}}\Big]du + \int_0^\infty \frac{2u\,e^{-u/2}}{1-e^{-2u}}du - \int_L^\infty \frac{2(u-L)e^{-u/2}}{1-e^{-2u}}du.$$

The first integral is $\psi(\frac14) = -\gamma_E - \frac{\pi}{2} - 3\log 2$. Expanding $(1-e^{-2u})^{-1} = \sum_k e^{-2ku}$, the second is $2\sum_{k\ge0}(2k+\frac12)^{-2} = \frac12\psi'(\frac14) = \frac{\pi^2}{2} + 4G$, and the third is $2\sum_{k\ge0} N^{-(2k+1/2)}(2k+\frac12)^{-2}$, whose term $k=0$ cancels the $8N^{-1/2}$ of the pole. □

A second derivation, independent of Proposition 1, integrates von Mangoldt's explicit formula for $\sum_{n<N}\Lambda(n)n^{-1/2}$ twice in $\log N$ [In, MV]: it gives $C(N) = R(N) - M(N)$ with

$$R(N) = 8(\sqrt{N}-1) - 2\frac{\zeta'}{\zeta}(\tfrac12)\log N + 2\sum_{m\ge1}\frac{1-N^{-(2m+1/2)}}{(2m+\frac12)^2}, \qquad M(N) = 2\sum_\rho \frac{1-N^{it_\rho}}{t_\rho^2},$$

and comparing with (1) gives $\log\pi + \gamma_E + \frac{\pi}{2} + 3\log 2 = 2\frac{\zeta'}{\zeta}(\frac12)$, which is $\xi'(\frac12) = 0$. So the remainder is the smoothed sum of the smooth part of the partial sum (the pole, the constant $\frac{\zeta'}{\zeta}(\frac12)$ and the trivial zeros), and the margin is the smoothed sum of its oscillating part.

2.1. **RH in one variable.**

Theorem 3 (Suzuki [Su1, Su2])**.** *RH holds if and only if $M(N) \ge 0$ for all sufficiently large N. Under RH,*

$$M(N) = 4\sum_{\gamma>0}\frac{1-\cos(\gamma\log N)}{\gamma^2} \in \Big[0,\ 8\sum_{\gamma>0}\gamma^{-2}\Big], \tag{3}$$

so then $M(N) \ge 0$ for every $N > 1$.

Proof. By Proposition 1, $M(e^L) = \sum_\rho F_L(t_\rho)$. Under RH every term is ≥ 0, and pairing $\pm\gamma$ gives (3). Conversely, suppose $M(e^L) \ge 0$ for $L \ge L_0$. For $\operatorname{Re} s > \frac12$, since $|F_L(t_\rho)| \le (2+2e^{L/2})|t_\rho|^{-2}$ and $\sum|t_\rho|^{-2} < \infty$, Fubini gives

$$\int_0^\infty M(e^L)e^{-sL}dL = \sum_\rho \frac{2}{s(s^2+t_\rho^2)} =: \mathcal{G}(s),$$

meromorphic on $\mathbb{C}$ with poles at 0 and at $s = \pm it_\rho$. A pole s fixes $t_\rho^2 = -s^2$, so only $\pm t_\rho$ meet there, with equal nonzero residues, and no cancellation occurs. An off-line zero $\frac12 + \delta + i\gamma_0$, $\delta > 0$, has $\gamma_0 \ne 0$ because ζ has no real zeros in $(0,1)$, and gives a pole at $\delta + i\gamma_0$ off the real axis; the zeros on the line give poles on the imaginary axis. So $\mathcal{G}$ is analytic on $(0,\infty)$. By Landau's theorem [In, MV], the Laplace integral of a function that is ≥ 0 on $[L_0,\infty)$ has a singularity at the real point of its abscissa of convergence; it differs from $\mathcal{G}$ by an entire function, so the abscissa is ≤ 0, and $\mathcal{G}$ has no pole with $\operatorname{Re} s > 0$. □

Status: proved. It is Landau's oscillation method for a Riesz mean of $\psi(x) - x$ weighted by $n^{-1/2}$, and the theorem is Suzuki's: with $M = 2\Psi(\log N)$, RH holds if and only if $\Psi \ge 0$, and if and only if $\Psi = O(1)$ [Su1, Theorems 1.6, 1.7]; the criterion with the main term alone is [Su2, Theorem 1]. Our proof is independent. The constants in (3) are values of ξ at the centre: Hadamard's product gives $(\log\xi)''(s) = -\sum_\rho(s-\rho)^{-2}$, and at $s = \frac12$

$$(\log\xi)''(\tfrac12) = \sum_\rho\frac{1}{t_\rho^2} = -8 + \frac{\pi^2}{4} + 2G + \Big(\frac{\zeta''}{\zeta} - \Big(\frac{\zeta'}{\zeta}\Big)^2\Big)(\tfrac12) = 0.0462100, \tag{4}$$

unconditionally; under RH it is $2\sum_{\gamma>0}\gamma^{-2}$, and with the 4520 zeros below 5000 and the tail the two agree to $6\cdot10^{-10}$.

Corollary 4 (the sandwich)**.** *The following are equivalent: (1) RH; (2) $M(N) \ge 0$ for all large N; (3) $M(N) \le B$ for some B and all large N; (4) for every $N > 1$,*

$$R(N) - 4(\log\xi)''(\tfrac12) \ \le\ C(N) \ \le\ R(N), \qquad 4(\log\xi)''(\tfrac12) = 0.18484. \tag{5}$$

Proof. (1)⇒(4) is (3) with (4); (4)⇒(2),(3) are immediate; (2)⇒(1) is Theorem 3; for (3)⇒(1) apply the same argument to $B - M(e^L) \geq 0$, whose Laplace integral has the same poles off the real axis. □

Status: proved; the equivalences (1)–(3) and the bounds are in [Su1], and the evaluation (4) of $\sum_\rho t_\rho^{-2}$ is classical [Vo]. Writing the window this way makes (5) checkable without the zeros, which is what the verification below uses. The two sides are not redundant in general: a real zero $\beta \in (\frac{1}{2}, 1)$ would push M upwards only, and only the upper side sees it. For ζ, which has no such zeros, either side alone is equivalent to RH; for the Beurling system and the fake curve of [CE13] the lower side never fails and the upper side fails for all large N.

2.2. **What is known unconditionally.**

Proposition 5 (unconditional bound, crude form)**.** *For every $N > 1$, $M(N) \geq -8(3+\sqrt{N})S \geq -2.29 \cdot 10^{-11}(3+\sqrt{N})$, with $S = \sum_{\operatorname{Im}\rho > T_0} (\operatorname{Im}\rho)^{-2} \leq 2.854 \cdot 10^{-12}$.*

Proof. In Proposition 1 the zeros on the line contribute ≥ 0. By Platt and Trudgian [PT] every zero with $0 < \operatorname{Im}\rho \leq T_0$ is on the line, so an off-line zero belongs to a quadruple $\rho, \bar\rho, 1-\rho, 1-\bar\rho$ with $\rho = \beta + i\gamma$, $\gamma > T_0$, $0 < \delta = \beta - \frac{1}{2} < \frac{1}{2}$, and its t_ρ is one of $\pm\gamma \pm i\delta$. With $F_L(z) = (2 - e^{iLz} - e^{-iLz})/z^2$, $|F_L(z)| \leq (2 + N^\delta + N^{-\delta})/|z|^2 \leq (3+\sqrt{N})/\gamma^2$, and summing over the off-line ordinates, which come in quadruples $\pm\gamma \pm i\delta$, gives at most $8(3+\sqrt{N})S$ (a conservative constant). Stieltjes integration and Trudgian's bound for $N(u)$ [Tr] give $S \leq \int_{T_0}^\infty 2N(u)u^{-3}du \leq 2.854 \cdot 10^{-12}$. □

The factor $\sqrt{N}$ is the price of a hypothetical zero near $\operatorname{Re} s = 1$. A zero-density estimate says that such zeros are rare, and it improves the bound by many orders of magnitude. Let $N(\sigma, T)$ be the number of zeros $\beta + i\gamma$ with $\beta > \sigma$, $0 < \gamma < T$. Kadiri, Lumley and Ng [KLN] prove, for $T \geq H_0 = 3.0610046 \cdot 10^{10}$ and each σ in their Table 1 ($\sigma = 0.60, 0.65, \ldots, 0.85, 0.86, \ldots, 0.99$), with constants A_σ, B_σ and $k_\sigma \leq 1$ listed there,

$$N(\sigma,T) \leq A_\sigma (\log k_\sigma T)^{2\sigma} (\log T)^{5-4\sigma} T^{\frac{8}{3}(1-\sigma)} + B_\sigma (\log T)^2 \leq A_\sigma (\log T)^{5-2\sigma} T^{\frac{8}{3}(1-\sigma)} + B_\sigma (\log T)^2. \tag{6}$$

Theorem 6 (explicit unconditional lower bound)**.** *Let $T_0 = 3 \cdot 10^{12}$, $\widetilde{N}(t) = \frac{t}{2\pi}\log\frac{t}{2\pi}$, and let $\frac{1}{2} = \sigma_0 < \sigma_1 < \cdots < \sigma_J < \sigma_{J+1} = 1$ be $\frac{1}{2}$, the twenty values of σ in* [KLN, Table 1]*, and 1. Put*

$$S_{1/2} = \int_{T_0}^\infty \widetilde{N}(t)\, t^{-3}\, dt, \qquad S_j = 2\int_{T_0}^\infty \min\Big(N_{\mathrm{KLN}}(\sigma_j, t), \tfrac{1}{2}\widetilde{N}(t)\Big) t^{-3}\, dt \quad (1 \leq j \leq J),$$

with N_{KLN} the right side of (6). Then for every $N > 1$

$$M(N) \geq -E(N), \qquad E(N) = 4\Big[(3 + N^{\sigma_1 - \frac{1}{2}})S_{1/2} + \sum_{j=1}^{J} \big(N^{\sigma_{j+1} - \frac{1}{2}} - N^{\sigma_j - \frac{1}{2}}\big) S_j\Big]. \tag{7}$$

One may take $S_{1/2} = 1.4822 \cdot 10^{-12}$, $S_j = S_{1/2}$ to four digits for $\sigma_j \leq 0.75$, $S(0.80) = 4.71 \cdot 10^{-13}$, $S(0.85) = 8.63 \cdot 10^{-15}$, $S(0.90) = 1.57 \cdot 10^{-16}$, $S(0.95) = 2.85 \cdot 10^{-18}$, $S(0.99) = 1.14 \cdot 10^{-19}$, and $E(N)$ takes the values of Table 1. In particular $|\min(M(N), 0)| < 10^{-3}$ for $N < 5 \cdot 10^{24}$ and $E(N) < 0.043$ for $N < 2.3 \cdot 10^{29}$.

Proof. As in Proposition 5, $M(N) \geq -\sum' |F_L(t_\rho)|$, the sum over the off-line zeros, all of which have $|\operatorname{Im}\rho| > T_0$, and $|F_L(t_\rho)| \leq (3 + N^{|\delta|})/\gamma^2$. A zero $\rho = \beta + i\gamma$ with $\beta > \frac{1}{2}$, $\gamma > T_0$, brings the quadruple $\rho, \bar\rho, 1-\rho, 1-\bar\rho$, so

$$\sum{}' |F_L| \leq 4 \sum_{\beta > \frac{1}{2},\ \gamma > T_0} \frac{3 + N^{\beta - 1/2}}{\gamma^2}.$$

Write $\Sigma(\sigma)=\sum_{\beta>\sigma,\gamma>T_0}\gamma^{-2}$. Since $N^{\beta-1/2}\le N^{\sigma_{j+1}-1/2}$ when $\sigma_j<\beta\le\sigma_{j+1}$, Abel summation over the bins gives

$$\sum_{\beta>\frac12,\gamma>T_0}\frac{N^{\beta-1/2}}{\gamma^2}\le N^{\sigma_1-\frac12}\Sigma(\tfrac12)+\sum_{j=1}^{J}\big(N^{\sigma_{j+1}-\frac12}-N^{\sigma_j-\frac12}\big)\Sigma(\sigma_j).$$

It remains to show $\Sigma(\frac12)\le S_{1/2}$ and $\Sigma(\sigma_j)\le S_j$. Since $N(\sigma,T_0)=0$, Stieltjes integration gives $\Sigma(\sigma)=2\int_{T_0}^{\infty}N(\sigma,t)t^{-3}dt$. By the symmetry $\beta\leftrightarrow1-\beta$, at most half of the zeros up to height t lie to the right of the line, so $N(\sigma,t)\le\frac12 N(t)$; and $N(t)\le\widetilde{N}(t)$ for $t\ge T_0$, since $N(t)=\frac{t}{2\pi}\log\frac{t}{2\pi e}+\frac78+\theta(0.112\log t+0.278\log\log t+2.51+0.2/t)$ with $|\theta|\le1$ [Tr], and the term $-\frac{t}{2\pi}$ dominates. This gives $\Sigma(\frac12)\le S_{1/2}$. For $\sigma_j\ge0.6$, (6) applies since $T_0>H_0$, and $\Sigma(\sigma_j)\le S_j$. The values used come from adaptive quadrature in $u=\log t$ (`ebound.py`); they are rigorous upper bounds, because each integrand $g(u)$ is positive and decreasing for $t\ge T_0$ (both $N_{\mathrm{KLN}}(\sigma,t)t^{-2}$ and $\widetilde{N}(t)t^{-2}$ decrease), so a left Riemann sum of step 10^{-4} on $[\log T_0,\log T_0+120]$ plus an explicit tail bounds the integral from above, and every quadrature value exceeds that bound (`ebound_rig.py`; the ratio is at most 0.9984). For instance $S_{1/2}=(\log\frac{T_0}{2\pi}+1)/(2\pi T_0)=1.47970\cdot10^{-12}$ exactly, the Riemann sum gives $1.47977\cdot10^{-12}$ and the value used is $1.48222\cdot10^{-12}$. □

Corollary 7 (unconditional two-sided bound)**.** *For every $N>1$,*

$$-E(N)\le M(N)\le 4(\log\xi)''(\tfrac12)+E(N)=0.18484+E(N).$$

Proof. The lower side is Theorem 6. For the upper side split $M=\sum_\rho F_L(t_\rho)$ into zeros on and off the line. A pair $\pm\gamma$ on the line contributes $4(1-\cos\gamma L)/\gamma^2\le8/\gamma^2$, and the off-line zeros contribute at most $E(N)$ in absolute value, as in the proof of Theorem 6. Identity (4) holds without RH, with $t_\rho=(\rho-\frac12)/i$ complex for an off-line zero. A quadruple $\pm\gamma\pm i\delta$ contributes to it $4(\gamma^2-\delta^2)/(\gamma^2+\delta^2)^2>0$, since $\delta<\frac12<\gamma$. Hence $\sum_{\text{on the line}}\gamma^{-2}\le(\log\xi)''(\frac12)$, and the zeros on the line contribute at most $4(\log\xi)''(\frac12)$. □

So without any hypothesis (5) holds up to $E(N)$ on both sides; RH removes $E(N)$.

TABLE 1. The off-line contribution: crude bound of Proposition 5 against $E(N)$ of Theorem 6.

N	10^{10}	10^{15}	10^{20}	10^{25}	10^{27}	10^{30}
$8(3+\sqrt{N})S$	$2.3\cdot10^{-6}$	$7.2\cdot10^{-4}$	0.23	72	720	$2.3\cdot10^{4}$
$E(N)$	$1.0\cdot10^{-8}$	$4.7\cdot10^{-7}$	$2.4\cdot10^{-5}$	$1.3\cdot10^{-3}$	$6.3\cdot10^{-3}$	0.072

The crude bound is below the least margin observed (0.0427, Proposition 8) as long as $N<3.5\cdot10^{18}$; the explicit bound (7) stays below it up to $2.3\cdot10^{29}$. Beyond that it grows like $N^{0.35}$ (the bin $0.80<\beta\le0.85$ dominates) and says nothing. This is the familiar fact that verifying RH to a height H, together with zero-density estimates, gives bounds of RH quality only in a range polynomial in H [Jn]; it comes from the zeros, not from the primes. Theorem 6 holds whether or not RH is true, and it is silent for $N>10^{30}$.

2.3. **Verification.**

Proposition 8 (the sandwich verified)**.** (1) *For every real N with $1<N\le10^{10}$, $0<M(N)<0.14$; more precisely $M(N)\ge0.0148\log N$ for $1<N\le e^{1/15}$, $M(N)\ge0.0276$ for $e^{1/30}\le N\le10^{10}$, $M(N)\ge0.0429$ for $10^3\le N\le10^{10}$, and $M(N)\le0.1379$ throughout.*

(2) *For every real N with $10^{10}\le N\le10^{12}$, $0.04268\le M(N)\le0.13122$. The least value of M on $(10^3,10^{12}]$ is attained near $N=1.056\cdot10^{11}$.*

(3) $M(N) \geq 0.0415$ *for every real* N *with* $10^8 \leq N \leq 10^{15}$, $M(N) \geq 0.0481$ *for* $10^{15} \leq N \leq 10^{27}$, *and* $M(N) > 0$ *for* $10^{15} \leq N \leq 2.15 \cdot 10^{29}$; $M(N) \leq 0.1377$ *for* $10^8 \leq N \leq 10^{27}$ *and* $M(N) < 0.18484$ *for* $10^8 \leq N \leq 4.28 \cdot 10^{29}$; *so* (5) *holds on* $[10^8, 2.15 \cdot 10^{29}]$. *Here the* 4520 *zeros below* 5000 *are certified: with ball arithmetic,* $N(5000) = 4520$ *and each ordinate used lies within* $9.1 \cdot 10^{-12}$ *of the true one.*

Since $0.14 < 0.18484$, *parts (1) and (3) give the two-sided inequality* (5) *for every real* N *with* $1 < N \leq 2.15 \cdot 10^{29}$; *part (2) checks the range* $[10^{10}, 10^{12}]$ *independently of the zeros.*

Proof. (1) Write $m(L) = M(e^L)$ and $r(L) = R(e^L)$. *Near* $N = 1$, no prime power lies in $(1, 2)$ and $M = R$; elementary bounds on the series in (1) give $r(L) \geq 0.01488L$ for $L \leq \frac{1}{15}$ and $r(L) \leq 0.1016$ for $L \leq \frac{1}{30}$, and on $[\frac{1}{30}, \log 2]$ monotonicity of each piece on 3300 intervals gives $0.0670 \leq M \leq 0.1282$. *From* 2 *to* 10^{10}: between consecutive prime powers $q < q'$, $m(L) = r(L) - 2AL + 2B$ with $A = \sum_{n \leq q} \Lambda(n) n^{-1/2}$ and $B = \sum_{n \leq q} \Lambda(n) n^{-1/2} \log n$ constant, and $m'' = r'' = 2e^{L/2} - 2e^{-L/2}/(e^{2L} - 1) > 0$ for $N \geq 2$, so m is convex there and its minimum on $[\log q, \log q']$ is bounded below by the tangent lines at the two ends; its maximum is at an end. *Arithmetic*: Python's `decimal` at 34 digits, whose $+, -, \times, /, \sqrt{}, \log, \exp$ are correctly rounded; π, γ_E, G entered to 48 digits; a margin of 10^{-15} is subtracted from every lower bound and added to every upper bound. The computation ran over all 455 062 595 prime powers below 10^{10}, from a segmented sieve; 8 intervals below $N = 100$ needed subdivision. The least lower bound for $N \geq 10^3$ is 0.042997, on $[34\,186\,367, 34\,186\,499]$, where the minimum of M below 10^{10} is attained; the largest upper bound is 0.137816 at $N = 319\,439$.

(2) Same convexity argument, over all 37 607 992 088 prime powers below 10^{12} (the count equals $\pi(10^{12}) = 37\,607\,912\,018$ plus the 80 070 powers p^k, $k \geq 2$, which checks the sieve). On $[\log q, \log q']$, $\min m \geq \max\big(m(a) - (b-a)\max(0, -m'(a)),\ m(b) - (b-a)\max(0, m'(b))\big)$. The arithmetic is IEEE double precision. The operations $+, -, \times, /, \sqrt{}$ are correctly rounded; the library logarithm was checked at run time to be within 0.5 ulp of the correctly rounded value on $2 \cdot 10^4$ test points, and an error of 2 ulp is assumed in the bounds. The sums A and B are accumulated in segments of length $5 \cdot 10^7$, each total exactly rounded (`math.fsum`) and the totals added exactly in rational arithmetic. Every lower bound is lowered, and every upper bound raised, by an explicit bound for all rounding errors (the error in $A(10^{12}) = 1\,999\,997.3576058$ is at most $2.1 \cdot 10^{-9}$, and the error of one evaluation of m is of order 10^{-7}). The least lower bound is 0.042686, near $N = 105\,606\,345\,001$; the largest upper bound on $[10^{10}, 10^{12}]$ is 0.131211, near $N = 742\,864\,907\,663$. The same program reproduces part (1) on $[10^3, 10^{10}]$ (0.042997, 0.137816, 455 062 595 prime powers). The minimum near $1.056 \cdot 10^{11}$ is also visible from the zeros: there the sum over the 4520 zeros below 5000 in part (3) is 0.041711, and the difference 0.00098 is the mean contribution $4 \sum_{\gamma > 5000} \gamma^{-2} = 0.00098$ of the omitted zeros.

(3) Unconditionally $M = \sum_\rho F_L(t_\rho)$; the zeros up to $3 \cdot 10^{12}$ are on the line and contribute ≥ 0, so dropping those not in the list and bounding the rest as in Proposition 5, $M(N) \geq 4 \sum_{\gamma \in \text{list}} (1 - \cos \gamma L)/\gamma^2 - 8(3 + \sqrt{N})S$. The sum is Lipschitz with constant $4 \sum 1/\gamma = 14.132$; on a grid of 805 906 points covering $[\log 10^8, \log 10^{15}]$, the least value of the sum on a cell, minus the Lipschitz loss and minus $8(3 + \sqrt{N})S$ at the right end of the cell (at most $7.2 \cdot 10^{-4}$), is 0.04156; it is attained near $N = 1.06 \cdot 10^{11}$, where the sum is 0.041711, the Lipschitz loss $1.41 \cdot 10^{-4}$ and the tail term $7 \cdot 10^{-6}$. For $10^{15} \leq N \leq 10^{30}$ the same computation with $E(N)$ of Theorem 6 in place of $8(3 + \sqrt{N})S$, on a grid of step $2 \cdot 10^{-5}$ in L (1 726 941 points), gives the least lower bound 0.048104 up to 10^{27}, 0.047549 up to 10^{28} and 0.034188 up to 10^{29}; the first cell where $W - E \leq 0$ is at $N = 2.159 \cdot 10^{29}$, where $E = 0.0418$ and the sum is near its least value on the range, 0.038908 at $N = 2.18 \cdot 10^{29}$ (Figure 1). For the upper side, the zeros on the line above 5000 contribute at most $8 \sum_{\gamma > 5000} \gamma^{-2}$, and by Stieltjes integration with $N(5000) = 4520$ and Trudgian's bound for $N(t)$ this is at most $8 \cdot 2.446 \cdot 10^{-4} = 1.957 \cdot 10^{-3}$. (As a check, $2(\sum_{\text{list}} 4\gamma^{-2} + 4 \cdot 2.446 \cdot 10^{-4}) = 0.184841$, against $4(\log \xi)''(\frac{1}{2}) = 0.184840$.) So $M \leq W + 1.957 \cdot 10^{-3} + E(N)$. The largest upper bound on the cells, the maximum of W plus the Lipschitz gain, plus $1.957 \cdot 10^{-3} + E(N)$, is 0.137298 up to 10^{20}, 0.137626 up to 10^{27} and 0.161553 up to 10^{29}, and it first reaches 0.18484 at $N = 4.28 \cdot 10^{29}$; the maximum

of the cell upper bounds on $[10^8, 10^{30}]$ is 0.135632 (W itself peaks at 0.135491). An error δ in each ordinate changes the sum by at most $4\delta L \sum \gamma^{-2} \le 3.2\,\delta$ for $N \le 10^{30}$. □

Status: (1) is computer-assisted, with correctly rounded arithmetic, and was checked in three independent ways (an extended-precision scan of A and B to 10^{10}, 45-digit values below $2 \cdot 10^6$, and an independent sieve on three windows). (2) is computer-assisted in double precision with a priori error bounds; it relies on the stated accuracy of the library logarithm, which was tested but not proved. (3) uses the zeros below 5000, certified with Arb [Arb] through python-flint: `acb.zeta_zeros` returns balls guaranteed to contain the n-th zeros, isolated by sign changes of Hardy's Z with the count checked by Turing's method. The 4520-th zero lies in $[4999.32968137297, 4999.32968137298]$ and the 4521-st above 5000.2343169313, so the list is complete, and it is accurate to $9.1 \cdot 10^{-12}$, which changes W by less than $3 \cdot 10^{-11}$. The grid evaluation of W in double precision has error below 10^{-10} (each term $4(1-\cos\gamma L)/\gamma^2$ to a few units in the last place of $\gamma L \le 3.5 \cdot 10^5$). Neither reaches the statement for all large N: the primes cost one evaluation per prime power, and the zeros stop near $2 \cdot 10^{29}$ by Theorem 6.

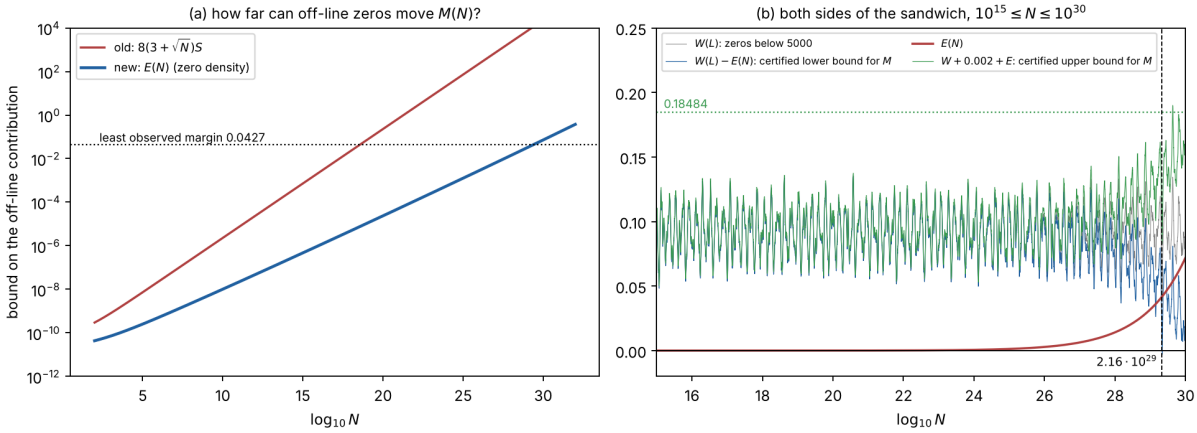


FIGURE 1. (a) The bound on how far off-line zeros could move $M(N)$: the crude bound of Proposition 5 and $E(N)$ of Theorem 6, which uses the zero density of [KLN]. (b) For $10^{15} \le N \le 10^{30}$: the contribution W of the 4520 zeros below 5000 (grey), the bound $E(N)$ (red) the certified lower bound $W - E$ for M (blue), positive up to $2.16 \cdot 10^{29}$ (dashed), and the upper bound $W + 1.96 \cdot 10^{-3} + E$ (green), below 0.18484 up to $4.28 \cdot 10^{29}$.

The wave. Formula (3) says what the primes must produce: an almost periodic function of $\log N$ with mean $4\sum_{\gamma>0}\gamma^{-2} = 0.0924$, frequencies at the ordinates and amplitudes $4/\gamma^2$. From the 5 762 859 prime powers below 10^8, at 6000 values of $\log N$, M agrees with the right side of (3) truncated at the zeros below 5000 to $5.4 \cdot 10^{-5}$; its spectrum in $\log N$, from the primes alone, has peaks at 14.132, 21.023, 25.014, 30.425, ... with heights agreeing with $4/\gamma^2$ to 10^{-5} (Figure 2). Over $4 \cdot 10^5$ values of $\log N$ in $[\log 2, \log 10^8]$ the mean is 0.092367 (predicted $2(\log\xi)''(\frac{1}{2}) = 0.092420$) and the standard deviation 0.017219 (predicted $\sqrt{8\sum\gamma^{-4}} = 0.017245$ for independent phases).

A least-squares fit of $m(L) = M(e^L)$ on the first 40 ordinates, from the prime powers below 10^7, gives cosine amplitudes equal to $-4/\gamma_k^2$ to within 0.4% and sine amplitudes below 10^{-5}, as (3) predicts under RH. The decay $4/\gamma^2$ is the regularity of m: it is continuous, and its slope jumps by $-2\Lambda(q)/\sqrt{q}$ at $L = \log q$ (Proposition 9).

Where the difficulty sits. By blocks at $N = 0.999 \cdot 10^8$: $R(N) = 79853.63$, the primes give 79665.63, the prime squares 130.68, the higher powers 57.21, and $M = 0.114$. The squares are controlled without RH; the whole difficulty is in the sum over primes, which has to be known to $O(1)$ out of

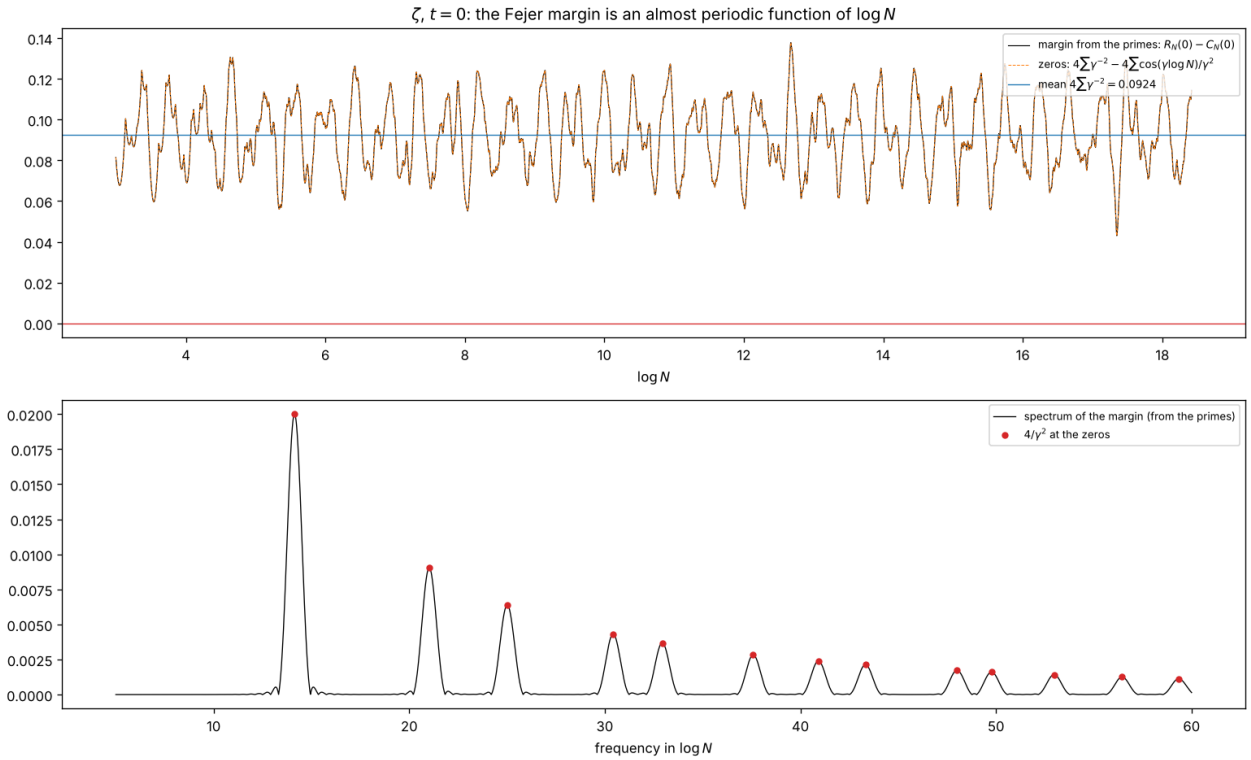


FIGURE 2. Top: the margin $M(N)$ computed from the primes below 10^8 (black) and the right side of (3) from the 4520 zeros below 5000 (dashed); mean 0.0924, least value 0.043. Bottom: its spectrum in the variable $\log N$, from the primes, against $4/\gamma^2$ at the zeros.

$8\sqrt{N}$. Without RH the best bound for that sum is of size $\sqrt{N}\exp(-c(\log N)^{3/5}(\log\log N)^{-1/5})$, and closing the gap to $O(1)$ is RH. For the Davenport–Heilbronn function the same construction gives its own window $w_f = 0.78456$; its margin computed from its coefficients up to $2 \cdot 10^7$ stays inside, and the model with its zeros predicts the first exit near $N \approx 2.3 \cdot 10^8$ [CE13]: in one variable the criterion is the simplest, but the slowest to detect an off-line zero.

3. The local minima and the slope

The proof of Proposition 8 uses that the margin is convex between consecutive prime powers. This has a sharper consequence: the minima of M are a discrete set, one at most between two prime powers, with an explicit location and value. The criterion of Theorem 3, a statement about every real N, becomes a statement about countably many points given by sums over primes. We keep the notation of the proof of Proposition 8: $m(L) = M(e^L)$, $r(L) = R(e^L)$, and for consecutive prime powers $q < q'$

$$A_q = \sum_{n \le q} \frac{\Lambda(n)}{\sqrt{n}}, \qquad B_q = \sum_{n \le q} \frac{\Lambda(n)\log n}{\sqrt{n}}, \qquad m(L) = r(L) - 2A_q L + 2B_q \quad (\log q \le L \le \log q').$$

Write $c_0 = \log\pi + \gamma_E + \frac{\pi}{2} + 3\log 2 = 5.3721834$ and $K_0 = -16 + \frac{\pi^2}{2} + 4G = -7.4013354$, so that by (1)

$$r(L) = 8e^{L/2} - c_0 L + K_0 - \tau(L), \qquad \tau(L) = 2\sum_{k \ge 1} \frac{e^{-(2k+\frac{1}{2})L}}{(2k+\frac{1}{2})^2}. \tag{8}$$

Proposition 9 (the local minima of the margin)**.** (1) *On $L > L_c = 0.2812$ ($N > 1.3247$), $r''(L) = 2e^{L/2} - 2e^{-L/2}/(e^{2L}-1) > 0$; so m is strictly convex on every interval between consecutive prime powers.*

(2) *At a prime power $q = p^k$ the derivative of m jumps down: $m'(\log q^+) - m'(\log q^-) = -2\Lambda(q)/\sqrt{q} < 0$. Hence no prime power is a local minimum of m.*

(3) *For $N \geq 2$, the local minima of M are exactly the points $N_q^* = e^{L_q^*}$ with $r'(L_q^*) = 2A_q$ and $q < N_q^* < q'$, one at most between two consecutive prime powers; at such a point*

$$M(N_q^*) = 2B_q + K_0 - 8u_q\big(\log u_q - 1\big) + \epsilon_q, \qquad u_q = \frac{2A_q + c_0}{4}, \tag{9}$$

with $N_q^* = u_q^2(1 + O(u_q^{-5}))$ and ϵ_q coming from τ: $|\epsilon_q| \leq 0.035$, and $|\epsilon_q| \leq 3 \cdot 10^{-6}$ for $N_q^* > 100$.

(4) *On $1 < N < 2$, where $M = R$, the only local minimum is at $N = 1.5904$, with $M = 0.0793$, and $M(N) \to 0$ as $N \to 1^+$.*

Proof. (1) Differentiating (8), $r''(L) = 2e^{L/2} - 2\sum_{k\geq 1} e^{-(2k+\frac{1}{2})L} = 2e^{L/2} - 2e^{-L/2}/(e^{2L}-1)$, which is positive exactly when $e^L(e^{2L}-1) > 1$, that is for $e^L > 1.32472$ (the real root of $x^3 - x = 1$), $L > 0.2812$. Between prime powers $m'' = r''$. (2) $C(e^L) = 2\sum_n \Lambda(n) n^{-1/2}(L - \log n)_+$ gains the slope $2\Lambda(q)/\sqrt{q}$ at $L = \log q$, and r is smooth. A local minimum at $\log q$ would need $m'(\log q^-) \leq 0 \leq m'(\log q^+)$, which the jump forbids. (3) By (1) and (2), for $L \geq \log 2$ every local minimum is an interior critical point of a strictly convex piece, so it is the unique solution of $m' = r' - 2A_q = 0$ in that piece, if it lies there. With $r'(L) = 4e^{L/2} - c_0 + \tau_1(L)$, $\tau_1 = -\tau'$, the equation is $4e^{L/2} = 2A_q + c_0 - \tau_1(L)$, so $e^{L^*/2} = u_q - \tau_1(L^*)/4$. Inserting $e^{L^*/2} = u_q$ and $L^* = 2\log u_q$ in $m = r - 2A_qL + 2B_q$ gives $8u_q - 2\log u_q\,(c_0 + 2A_q) + K_0 + 2B_q = 2B_q + K_0 - 8u_q(\log u_q - 1)$; the difference with the exact value comes from τ and τ_1, of order $N^{-5/2}$, and its size is the computed one. (4) On $(1,2)$ the sum over primes is empty, $r'' < 0$ for $N < 1.3247$ and $r'' > 0$ after, r' vanishes once on $(1.3247, 2)$, at $N = 1.5904$, and $r(0) = 0$ with $r(L) \geq 0.0148L$ near 0 (Proposition 8). □

Status: proved; the bounds for ϵ_q and the values in (4) are numerical, from the exact solution of $r'(L) = 2A_q$ by Newton's method. Part (1) is the convexity already used in Proposition 8; parts (2) and (3) are new to this paper as far as we know.

Corollary 10 (a discrete criterion)**.** *RH holds if and only if $M(N_q^*) \geq 0$ for every prime power q with $q < N_q^* < q'$, that is, if and only if*

$$2B_q \;\geq\; 8u_q\big(\log u_q - 1\big) - K_0 - \epsilon_q \qquad \text{whenever } q < N_q^* < q'. \tag{10}$$

Proof. If RH holds, $M \geq 0$ everywhere by Theorem 3. Conversely, assume (10) and suppose $M(N_0) < 0$ for some $N_0 \geq 2$ ($M > 0$ on $(1,2)$ by Proposition 8). Let $L_0 = \log N_0$. If $m'(L_0^-) > 0$, follow m to the left: on the convex piece containing L_0 it decreases towards the left end unless it meets a critical point, which would be a local minimum with value $< m(L_0) < 0$; at a prime power the left derivative is larger than the right one by (2), so m keeps decreasing to the left. Either we meet a local minimum with negative value, contradicting (10), or we reach $L = \log 2$ with $m(\log 2) < 0$, contradicting Proposition 8. If $m'(L_0^+) < 0$, follow m to the right in the same way: either a local minimum with negative value appears, or m decreases for all $L > L_0$; in the second case $M \leq M(N_0)$ for all $N \geq N_0$, so M is bounded above, and Corollary 4(3) gives RH, hence $M \geq 0$, a contradiction. One of the two cases occurs, since $m'(L_0^-) \geq m'(L_0^+)$ and they cannot be both ≤ 0 and ≥ 0 unless L_0 is itself a critical point, which is then a local minimum with negative value. □

So RH is equivalent to countably many inequalities between two explicit sums over the prime powers below q, at the prime powers q for which the stationary point u_q^2 falls before the next prime power. The reduction removes the need to look at every real N; it does not make the inequality easier. Both sides of (10) are of size $4\sqrt{q}\log q$, and their difference is $M(N_q^*) \in [0, 0.185]$ under RH.

Proposition 11 (no margin)**.** *If RH holds, then* $\liminf_{N\to\infty} M(N) = 0$*, and the local minima of Proposition 9 accumulate at* 0*:* $\liminf M(N_q^*) = 0$ *over the prime powers* q *with* $q < N_q^* < q'$*.*

Proof. Under RH, $m(L) = 4\sum_{\gamma>0}(1-\cos\gamma L)/\gamma^2$ by (3), an absolutely convergent trigonometric series, so m is uniformly almost periodic with mean $2(\log\xi)''(\frac12) = 0.0924$. Given $\varepsilon \in (0, 0.09)$, choose T with $8\sum_{\gamma>T}\gamma^{-2} < \varepsilon/2$. By Dirichlet's simultaneous approximation theorem there are arbitrarily large L_0 with $\|\gamma L_0/2\pi\|$ so small for every $0 < \gamma \le T$ that $4\sum_{\gamma\le T}(1-\cos\gamma L_0)/\gamma^2 < \varepsilon/2$; then $m(L_0) < \varepsilon$. Since m is almost periodic with mean $0.0924 > \varepsilon$, the set $\{m > \varepsilon\}$ is relatively dense, so there are $a < L_0 < b$ with $m(a), m(b) > \varepsilon$. The minimum of m on $[a, b]$ is $< \varepsilon$, is not at a or b, and is not at a prime power by Proposition 9(2); so it is a local minimum N_q^* with $M(N_q^*) < \varepsilon$, and $N_q^* \ge e^a$ is arbitrarily large. □

Status: proved, conditional on RH, and needs no linear independence of the ordinates. So no proof of $M \ge 0$ can rest on a uniform positive lower bound: under RH the margin is positive without margin, as the least eigenvalues of the windows are in Section 6.

The minima from the primes. Below 10^7 there are 665 134 prime powers and 10 341 of the intervals between them, 1.6%, contain a local minimum. The least is $M(N_q^*) = 0.055041$ at $N_q^* = 3099.71$, between $q = 3089$ and $q' = 3109$; the largest is 0.137816 at $N_q^* = 319\,433.67$, next to the maximum of M found in Proposition 8; the mean is 0.091059, against the mean 0.0924 of M. The first ones are $M = 0.0718, 0.0754, 0.0862, 0.0652$ at $N_q^* = 2.448, 3.598, 4.298, 5.934$. The minimum 0.042997 of M below 10^{10}, near $N = 3.4\cdot10^7$, is necessarily one of these points.

Both sides of (10) are of size $4\sqrt{q}\log q$, and their difference is at most 0.185. In this subsection we remove the common smooth part of the two sums A and B and look at what remains. The criterion becomes a condition on the fluctuation of one weighted Chebyshev sum, B, at the points where the fluctuation of the other, A, vanishes. The identities below are exact and elementary, and they contain no information beyond Corollary 10. What they do is separate the two roles that the primes play at a minimum, where it lies and how deep it is, and show the size of the problem.

For real $x \ge 1$ let $A(x) = \sum_{n\le x}\Lambda(n)n^{-1/2}$ and $B(x) = \sum_{n\le x}\Lambda(n)n^{-1/2}\log n$, so that $A_q = A(q)$ and $B_q = B(q)$, and define their fluctuations

$$\alpha(x) = A(x) - \Big(2\sqrt{x} - \frac{c_0}{2}\Big), \qquad \beta(x) = B(x) - \big(2\sqrt{x}\log x - 4\sqrt{x}\big). \tag{11}$$

The smooth parts are the integrals $\int_0^x t^{-1/2}\,dt = 2\sqrt{x}$ and $\int_0^x t^{-1/2}\log t\,dt = 2\sqrt{x}\log x - 4\sqrt{x}$, that is, the two sums with $d\psi(t)$ replaced by dt; the constant $-c_0/2$ in α is the one that the closed form (8) imposes. By the prime number theorem and partial summation, $\alpha(x) = o(\sqrt{x})$ and $\beta(x) = o(\sqrt{x}\log x)$; under RH, $\psi(x)-x = O(\sqrt{x}\log^2 x)$ gives $\alpha(x) = O(\log^2 x)$ and $\beta(x) = O(\log^3 x)$. Between two consecutive prime powers A and B are constant, so α decreases, with $d\alpha/d\log x = -\sqrt{x}$, and at a prime power q it jumps up by $\Lambda(q)/\sqrt{q}$. It is a sawtooth (Figure 3(a)).

Proposition 12 (the two sums give the slope and the value of the margin)**.** *Let* $N \ge 2$ *not be a prime power,* $L = \log N$*, and* $\tau_1 = -\tau'$*, that is,* $\tau_1(L) = 2\sum_{k\ge1} e^{-(2k+\frac12)L}/(2k+\frac12)$*. Then*

$$\alpha(N) = \frac{\tau_1(L) - m'(L)}{2}, \tag{12}$$

$$\beta(N) - L\,\alpha(N) = \frac{m(L) - K_0 + \tau(L)}{2}. \tag{13}$$

Equivalently, $M(N) = K_0 + 2\big(\beta(N) - \alpha(N)\log N\big) - \tau(\log N)$ *and* $\beta(N) = \frac12\big(m - Lm' - K_0 + \tau + L\tau_1\big)$*.*

Proof. Let $q < N < q'$ be the consecutive prime powers around N. On $(\log q, \log q')$ we have $A(e^L) = A_q$, $B(e^L) = B_q$ and $m(L) = r(L) - 2A_qL + 2B_q$. Differentiating (8), $r'(L) = 4e^{L/2} - c_0 + \tau_1(L)$, so

$$m'(L) = 4e^{L/2} - c_0 + \tau_1(L) - 2A_q = \tau_1(L) - 2\alpha(N),$$

which is (12). Write $A_q = 2e^{L/2} - \frac{c_0}{2} + \alpha$ and $B_q = 2Le^{L/2} - 4e^{L/2} + \beta$, with $\alpha = \alpha(N)$, $\beta = \beta(N)$. Inserting these and (8) in $m = r - 2A_qL + 2B_q$ gives

$$m = 8e^{L/2} - c_0L + K_0 - \tau - 2L\Big(2e^{L/2} - \frac{c_0}{2} + \alpha\Big) + 2\big(2Le^{L/2} - 4e^{L/2} + \beta\big) = K_0 - \tau + 2(\beta - L\alpha).$$

The terms in $e^{L/2}$, $Le^{L/2}$ and L cancel, and what remains is (13). □

Status: proved, elementary. So α is, up to the factor $-\frac{1}{2}$ and the trivial-zero term τ_1, the slope of the margin in the variable $\log N$. The combination $\beta - \alpha \log N$ is its value, up to the constant K_0 and the factor 2. The trivial-zero terms are small: $\tau(\log 2) = 0.0615$ and $\tau_1(\log 2) = 0.165$, and both are $O(N^{-5/2})$.

Corollary 13 (the minima as zeros of α)**.** *For $N \geq 2$, the local minima of M are the points N_q^* of an interval (q, q') between consecutive prime powers at which the sawtooth α crosses the level $\frac{1}{2}\tau_1(\log N)$, a quantity below 0.083 that tends to 0. At such a point*

$$M(N_q^*) = K_0 + 2\,\beta(N_q^*) - \tau(L_q^*) - L_q^*\,\tau_1(L_q^*), \qquad L_q^* = \log N_q^*. \tag{14}$$

Consequently RH holds if and only if

$$\beta(N_q^*) \;\geq\; -\frac{K_0}{2} + \frac{\tau(L_q^*) + L_q^*\tau_1(L_q^*)}{2} \qquad \text{at every local minimum } N_q^*, \tag{15}$$

where $-K_0/2 = 8 - \pi^2/4 - 2G = 3.7006677$ and the correction is below $1.8 \cdot 10^{-5}$ for $N_q^* > 100$.

Proof. By Proposition 9(3) the local minima are the solutions of $m'(L) = 0$ inside the intervals $(\log q, \log q')$, and by (12) these are the points where $\alpha = \frac{1}{2}\tau_1$. On $L \geq \log 2$ the function $\alpha - \frac{1}{2}\tau_1$ is strictly decreasing between prime powers, since its derivative in L is $-e^{L/2} + \sum_{k\geq 1} e^{-(2k+\frac{1}{2})L} < 0$; so there is at most one such point in each interval, and α crosses the level from above. At that point (13) gives $m = K_0 - \tau + 2\beta - 2L\alpha = K_0 - \tau + 2\beta - L\tau_1$, which is (14). The equivalence with RH is Corollary 10. The bound for the correction is the computed maximum of $\frac{1}{2}(\tau + L\tau_1)$ over the minima with $N_q^* > 100$. □

In words: a minimum of the margin lies where the staircase A comes down to its smooth curve $2\sqrt{x} - \frac{c_0}{2}$, inside an interval free of prime powers. Its depth is twice the excess of the staircase B over its smooth curve at that point, minus 7.40. The large term $\alpha \log N$ in (13) is switched off exactly at the minima, which is why the value of a minimum depends on β alone. One function of the primes locates the minima and the other measures them.

The numbers. On the 10 341 local minima below 10^7 the identities hold to the precision of the computation: $|\alpha(N_q^*) - \frac{1}{2}\tau_1| \leq 5.5 \cdot 10^{-12}$, and (14) holds to $2.2 \cdot 10^{-10}$. The values of β at the minima lie in $[3.72819, 3.79680]$, with mean 3.74621, against the threshold 3.70067 of (15) (Figure 3(c)). The least value is at $N_q^* = 3099.71$, the least minimum of M below 10^7, and its excess over the threshold, 0.02752, is half of $M(N_q^*) = 0.05504$. The least value of M below 10^{10}, 0.042997, corresponds to $\beta = 3.72217$.

The minima one after another. Let $N_1^* < N_2^* < \cdots$ be the local minima of M on $[2, \infty)$, and write $L_j = \log N_j^*$, $u_j = \sqrt{N_j^*}$ and $\beta_j = \beta(N_j^*)$.

Proposition 14 (the recurrence of the minima)**.** *The first minimum is $N_1^* = 2.44777$, with $\beta_1 = (\log 2)^2/\sqrt{2} - 4u_1(\log u_1 - 1) = 3.796801$ and $M(N_1^*) = 0.071795$. For $j \geq 1$, N_{j+1}^* is the stationary point u^2, $4u = 2A_q + c_0 - \tau_1(2\log u)$, of the first interval (q, q') after N_j^* that contains it. With dA the measure of mass $\Lambda(n)/\sqrt{n}$ at each prime power n,*

$$\beta_{j+1} - \beta_j = \int_{N_j^*}^{N_{j+1}^*} \log\frac{x}{N_j^*}\,\Big(dA(x) - \frac{dx}{\sqrt{x}}\Big) + \frac{L_j}{2}\big(\tau_1(L_{j+1}) - \tau_1(L_j)\big), \tag{16}$$

and the two measures have equal mass on (N_j^*, N_{j+1}^*) *up to* $\frac{1}{2}(\tau_1(L_{j+1}) - \tau_1(L_j))$. *RH holds if and only if the running sum* β_J *never falls below the threshold of* (15).

Proof. The position follows from Proposition 9(3) and Corollary 13. Subtracting $\alpha(N_j^*) = \frac{1}{2}\tau_1(L_j)$ from $\alpha(N_{j+1}^*) = \frac{1}{2}\tau_1(L_{j+1})$ gives $A(N_{j+1}^*) - A(N_j^*) - 2(u_{j+1} - u_j) = \frac{1}{2}(\tau_1(L_{j+1}) - \tau_1(L_j))$, which is the statement about the masses. With $S(u) = 4u(\log u - 1)$ we have $\beta_{j+1} - \beta_j = (B(N_{j+1}^*) - B(N_j^*)) - (S(u_{j+1}) - S(u_j))$. Write $\log n = \log(n/N_j^*) + L_j$ in the first difference and $\int x^{-1/2} \log x \, dx = \int x^{-1/2} \log(x/N_j^*) \, dx + L_j \int x^{-1/2} dx$ in the second. The terms in L_j combine to L_j times the mass difference, which gives (16). The last statement is Corollary 13. □

Status: proved, elementary. Between two minima, the primes and the smooth density $x^{-1/2}dx$ contribute the same mass, and β changes by the difference of their first moments in $\log(x/N_j^*)$. RH says that, starting from $\beta_1 = 3.796801$, these mass-balanced moments never lose more than $\beta_1 + K_0/2 = 0.0961$. Run forward with the prime powers alone, the recurrence finds the same 10 341 minima below 10^7, with β correct to $1.7 \cdot 10^{-11}$; 5 258 steps go up and 5 082 go down. Like Corollary 10, it is exact but contains no information beyond Landau's form of the criterion [In].

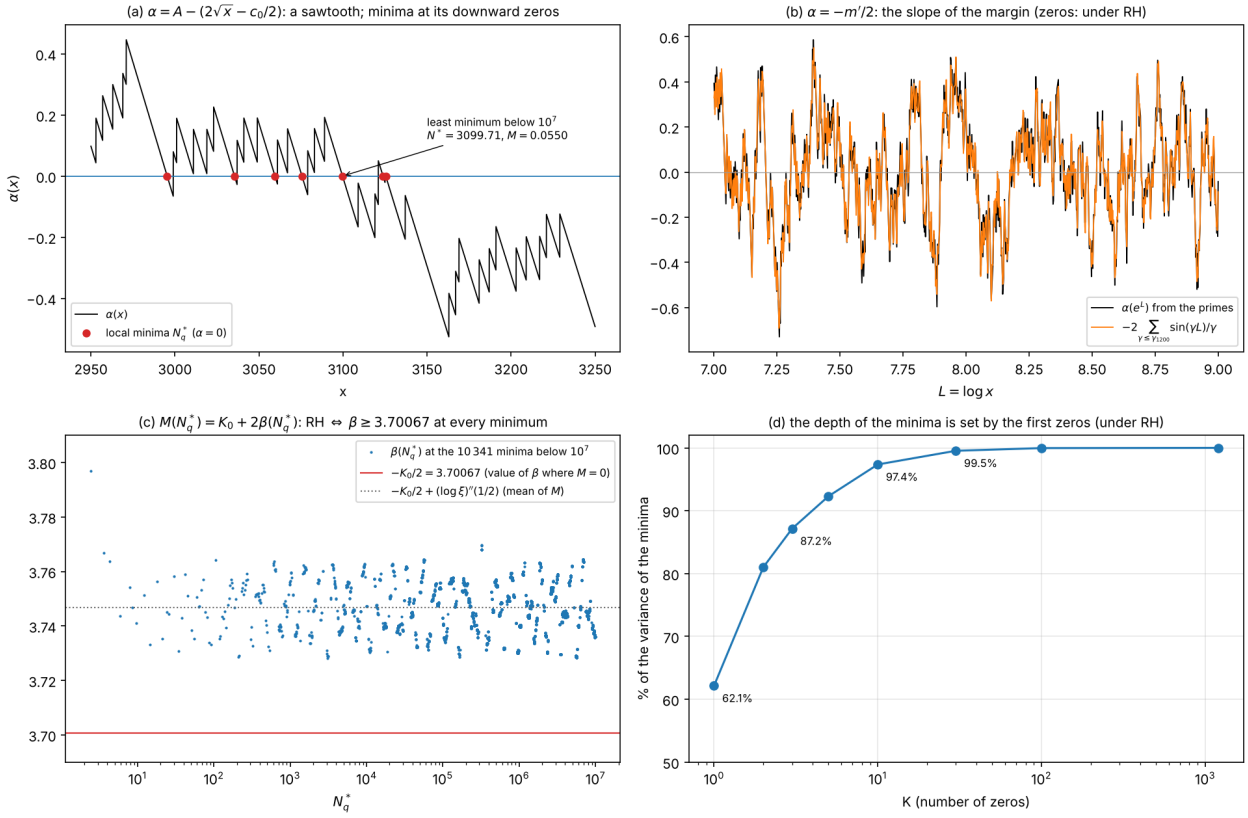


FIGURE 3. The minima read from the slope. (a) The sawtooth $\alpha(x)$ on $[2950, 3250]$; its downward zeros (red) are the local minima N_q^*, among them the least minimum below 10^7. (b) $\alpha(e^L)$ from the primes and $-2\sum_{\gamma \le \gamma_{1200}} \sin(\gamma L)/\gamma$ from the zeros. (c) $\beta(N_q^*)$ at the 10 341 minima below 10^7; the red line is $M = 0$ by (14). (d) Share of the variance of the minima reproduced by the first K zeros.

4. The margin as an energy

By Proposition 12, α is, up to the factor $-\frac{1}{2}$ and the term τ_1, the slope of the margin. Under RH, α is almost periodic in $\log N$, and the margin turns out to be the mean square of its increments. This makes both sides of the sandwich of Corollary 4 visible: $M \geq 0$ because it is a mean of squares, and $M \leq 4(\log\xi)''(\frac{1}{2})$ because the variance of α is $(\log\xi)''(\frac{1}{2})$. The same identity turns RH into a statement about energy: α has finite mean square in $\log x$.

Write $\langle f\rangle = \lim_{U\to\infty}\frac{1}{U}\int_0^U f(u)\,du$ when the limit exists. B^2 is the Besicovitch space of almost periodic functions, the closure of trigonometric polynomials in the seminorm $\langle|f|^2\rangle^{1/2}$. For $f\in B^2$ with Fourier series $\sum_\omega a_\omega e^{i\omega u}$, Parseval's identity $\langle|f|^2\rangle = \sum|a_\omega|^2$ holds [Bes].

Proposition 15 (the margin is the variogram of α)**.** *Assume RH. Then $u\mapsto\alpha(e^u)$ belongs to B^2, has mean 0, and has the Fourier series*

$$\alpha(e^u) \sim \sum_\gamma \frac{i\,m_\gamma}{\gamma}\,e^{i\gamma u} = -2\sum_{\gamma>0}\frac{m_\gamma}{\gamma}\sin(\gamma u),$$

where γ runs over the distinct ordinates of the zeros (both signs in the first sum) and m_γ is the multiplicity. Consequently, for every $L>0$,

$$V(L) := \big\langle|\alpha(e^{u+L})-\alpha(e^u)|^2\big\rangle = 4\sum_{\gamma>0}\frac{m_\gamma^2\,(1-\cos\gamma L)}{\gamma^2} \;\geq\; M(e^L), \tag{17}$$

$$\langle\alpha^2\rangle = 2\sum_{\gamma>0}\frac{m_\gamma^2}{\gamma^2} \;\geq\; (\log\xi)''(\tfrac{1}{2}) = 0.046210\ldots, \tag{18}$$

with equality in both, for all L, if and only if all zeros are simple. In that case $M(e^L)=V(L)$ is a mean of squares, and the sandwich $0\leq M\leq 4(\log\xi)''(\frac{1}{2})$ is the elementary inequality

$$0\leq\langle|f(u+L)-f(u)|^2\rangle\leq 2\langle|f(u+L)|^2\rangle+2\langle|f(u)|^2\rangle = 4\langle|f|^2\rangle.$$

Proof. Let $\phi(u)=e^{-u/2}(\psi(e^u)-e^u)$. Under RH, $\phi\in B^2$ with Fourier series $-\sum_\rho e^{i\gamma_\rho u}/\rho$ [RuSa, ANS]. This is the explicit formula $\psi(x)=x-\sum_\rho x^\rho/\rho-\log 2\pi-\frac{1}{2}\log(1-x^{-2})$, with the series converging in B^2 because $\sum_\rho|\rho|^{-2}<\infty$.

Partial summation with $A(x)=\sum_{n\leq x}\Lambda(n)n^{-1/2}$ gives, for $x=e^u$ not a prime power,

$$A(x)=\frac{\psi(x)}{\sqrt{x}}+\frac{1}{2}\int_1^x\frac{\psi(t)}{t^{3/2}}dt, \qquad \alpha(e^u)=\phi(u)+\frac{1}{2}\int_0^u\phi(v)\,dv+\frac{c_0}{2}-1.$$

Integrating the explicit formula termwise, which is justified as in [In, Ch. IV], gives

$$\frac{1}{2}\int_0^u\phi = -\frac{1}{2}\sum_\rho\frac{e^{i\gamma_\rho u}-1}{i\gamma_\rho\,\rho}+\kappa+o(1)$$

for a constant κ. Here the series converges absolutely and uniformly, since $\sum|\gamma\rho|^{-1}<\infty$. Hence $\alpha(e^u)\in B^2$, and its coefficient at $e^{i\gamma u}$ is

$$-m_\gamma\Big(\frac{1}{\rho}+\frac{1}{2i\gamma\rho}\Big) = -m_\gamma\,\frac{2i\gamma+1}{2i\gamma\rho} = -\frac{m_\gamma}{i\gamma} = \frac{i\,m_\gamma}{\gamma}.$$

The mean is 0. Indeed, for $\Re s>\frac{1}{2}$, the Laplace transform of $\alpha(e^u)$ is

$$\mathcal{L}[\alpha](s) = -\frac{\zeta'}{\zeta}\big(s+\tfrac{1}{2}\big)\frac{1}{s}-\frac{2}{s-\frac{1}{2}}+\frac{c_0}{2s}.$$

By the Abelian theorem [Wi], the mean equals $\lim_{\sigma\to0^+}\sigma\mathcal{L}[\alpha](\sigma) = -\frac{\zeta'}{\zeta}(\frac{1}{2})+\frac{c_0}{2}$, which is 0 by the identity $c_0=2\frac{\zeta'}{\zeta}(\frac{1}{2})$ of Section 2.

Translation and difference preserve B^2. So $\alpha(e^{u+L}) - \alpha(e^u)$ has coefficients $\frac{im_\gamma}{\gamma}(e^{i\gamma L} - 1)$, and Parseval gives (17), because $|e^{i\gamma L} - 1|^2 = 2(1 - \cos\gamma L)$ and each $\gamma > 0$ appears with its negative. Parseval for α itself gives (18).

For the comparisons, (3) counts zeros with multiplicity, so $M(e^L) = 4\sum_{\gamma>0} m_\gamma(1 - \cos\gamma L)/\gamma^2$. Similarly $(\log\xi)''(\frac{1}{2}) = -\sum_\rho(\frac{1}{2} - \rho)^{-2} = 2\sum_{\gamma>0} m_\gamma/\gamma^2$. Since $m^2 \geq m$, with equality iff $m = 1$, both inequalities and the equality cases follow, because $1 - \cos\gamma L$ does not vanish identically in L. □

Proposition 16 (RH as finite energy)**.** *The following are equivalent:*

(i) *the Riemann hypothesis;*
(ii) *α has bounded mean energy in $\log x$: $\frac{1}{U}\int_0^U \alpha(e^u)^2\,du = O(1)$;*
(iii) *$\int_0^U \alpha(e^u)^2\,du = e^{o(U)}$, that is, $\int_1^X \alpha(x)^2\frac{dx}{x} = X^{o(1)}$.*

Proof. (i)⇒(ii) is Proposition 15. In substance it is Cramér's theorem that, under RH, $\int_2^X(\psi(x) - x)^2x^{-2}dx \sim c\log X$ [In].

(ii)⇒(iii) is trivial.

(iii)⇒(i): let $I(U) = \int_0^U \alpha(e^u)^2du = e^{o(U)}$ and $\sigma > 0$. Integration by parts against dI shows that $\int_0^\infty \alpha(e^u)^2e^{-2\sigma' u}du < \infty$ for every $\sigma' > 0$. By Cauchy–Schwarz with $0 < \sigma' < \sigma$,

$$\int_0^\infty |\alpha(e^u)|e^{-\sigma u}du \leq \Big(\int_0^\infty \alpha^2e^{-2\sigma' u}du\Big)^{1/2}\Big(\int_0^\infty e^{-2(\sigma-\sigma')u}du\Big)^{1/2} < \infty.$$

So $\mathcal{L}[\alpha](s)$ converges absolutely in $\Re s > 0$ and is holomorphic there. For $\Re s > \frac{1}{2}$ it equals $-\frac{\zeta'}{\zeta}(s + \frac{1}{2})\frac{1}{s} - \frac{2}{s-\frac{1}{2}} + \frac{c_0}{2s}$, by partial summation from $\sum_n \Lambda(n)n^{-1/2-s} = -\frac{\zeta'}{\zeta}(s + \frac{1}{2})$.

By analytic continuation, $-\frac{\zeta'}{\zeta}(w)/(w - \frac{1}{2}) - 2/(w - 1)$ is holomorphic in $\Re w > \frac{1}{2}$. The pole of $-\frac{\zeta'}{\zeta}$ at $w = 1$ is cancelled, since its residue 1 divided by $w - \frac{1}{2} = \frac{1}{2}$ gives $\frac{2}{w-1}$. So ζ has no zeros in $\Re w > \frac{1}{2}$. □

Remark 17 (what Proposition 16 changes, and what it does not). Proposition 16 replaces the positivity $M \geq 0$ by an energy bound for the slope, $\int_1^X \alpha(x)^2\frac{dx}{x} = X^{o(1)}$. Positivity is then automatic: under (ii), Proposition 15 applies and M is a mean of squares. The bound in (iii) is weak: it allows sub-exponential growth in $\log X$. Yet it is equivalent to RH.

Unconditionally, $\alpha(x) \ll x^{\Theta-1/2+\varepsilon}$, where Θ is the supremum of the real parts of the zeros. This gives only $\int_1^X \alpha^2\frac{dx}{x} \ll X^{2\Theta-1+\varepsilon}$, and the gap between $X^{2\Theta-1}$ and $X^{o(1)}$ is exactly the gap between the known zero-free regions and RH.

By the Paley–Wiener theorem, (ii) says that $\mathcal{L}[\alpha]$ belongs to the Hardy space H^2 of every half-plane $\Re s > \sigma > 0$. Its holomorphy there is the content of RH. Its L^2 size on vertical lines is a mean-square bound for $\frac{\zeta'}{\zeta}$ to the right of the critical line, which is known under RH [MV]. So an unconditional proof of (iii) would have to control $\frac{\zeta'}{\zeta}$ in mean square on every line $\Re w = \frac{1}{2} + \sigma$ and show that it extends holomorphically. No method we know does the first without the second. We do not claim progress on it.

Numerically, with the prime powers below 10^8, $M(e^L)$ and the variogram of α agree to three digits at $L = 0.2, 0.6, \ldots, 3$, and the empirical variance of $\alpha(e^u)$ is 0.04625, against $(\log\xi)''(\frac{1}{2}) = 0.04621$.

4.1. **The energy programme: dyadic blocks and averages in L.** Proposition 16 reduces RH to an energy bound for α. This subsection follows the two natural ways of attacking that bound: in dyadic blocks, using the recurrence of the minima, and by averaging the margin in L. Each route gives exact statements and a clear account of where it stops. Neither route proves RH.

An exact identity: the energy of the slope is a C-transform of the margin. Write $m(L) = M(e^L)$, so that $m' = \tau_1 - 2\alpha$ by Proposition 12.

Theorem 18 (energy identity)**.** *For $1 < X < Y$, neither of them a prime power, and without any hypothesis,*

$$\int_{\log X}^{\log Y} m'(L)^2\,dL = 2\sum_{X<q<Y}\frac{\Lambda(q)}{\sqrt{q}}\,M(q) \;-\; \int_{\log X}^{\log Y} m(L)\,r''(L)\,dL \;+\; \Big[m\,m'\Big]_{\log X}^{\log Y},$$

where q runs over prime powers and $r''(L) = 2e^{L/2} - 2\sum_{k\ge1} e^{-(2k+\frac12)L}$.

Proof. Between consecutive prime powers, $m = r - 2A_qL + 2B_q$ is smooth with $m'' = r''$. At a prime power q, m is continuous, and $m' = r' - 2A$ jumps by $-2\Lambda(q)/\sqrt{q}$. On each gap, $\int m'^2 = [mm'] - \int m\,r''$. Summing over the gaps, the interior boundary terms combine into $\sum_q m(\log q)\big(m'(\log q^-) - m'(\log q^+)\big) = 2\sum_q \frac{\Lambda(q)}{\sqrt{q}}M(q)$. □

The prime measure $\frac{\Lambda(q)}{\sqrt{q}}\delta_{\log q}$ and its smooth density $e^{L/2}dL = \frac12 r''dL + O(e^{-5L/2})$ both look at the margin. Theorem 18 says that the energy of the slope is twice the difference between what the primes see and what the continuum sees, which is the C-transform of M with the weight $\Lambda(n)/\sqrt{n}$.

Corollary 19 (the rate is the window width)**.** *Assume RH and that the zeros are simple. Then*

$$2\sum_{q<Y}\frac{\Lambda(q)}{\sqrt{q}}M(q) - \int_0^{\log Y} m\,r''\,dL = 4\langle\alpha^2\rangle\log Y + o(\log Y) + O\big(1+|\alpha(Y)|\big),$$

with $4\langle\alpha^2\rangle = 4(\log\xi)''(\frac12) = 0.18484\ldots$ In particular, the left side divided by $\log Y$ tends to 0.18484 along every sequence with $\alpha(Y) = o(\log Y)$, and so on a set of Y of logarithmic density 1. In words: per unit of $\log N$, the primes see the margin more than the continuum does by exactly the width of the window of Corollary 4.

Proof. Under RH, $0 \le m \le 0.18484$ (Corollary 4), and $m' = \tau_1 - 2\alpha$ with $\tau_1(L) = O(e^{-5L/2})$. So the boundary term of Theorem 18 is $O(1+|\alpha(Y)|)$, and $\int_0^{\log Y} m'^2 = 4\int_0^{\log Y}\alpha^2 + O(1) = 4\langle\alpha^2\rangle\log Y + o(\log Y)$ by Proposition 15. For simple zeros, (18) gives the value. Finally, $\alpha(e^u)$ has a limiting distribution under RH [ANS], so $|\alpha(Y)| \le K$ on a set of logarithmic density tending to 1 as $K\to\infty$. (The boundary term cannot be dropped altogether: α is unbounded, by Littlewood's theorem.) □

Theorem 20 (a lower bound for the energy, without hypothesis)**.** *Unconditionally,*

$$\limsup_{X\to\infty}\frac{1}{\log X}\int_1^X \alpha(x)^2\frac{dx}{x} \;\ge\; (\log\xi)''(\tfrac12) = 0.046210\ldots,$$

and the limit exists and equals $(\log\xi)''(\frac12)$ if and only if RH holds and all zeros are simple.

Proof. If RH fails, Proposition 16 shows that $\int_1^X \alpha^2\frac{dx}{x}$ is not $X^{o(1)}$, so the lim sup is infinite. If RH holds, Proposition 15 gives the limit $\langle\alpha^2\rangle = 2\sum_{\gamma>0} m_\gamma^2/\gamma^2$. This is $\ge 2\sum_{\gamma>0} m_\gamma/\gamma^2 = (\log\xi)''(\frac12)$, with equality iff every $m_\gamma = 1$. □

So the energy of the slope is a computable test of "RH and simple zeros" together, with an excess $2\sum(m_\gamma^2 - m_\gamma)/\gamma^2$ for multiple zeros. The empirical value is 0.04625 up to 10^8, and 0.0465 from the block energies to 10^{10} below.

Dyadic blocks. Let $E_k = \int_{k\log 2}^{(k+1)\log 2}\alpha(e^u)^2du$ be the energy in the block $[2^k, 2^{k+1})$; by Proposition 16, RH is equivalent to $E_k = 2^{o(k)}$. We computed E_k and $W_k = \int m'^2$ for every block below 10^{10}, exactly between consecutive prime powers, from all 455 062 595 prime powers (Figure 4(a)). The block energies do not grow: the mean of $E_k/\log 2$ over $6 \le k \le 32$ is 0.0465, against 0.0462, and $W_k/\log 2$ has mean 0.1858, against the rate 0.1848 of Corollary 19. Theorem 18 holds block by block to $4\cdot 10^{-11}$ for $k \ge 15$.

Why this does not give a proof: on a block, $\alpha(x) = \alpha(2^k) + D_k(x)$, where D_k depends only on the prime powers of the block. Its energy is the energy of the normalized error $\big(\psi(x) - \psi(2^k) - (x - 2^k)\big)/2^{k/2}$ of the prime number theorem on an interval as long as its position. A bound $2^{o(k)}$ for it is already RH. The recurrence of Proposition 14 moves α exactly from one prime power to the next and reproduces D_k, so it gives nothing more. Finally, no computation can see an off-line zero. Such a zero has $\gamma > T_0$, so its term in α is at most $10^{-12}x^{\beta-1/2}$, which reaches the size 0.2 of α only beyond $x = e^{54}$ even for $\beta = 1$.

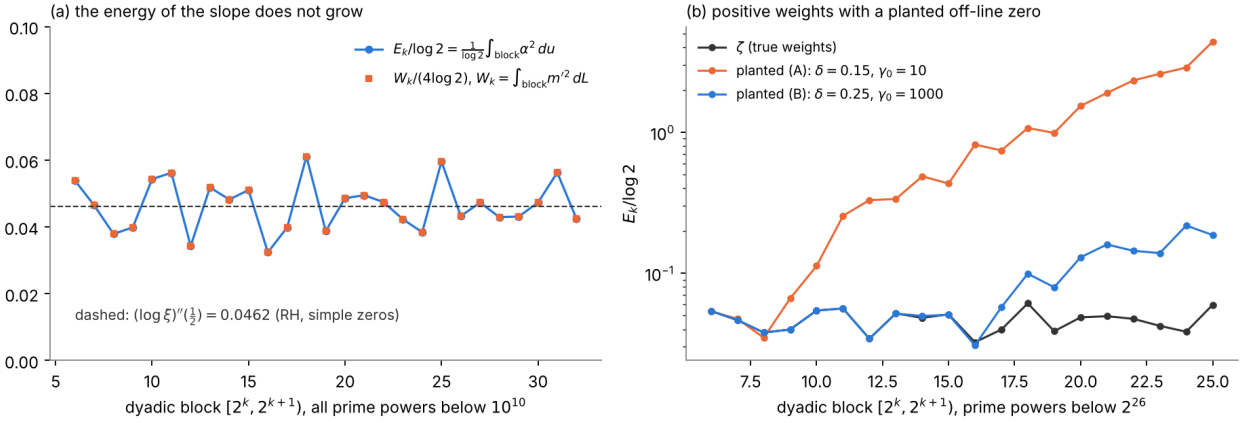


FIGURE 4. (a) Energy per dyadic block of the slope, $E_k/\log 2$, and of the margin's derivative, $W_k/(4\log 2)$, from all prime powers below 10^{10}, against $(\log \xi)''(\frac{1}{2})$. (b) The same block energy, below 2^{26}, for the true weights and for the weights (19) with a planted zero: (A) $\delta = 0.15$, $\gamma_0 = 10$; (B) $\delta = 0.25$, $\gamma_0 = 1000$ (Theorem 23).

Averages in L. Averaging the margin in L is a second-order Riesz mean:

$$M_2(N) := \int_0^{\log N} M(e^L)\, dL = R_2(N) - \sum_{n<N} \frac{\Lambda(n)}{\sqrt{n}} \log^2 \frac{N}{n}, \qquad R_2(N) = \int_0^{\log N} R(e^L)\, dL.$$

Under RH and simple zeros, $M_2(N) = 4\sum_{\gamma>0} \gamma^{-2} \log N - 4\sum_{\gamma>0} \sin(\gamma \log N)/\gamma^3 = 2\langle \alpha^2 \rangle \log N + O(1)$, which is the identity $\langle M \rangle = 2\langle \alpha^2 \rangle$.

Proposition 21 (averaging in L)**.** *The following are equivalent:*

(i) *RH;*
(ii) *for every $\varepsilon > 0$, $M_2(N) \geq -C_\varepsilon N^\varepsilon$;*
(iii) $|M_2(N)| = N^{o(1)}$.

Proof. (i)⇒(ii),(iii): under RH, $0 \leq M \leq 0.18484$ by Corollary 4, so $0 \leq M_2(N) \leq 0.18484 \log N$.

(iii)⇒(i): integrating by parts, $\int_0^\infty M(e^L) e^{-sL} dL = s\int_0^\infty M_2(e^L) e^{-sL} dL$ converges for $\Re s > 0$, so the function $\mathcal{G}$ of the proof of Theorem 3 is holomorphic in $\Re s > 0$. But each zero $\rho = \frac{1}{2} + \delta + it$ with $\delta > 0$ gives $\mathcal{G}$ a pole at $s = \delta + it$.

(ii)⇒(i): apply Landau's theorem [Wi] to $M_2(e^L) + C_\varepsilon e^{\varepsilon L} \geq 0$. Its Laplace transform is $\mathcal{G}(s)/s + C_\varepsilon/(s - \varepsilon)$, which is holomorphic at every real point $s > \varepsilon$, because zeros of ζ are not real and $\zeta(\sigma) \neq 0$ for $\frac{1}{2} < \sigma < 1$. So the integral converges in $\Re s > \varepsilon$, and $\mathcal{G}$ has no pole there. □

Proposition 21 is weaker than Theorem 3 in its hypothesis, since it needs only a one-sided bound on an average and allows N^ε. It is equivalent all the same, and it is the same statement as Proposition 16, because $M_2 = 2\langle \alpha^2 \rangle \log N + O(1)$ under RH. The least eigenvalue of a window (Section 6) is an infimum of such energies over test functions, so it enters through the same door.

Remark 22 (where both routes stop). Every statement in this subsection that is equivalent to RH is equivalent through a Landau-type argument. The analytic content is always the same: the Laplace transform of m, of α, or of the kernel's remainder is $-\frac{\zeta'}{\zeta}$ shifted by $\frac{1}{2}$, divided by a polynomial in s, plus explicit terms. Each bound in the programme (energy, average, one-sided bound) is equivalent to the holomorphy of $\frac{\zeta'}{\zeta}$ in $\Re w > \frac{1}{2}$.

The two routes make this concrete. The block computation shows that "local" in $\log x$ means a prime number theorem on intervals as long as their position. Averaging shows that averaging changes the order of the Riesz mean but not the abscissa. Theorem 18 is the one exact link between the energy and the primes. To be useful it would need an independent lower bound for the continuum term or an upper bound for the prime term, and each of these is again a statement about $\frac{\zeta'}{\zeta}$ to the right of the line.

We record the programme as a reformulation with exact identities and verified numerics.

5. What the arguments cannot see, and what they must use

5.1. **What the energy programme cannot see: a planted zero.** Remark 22 says that every route of Section 4.1 ends at the holomorphy of $\frac{\zeta'}{\zeta}$. This subsection makes that statement precise in the opposite direction. We build a second system of weights on the prime powers. It satisfies every structural ingredient the programme uses: positivity, the energy identity, the prime number theorem with the classical error term, the local recurrence, and Chebyshev's identity up to $O(x^{1/2+\delta})$. Yet its slope carries an oscillation of an off-line zero. So any proof of the energy bound must use the exact values of Λ at a precision that these ingredients do not give.

The construction. Fix $\delta \in (0, \frac{1}{2})$, $\gamma_0 > 0$, a switch-on point x_0 and a smooth function χ with $\chi = 0$ on $(-\infty, 0]$, $\chi = 1$ on $[1, \infty)$. Put

$$f(x) = \frac{2}{\gamma_0}\, x^{\delta} \sin(\gamma_0 \log x)\, \chi\Big(\log \frac{x}{x_0}\Big), \qquad \phi = f + \lambda h,$$

where h is a fixed smooth bump with compact support in $(1, x_0)$ and λ is a real constant chosen below. Define new weights on the prime powers,

$$w(q) = \frac{\Lambda(q)}{\sqrt{q}}\,\big(1 + \varepsilon(q)\big), \qquad \varepsilon(q) = \sqrt{q}\,\phi'(q), \tag{19}$$

and let A_w, $\alpha_w = A_w - 2\sqrt{x} + \frac{c_0}{2}$, C_w, $M_w = R - C_w$ and m_w be built from w exactly as A, α, C, M and m are built from $\Lambda(n)n^{-1/2}$.

Theorem 23 (a planted zero)**.** *With the notation above, for x_0 large enough:*

(i) positivity, and the weights are asymptotically the true ones: $|\varepsilon(q)| \le C_\phi\, q^{\delta - \frac{1}{2}} < 1$, *so* $w(q) > 0$ *and* $w(q)\sqrt{q}/\Lambda(q) \to 1$*;*

(ii) the energy identity: *Theorem 18 holds verbatim for m_w, with $\Lambda(q)q^{-1/2}$ replaced by $w(q)$;*

(iii) the planted oscillation: *unconditionally,* $\alpha_w(x) = \alpha(x) + f(x) + o(x^{\delta})$*;*

(iv) the prime number theorem is unchanged at the classical scale: $\sum_{q \le x} \Lambda(q)(1 + \varepsilon(q)) = x + O\big(x\, e^{-c\sqrt{\log x}}\big)$*;*

(v) Chebyshev's identity up to $x^{1/2+\delta}$: *for a suitable* λ, $\sum_{q \le x} \Lambda(q)(1 + \varepsilon(q))\lfloor x/q \rfloor = \log \lfloor x \rfloor ! + O(x^{1/2+\delta})$*;*

(vi) one of the two systems violates the energy bound: *if* $\gamma_0 \ge 7\delta$ *(in particular if* $\gamma_0 \ge 3.5$*), then for all large X,*

$$\max\Big(\int_1^X \alpha^2 \frac{dx}{x},\ \int_1^X \alpha_w^2 \frac{dx}{x} \Big) \ \ge\ \frac{X^{2\delta}}{7\,\delta\,\gamma_0^2}.$$

Proof. (i) $|f'(x)| \le \frac{2}{\gamma_0}(\delta+\gamma_0+\|\chi'\|_\infty)x^{\delta-1}$ and h' is bounded with compact support, so $|\varepsilon(q)| \le C_\phi q^{\delta-1/2}$. On the support of h the bound is $|\lambda|\sqrt{q}\|h'\|_\infty$, which is <1 for the small λ needed in (v). For $q \ge x_0$, $C_\phi x_0^{\delta-1/2} < 1$ once x_0 is large.

(ii) The proof of Theorem 18 uses only that $m_w = r - 2A_wL + 2B_w$ is smooth with $m_w'' = r''$ between prime powers and that m_w' jumps by $-2w(q)$ at q. Nothing else about the weights enters.

(iii) $A_w(x) - A(x) = \sum_{q<x}\Lambda(q)\phi'(q) = \int_{2^-}^x \phi'(t)\,d\psi(t)$. Write $d\psi = dt + d(\psi(t)-t)$ and integrate the second part by parts:

$$A_w(x) - A(x) = \phi(x) + \phi'(x)\big(\psi(x)-x\big) - \int_2^x \big(\psi(t)-t\big)\phi''(t)\,dt.$$

Here $\phi(2)=0$, and for $x \ge x_0$, $\phi = f$. With $\psi(t)-t \ll te^{-c\sqrt{\log t}}$ and $\phi''(t) \ll \gamma_0 t^{\delta-2}$, the last two terms are $\ll x^\delta e^{-c'\sqrt{\log x}} = o(x^\delta)$.

(iv) The same computation without the factor $q^{-1/2}$ gives $\sum_{q\le x}\Lambda(q)\varepsilon(q) = \int_2^x \sqrt{t}\,\phi'(t)\,d\psi(t) \ll x^{1/2+\delta}$, which is smaller than the classical error term.

(v) Write $\lfloor x/q\rfloor = x/q - \{x/q\}$. The defect is $D(x) = x\sum_{q\le x}\Lambda(q)\varepsilon(q)/q - \sum_{q\le x}\Lambda(q)\varepsilon(q)\{x/q\}$. The second sum is $\ll \sum_{q\le x}\Lambda(q)q^{\delta-1/2} \ll x^{1/2+\delta}$. The series $S = \sum_q \Lambda(q)\varepsilon(q)/q$ converges absolutely, and its tail beyond x is $\ll x^{\delta-1/2}$. The bump changes S by $\lambda\sum_q \Lambda(q)h'(q)q^{-1/2} \ne 0$ for a suitable h, so one value of λ gives $S=0$, and then $D(x) \ll x^{1/2+\delta}$.

(vi) By (iii), $f = \alpha_w - \alpha + o(x^\delta)$, so $\int_1^X f^2\frac{dx}{x} \le 3\int\alpha_w^2\frac{dx}{x} + 3\int\alpha^2\frac{dx}{x} + o(X^{2\delta})$. Now $\frac{4}{\gamma_0^2}\int^{\log X} e^{2\delta u}\sin^2(\gamma_0 u)\,du = \frac{X^{2\delta}}{\delta\gamma_0^2}\Big[1 - \delta\,\Re\frac{e^{2i\gamma_0\log X}}{\delta+i\gamma_0}\Big] + O(1)$, and the bracket oscillates between $1 \pm \delta/\sqrt{\delta^2+\gamma_0^2}$ (it does not tend to 1); for $\gamma_0 \ge 7\delta$ it is at least $1 - \frac{1}{\sqrt{50}} > \frac{6}{7}\cdot 1.001$, so $\int_1^X f^2\frac{dx}{x} \ge \frac{6}{7}\cdot 1.001\,\frac{X^{2\delta}}{\delta\gamma_0^2}$ for large X, and the maximum is at least one sixth of this minus $o(X^{2\delta})$. For general $\gamma_0 > 0$ the same argument gives the constant $(1-\delta/\sqrt{\delta^2+\gamma_0^2})/6.01$ in place of $\frac{1}{7}$. □

What the theorem says. Every structural fact the programme uses holds for both systems: positivity of the weights, the energy identity, the recurrence of the minima (Proposition 14 is also pure bookkeeping), the prime number theorem with its classical error, and the arithmetic identity $\Lambda * 1 = \log$ up to $O(x^{1/2+\delta})$. Yet at most one of the two systems satisfies the energy bound of Proposition 16. So a proof of RH along these lines must use the exact values of Λ at relative precision $x^{\delta-1/2}$, for every $\delta > 0$. That is the strength of the statement $\psi(x) = x + O(x^{1/2+\delta})$ itself, the zero-free strip of width $\frac{1}{2} - \delta$.

Beurling systems give a stronger form of the same obstruction. Their weights come from a genuine Euler product, the integers satisfy $N(x) = cx + O(x^\theta)$ with $\theta < 1$, and yet the error term of the prime number theorem can be as large as $x\exp(-c\sqrt{\log x})$ [DMV]. The weights (19) do not come from an Euler product. Their point is that they stay within $q^{\delta-1/2}$ of the true ones and violate $\Lambda * 1 = \log$ only at the scale $x^{1/2+\delta}$. The information that rules them out is the exact identity $\sum_{n\le x}\Lambda(n)\lfloor x/n\rfloor = \log\lfloor x\rfloor!$, equivalently $\zeta(s) = \sum n^{-s}$ with its analytic continuation. No step of Sections 3 and 4 uses it beyond the prime number theorem.

The numbers. Figure 4(b) plants two zeros into the true prime powers below 2^{26}. The bump $h = \sin^2$ in $\log x$ on $[30, 300]$ fixes $S = 0$ on the computed range. In (A), $\delta = 0.15$, $\gamma_0 = 10$, $x_0 = 10^3$: $\max|\varepsilon(q)| = 0.112$, the Chebyshev defect satisfies $|D(x)| \le 0.31\,x^{0.65}$ for $10^6 \le x \le 6\cdot 10^7$, and the block energy grows like $2^{2\delta k}$. In (B), $\delta = 0.25$, $\gamma_0 = 1000$, $x_0 = 10^5$: $\max|\varepsilon(q)| = 0.077$, and the block energies agree with those of ζ within 0.002 up to $k = 16$. Case (B) is the realistic one; a true off-line zero has $\gamma_0 > T_0$ and is invisible below e^{54}.

Remark 24 (status). Theorem 23 is proved and unconditional. It does not say that RH is false, or that the energy programme is useless. It says exactly which input is missing: an argument that uses

$\Lambda = \mu * \log$ exactly, not only through the prime number theorem. The same test applies to any other argument in this paper: an argument that remains valid for the weights (19) cannot prove RH.

5.2. **Using $\Lambda = \mu * \log$ exactly.** Theorem 23 says that a proof of the energy bound must use the identity $\Lambda = \mu * \log$ exactly, not only through the prime number theorem. This subsection carries that identity through the slope α and asks what it gives.

The identity. For every $n \geq 1$,

$$\Lambda(n) = \sum_{d|n} \mu(d) \log \frac{n}{d}, \qquad \text{equivalently} \qquad \sum_{d|n} \Lambda(d) = \log n, \qquad -\frac{\zeta'}{\zeta}(s) = \frac{1}{\zeta(s)} \cdot \big(-\zeta'(s)\big). \tag{20}$$

The divisor sum is a massive cancellation (Figure 5(a)). For n with two distinct prime factors, the positive and negative terms are each as large as $\log n$, and they cancel exactly to 0. Summed up to x, (20) is Chebyshev's identity $\sum_{m \leq x} \psi(x/m) = \log \lfloor x \rfloor !$. Its Möbius inverse $\psi(x) = \sum_{d \leq x} \mu(d) \log \lfloor x/d \rfloor !$ adds terms of size $x \log x$ that cancel to x (Figure 5(b): at $x = 10^4$ the first term is 82 109 and the sum is 10 013.4). The cleanest exact form for what follows is

$$\psi(x) + \log x = \sum_{d \leq x} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor \log \frac{x}{d}, \tag{21}$$

which follows from (20) by summing $\sum_{d|n} \mu(d) \log(x/d) = [n = 1] \log x + \Lambda(n)$ over $n \leq x$. It holds to 10^{-11} in floating point at $x = 10^4$, 12 345.6 and 99 999.5.

The slope through Möbius. Write, for $x > 1$ not an integer,

$$m_1(x) = \sum_{d<x} \frac{\mu(d)}{d}, \quad m_2(x) = \sum_{d<x} \frac{\mu(d) \log d}{d}, \quad m_{1/2}(x) = \sum_{d<x} \frac{\mu(d)}{\sqrt{d}}, \quad S(y) = \sum_{m<y} \frac{\log m}{\sqrt{m}},$$

and $S_0(y) = 2\sqrt{y} \log y - 4\sqrt{y} - \zeta'(\tfrac{1}{2})$. Euler–Maclaurin gives $S(y) - S_0(y) \to 0$; the constant is $-\zeta'(\tfrac{1}{2})$ because $\sum_{m \leq y} m^{-s} \log m = \frac{y^{1-s} \log y}{1-s} - \frac{y^{1-s}}{(1-s)^2} - \zeta'(s) + o(1)$ for $0 < s < 1$.

Proposition 25 (the slope through Möbius)**.** *For every $x > 1$ that is not an integer, exactly,*

$$\begin{aligned} \alpha(x) &= \mathfrak{T}_1(x) + \mathfrak{T}_2(x) + \mathfrak{T}_3(x), \\ \mathfrak{T}_1 &= 2\sqrt{x}\,\big(m_1 \log x - m_2 - 2m_1 - 1\big), \qquad \mathfrak{T}_2 = -\zeta'(\tfrac{1}{2})\Big(m_{1/2} - \frac{1}{\zeta(\frac{1}{2})}\Big), \\ \mathfrak{T}_3 &= \sum_{d<x} \frac{\mu(d)}{\sqrt{d}}\,(S - S_0)\Big(\frac{x}{d}\Big). \end{aligned}$$

Moreover $(S - S_0)(y) = -B_1(\{y\}) \frac{\log y}{\sqrt{y}} + O\big(\frac{\log ey}{y^{3/2}}\big)$ with $B_1(t) = t - \frac{1}{2}$, so that

$$\mathfrak{T}_3(x) = -\frac{1}{\sqrt{x}} \sum_{d<x} \mu(d)\, B_1\Big(\Big\{\frac{x}{d}\Big\}\Big) \log \frac{x}{d} + \sum_{d<x} \frac{\mu(d)}{\sqrt{d}}\, O\Big(\frac{\log(ex/d)}{(x/d)^{3/2}}\Big). \tag{22}$$

Proof. By (20), $A(x) = \sum_{dm<x} \mu(d) \log m\, (dm)^{-1/2} = \sum_{d<x} \mu(d) d^{-1/2} S(x/d)$. Write $S = S_0 + (S - S_0)$. The S_0 part is

$$\sum_{d<x} \frac{\mu(d)}{\sqrt{d}} \Big[2\sqrt{\tfrac{x}{d}}\big(\log x - \log d\big) - 4\sqrt{\tfrac{x}{d}} - \zeta'(\tfrac{1}{2})\Big] = 2\sqrt{x}\,(m_1 \log x - m_2 - 2m_1) - \zeta'(\tfrac{1}{2})\, m_{1/2}.$$

Subtracting $2\sqrt{x}$ and adding $\frac{c_0}{2} = \frac{\zeta'}{\zeta}(\frac{1}{2})$ gives $\alpha = \mathfrak{T}_1 + \mathfrak{T}_2 + \mathfrak{T}_3$. For (22), apply Euler–Maclaurin to $f(t) = t^{-1/2} \log t$ with the first Bernoulli function: $S(y) = \int_1^y f + C - B_1(\{y\}) f(y) - \int_y^\infty B_1(\{t\}) f'(t)\, dt$. The last integral is $O(|f'(y)|)$ after one more integration by parts, and C is fixed by $S - S_0 \to 0$. $\square$

Status: proved; the identity is checked to $4.4 \cdot 10^{-9}$ at 400 points in $[10^3, 10^7]$ (Figure 5(c)). The three pieces behave very differently. $\mathfrak{T}_1$ and $\mathfrak{T}_2$ are each large: their root mean squares on $[10^3, 10^7]$ are 0.694 and 0.680, against 0.218 for α. But they are almost exact negatives of each other: their correlation is -0.9999999, and $\mathfrak{T}_1 + \mathfrak{T}_2$ has root mean square 0.0140. So α is essentially $\mathfrak{T}_3$, and $\mathfrak{T}_3$ is essentially the *Möbius-weighted sawtooth* in (22): the difference has root mean square 0.012. This difference in turn nearly cancels $\mathfrak{T}_1 + \mathfrak{T}_2$: α minus the sawtooth has root mean square 0.0020, against 0.249 for α at the same 50 points. To within 1% in root mean square on the range tested,

$$\alpha(x) \approx -\frac{1}{\sqrt{x}} \sum_{d<x} \mu(d) \Big(\Big\{ \frac{x}{d} \Big\} - \frac{1}{2} \Big) \log \frac{x}{d}.$$

The slope of the margin is the fractional parts of x/d, weighted by $\mu(d) \log(x/d)$. This is the one place where the additive structure of $\mathbb{N}$, through $\lfloor \cdot \rfloor$, enters the margin. The 1% agreement is not a coincidence. By Remark 28, the sawtooth sum is $-(\psi(x) - x)$ up to Mertens-type sums. Under RH, $(\psi(x) - x)/\sqrt{x}$ has the coefficients $-1/\rho = \frac{i}{\gamma} + O(\gamma^{-2})$, the same as α up to $O(\gamma^{-2})$.

The spectral content of the pieces. Under RH the coefficient of $e^{i\gamma u}$ in $\alpha(e^u)$ is $\frac{im_\gamma}{\gamma}$ (Proposition 15). It is rigid: it depends only on γ and the multiplicity. The pieces of Proposition 25 are not rigid.

Proposition 26 (the pieces carry $1/\zeta'(\rho)$)**.** *Assume RH, that the zeros are simple, and that the Perron expansions of m_1, m_2 and $m_{1/2}$ over the zeros converge in B^2 (as they do for $M(x)$ under Ng's hypothesis below). Then the coefficients of $e^{i\gamma u}$, γ the ordinate of $\rho = \frac{1}{2} + i\gamma$, are*

$$c_1(\gamma) = \frac{4(1 - i\gamma)}{(i\gamma - \frac{1}{2})^2 \, \zeta'(\rho)}, \qquad c_2(\gamma) = -\frac{\zeta'(\frac{1}{2})}{i\gamma \, \zeta'(\rho)}, \qquad c_3(\gamma) = \frac{i}{\gamma} - c_1(\gamma) - c_2(\gamma),$$

for $\mathfrak{T}_1$, $\mathfrak{T}_2$ and $\mathfrak{T}_3$. In particular, as $\gamma \to \infty$,

$$c_1(\gamma) + c_2(\gamma) = \frac{\big(4 + \zeta'(\frac{1}{2})\big)\, i}{\gamma \, \zeta'(\rho)} \big(1 + O(\gamma^{-1})\big), \qquad 4 + \zeta'(\tfrac{1}{2}) = 0.0774.$$

Proof. Perron's formula for $m_{1/2}$ is $\frac{1}{2\pi i} \int \frac{x^w}{w \, \zeta(\frac{1}{2}+w)} dw$. The zero ρ gives the residue $\frac{x^{i\gamma}}{i\gamma \zeta'(\rho)}$, so $\mathfrak{T}_2$ has c_2. For m_1 the residue is $\frac{x^{w_0}}{w_0 \zeta'(\rho)}$, $w_0 = \rho - 1$. For m_2, the Dirichlet series is $-(1/\zeta)'(1+w)$, which has a double pole with principal part $\frac{1}{\zeta'(\rho)(w - w_0)^2}$; the residue of $\frac{x^w}{w} \cdot \frac{1}{\zeta'(\rho)(w-w_0)^2}$ is $\frac{x^{w_0}}{\zeta'(\rho)} \big(\frac{\log x}{w_0} - \frac{1}{w_0^2} \big)$. Hence $m_1 \log x - m_2 - 2m_1$ contributes $\frac{x^{w_0}}{\zeta'(\rho)} \big(\frac{1}{w_0^2} - \frac{2}{w_0} \big)$. Multiplying by $2\sqrt{x}$ gives c_1, since $1 - 2w_0 = 2 - 2i\gamma$. The sum of the three coefficients is the coefficient of α, which gives c_3. The asymptotic form follows from $c_1 = \frac{4i}{\gamma \zeta'(\rho)}(1 + O(\gamma^{-1}))$. □

Status: the residues are a computation; the convergence in B^2 is an assumption of the same type as Ng's. A least-squares fit of the first fourteen frequencies to the computed pieces on $\log x \in [\log 10^3, \log 10^7]$ gives amplitudes $2|c|$ that agree with the predictions to the accuracy of the fit. At γ_1 the predicted and fitted values are 0.6998 and 0.7058 for $\mathfrak{T}_2$, 0.7145 and 0.7206 for $\mathfrak{T}_1$, and 0.0147 and 0.0148 for the residual $\mathfrak{T}_1 + \mathfrak{T}_2$, which is 50 times smaller than either piece.

What the factorization does to the energy. The identity (20) factors $-\frac{\zeta'}{\zeta}$ into $\frac{1}{\zeta}$ times $-\zeta'$. The second factor is harmless: it is holomorphic in $\Re s > \frac{1}{2}$ except at $s = 1$, and of polynomial growth there. The first factor carries all the zeros, with residues $\frac{1}{\zeta'(\rho)}$. The rigidity of α is the identity $\zeta'(\rho) \cdot \frac{1}{\zeta'(\rho)} = 1$, and it reappears only when the three pieces are added. The energy shows the difference.

Proposition 27 (the Möbius energy)**.** *Let $M(x) = \sum_{n \le x} \mu(n)$.*

(i) *Unconditionally: if $\int_1^X M(x)^2 x^{-2} dx = X^{o(1)}$, then RH holds.*

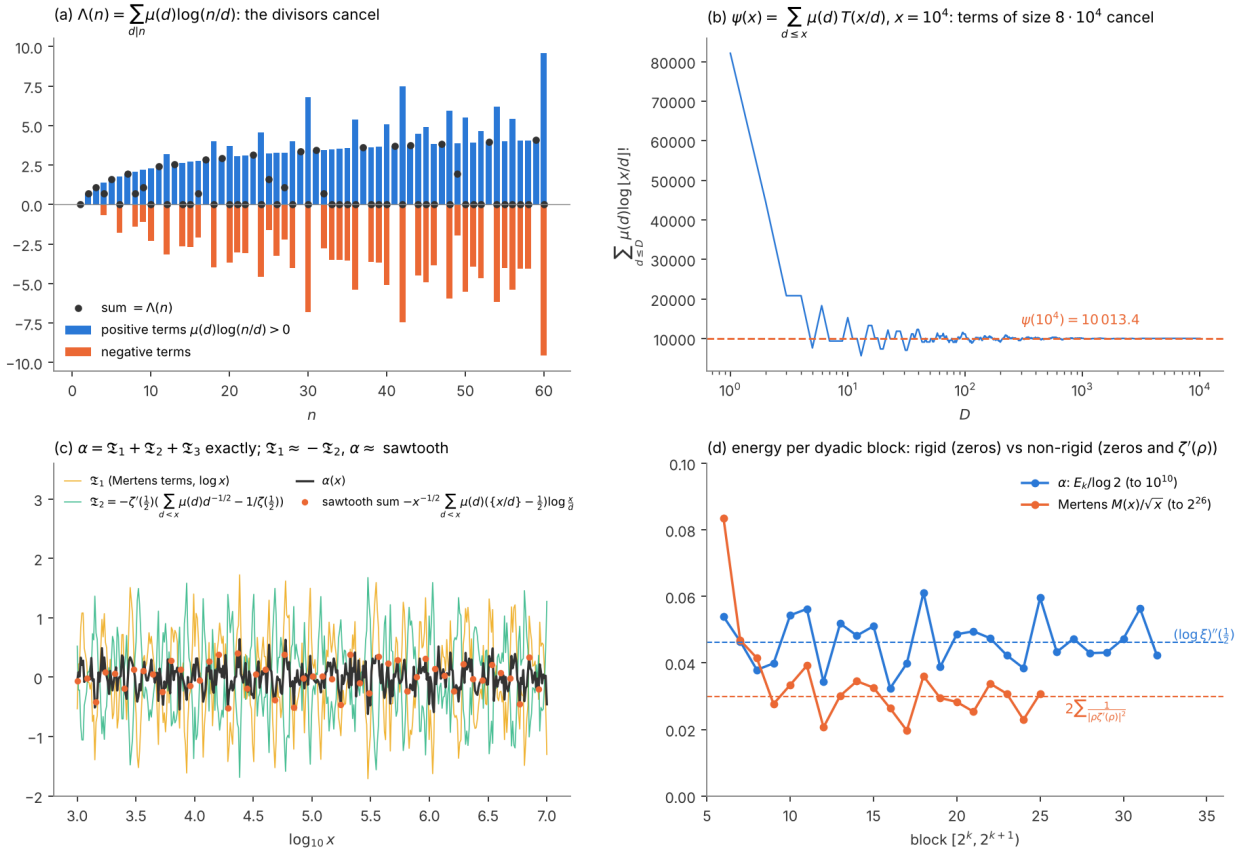


FIGURE 5. $\Lambda = \mu * \log$. (a) The divisor sum (20) for $n \le 60$: positive and negative terms, and their sum $\Lambda(n)$. (b) Partial sums of $\psi(x) = \sum_{d \le x} \mu(d) \log\lfloor x/d \rfloor!$ at $x = 10^4$. (c) The exact decomposition of Proposition 25: $\mathfrak{T}_1$ and $\mathfrak{T}_2$ are mirror images, and α follows the Möbius-weighted sawtooth (22) (dots). (d) Energy per dyadic block of α (to 10^{10}) and of the Mertens function $M(x)/\sqrt{x}$ (to 2^{26}), with their predicted means $(\log \xi)''(\frac{1}{2})$ and $2\sum_{\gamma>0} |\rho\zeta'(\rho)|^{-2}$.

(ii) *(Ng [Ng]) If RH holds and* $J_{-1}(T) = \sum_{0<\gamma\le T} |\zeta'(\rho)|^{-2} \ll T$, *then* $e^{-u/2}M(e^u)$ *has a limiting distribution with mean* 0 *and mean square*

$$\big\langle (M(x)/\sqrt{x})^2 \big\rangle = 2 \sum_{\gamma>0} \frac{1}{|\rho\, \zeta'(\rho)|^2}.$$

(iii) *Numerically, the sum over the first* 96 *zeros is* 0.0282. *The Gonek–Hejhal conjecture* $J_{-1}(T) = \sum_{0<\gamma\le T} |\zeta'(\rho)|^{-2} \sim \frac{3T}{\pi^3}$ [Go] *estimates the tail beyond* $\gamma = 229.3$ *at* $2 \cdot \frac{3}{\pi^3}/229.3 = 0.0008$, *for a total of* 0.0291. *The block energies of* $M(x)/\sqrt{x}$ *for* $k = 10, \dots, 25$ *have mean* 0.0296 *(Figure 5(d)).*

Proof. (i) For $\Re s > 1$, $\int_1^\infty M(x) x^{-s-1} dx = \frac{1}{s\zeta(s)}$. In the variable $u = \log x$, $\int_0^\infty M(e^u) e^{-u/2} e^{-wu} du = \frac{1}{(w+\frac{1}{2})\zeta(w+\frac{1}{2})}$. If the energy is $e^{o(U)}$, the Cauchy–Schwarz argument of Proposition 16 makes this Laplace integral converge absolutely in $\Re w > 0$. So $\frac{1}{\zeta}$ is holomorphic in $\Re s > \frac{1}{2}$. (ii) is proved in [Ng], and (iii) is a computation. □

The comparison is the point of this subsection. Under RH, the energy of α is $2\sum m_\gamma^2/\gamma^2$. It is finite, equal to $(\log \xi)''(\frac{1}{2})$ for simple zeros, and needs nothing more. The Möbius energy needs

control of $\zeta'(\rho)$ from below, which is not known even under RH. Through (20), the energy problem for α becomes an energy problem for M together with a harmless smooth factor. That problem is no easier: it is strictly less rigid.

Three levels of exactness. Which arithmetic facts exclude the systems of Theorem 23?

(0) *The planted weights* (19) have an Euler product only formally. Defining $\zeta_w(s) = \exp\big(\sum_n \Lambda(n)(1+\varepsilon(n))n^{-s}/\log n\big)$, the local factor at p is $\exp\big(\sum_k (1+\varepsilon(p^k))p^{-ks}/k\big)$. So $\mu_w(p) = -(1+\varepsilon(p))$ and $\mu_w(p^2) = \frac{1}{2}\big((1+\varepsilon(p))^2 - (1+\varepsilon(p^2))\big)$. For the system (A) of Figure 4(b), $\max|\mu_w(p)| = 1.112$ on $10^3 < p < 2^{13}$, where 55% of the primes have $|\mu_w(p)| > 1$, and $\max|\mu_w(p^2)| = 0.113$. The local facts $\mu(p) = -1$ and $\mu(p^2) = 0$ already exclude them.

(1) *Beurling systems* have $\mu_B(p) = -1$ and $\mu_B(p^k) = 0$ for $k \geq 2$, because they come from a genuine Euler product over generalized primes. Diamond, Montgomery and Vorhauer [DMV] construct such systems with integer counting function $N_B(x) = \kappa x + O(x^\theta)$, $\frac{1}{2} < \theta < 1$, whose primes satisfy $\pi_B(x) = \mathrm{li}(x) + \Omega\big(x\exp(-c\sqrt{\log x})\big)$. So the multiplicative structure, even with integers known to precision x^θ for any $\theta > \frac{1}{2}$, does not give the energy bound. In Proposition 25 for such a system, the sawtooth $B_1(\{x/d\})$ is replaced by the error of N_B at x/d, which can be as large as $(x/d)^\theta$ instead of $O(1)$.

(2) *The integers* $\mathbb{N}$: $N(x) = \lfloor x \rfloor$, with error $O(1)$. This is the only information left, and Proposition 25 shows where it enters: the sawtooth in (22) and the two constants $\zeta(\frac{1}{2})$, $\zeta'(\frac{1}{2})$.

Remark 28 (status: what an exact argument must do). By Proposition 25 and Theorem 23, a proof along the lines of this paper must show that the Möbius-weighted sawtooth

$$\mathcal{S}(x) = \sum_{d \leq x} \mu(d)\, B_1\Big(\Big\{\frac{x}{d}\Big\}\Big) \log\frac{x}{d}$$

is $x^{1/2+o(1)}$ in mean square on $[X, 2X]$, and it must use the cancellation between $\mu(d)$ and the fractional parts of x/d. The identity (21) shows that nothing is gained for free. Inserting $\lfloor y \rfloor = y - \frac{1}{2} - B_1(\{y\})$ gives, exactly,

$$\mathcal{S}(x) = x\sum_{d \leq x} \frac{\mu(d)}{d}\log\frac{x}{d} - \frac{1}{2}\sum_{d\leq x}\mu(d)\log\frac{x}{d} - \psi(x) - \log x,$$

so $\mathcal{S}$ is $-\big(\psi(x) - x\big)$ up to Mertens-type sums that are $O(x^{1/2+\varepsilon})$ precisely when RH holds. The identity $\Lambda = \mu * \log$ is reversible, and rewriting ψ through it changes the form of the problem but not its content.

The form is still informative. $\mathcal{S}$ is a type I sum: Möbius against the explicit, bounded, additive function $B_1(\{x/d\})$. Expanding B_1 in its Fourier series turns $\mathcal{S}$ into exponential sums $\sum_d \mu(d)e(hx/d)$. The best unconditional bounds for $\psi(x) - x$, and so for $\mathcal{S}$, save only a factor $\exp\big(-c(\log x)^{3/5}(\log\log x)^{-1/5}\big)$ over the trivial bound. That is the Vinogradov–Korobov zero-free region, seen from the margin. A saving of $x^{-1/2+\varepsilon}$ would be RH.

6. Windows and Weil positivity: a summary

The margin is Weil's form Q for one test function. Parts 13, 14 and 21 of this series [CE13, CE14, CE21] study Q on whole spaces of test functions. We summarize the results, with their status, and refer to those papers for proofs and data.

Dirichlet polynomials [CE13]. Because $\Lambda \geq 0$, the prime part of Q on Gaussian-smoothed Dirichlet polynomials is unbounded below. So positivity is a cancellation between the archimedean place and the coherent sum of all primes, not a diagonal dominance (proved, elementary). With ball arithmetic, Q is positive definite on these polynomials for every length $N \leq 179$ (computer-assisted, for smoothing width 0.005). No translation-invariant square $\|Xk\|^2$ with X a Fourier multiplier fits below Q on the whole line; in particular, no Euler-product or mollifier multiplier does (proved).

The margin as $W - E$ [CE14]. Following Weil's proof for curves, write $Q = W - E$ with the pole $W = 2\widehat{g}(\frac{i}{2})\widehat{g}(-\frac{i}{2})$. Then W is a hyperbolic plane, like the two fibres of $C \times C$, and $Q \geq 0$ on a space implies that E has at most one positive direction there (proved; not conversely). The archimedean place alone has $(t_0/\pi + o(1)) \log N$ positive directions, $t_0 = 6.2898$ (proved, of Szegő type). Numerically, the primes remove them only all together and only with their powers.

Windows [CE21]. In a window of length $L = \log N$ the form is one symbol. A positive completion of that symbol gives an exact sum of squares, and a coercive form has the completion $\pi/|E_L|^2$ with E_L the de Branges function of the window (proved, classical in substance). For even k of bounded variation supported in $[-\frac{L}{2}, \frac{L}{2}]$, with $K = \widehat{k}$, the explicit formula applied to $k * \tilde{k}$ gives

$$Q(k) = \sum_{\rho} K(t_\rho)^2,$$

so under RH the least eigenvalue of a window is the minimum of $2\sum_{\gamma>0} K(\gamma)^2/\|k\|^2$ (proved). Connes observed that the minimizer reproduces the zeros [Co]. In high precision, it does so to 10^{-2} up to 1.3 to 1.9 times $T^* = 2\pi N$, the height where the density $\frac{1}{2\pi}\log\frac{t}{2\pi}$ of the zeros equals the Nyquist density $L/2\pi$ of the window (numerical; Figure 6). For the Davenport–Heilbronn function, the agreement stops at the corresponding scale $2\pi N/5$. The zeros below T^* carry a fraction between 10^{-25} and 10^{-8} of the least eigenvalue.

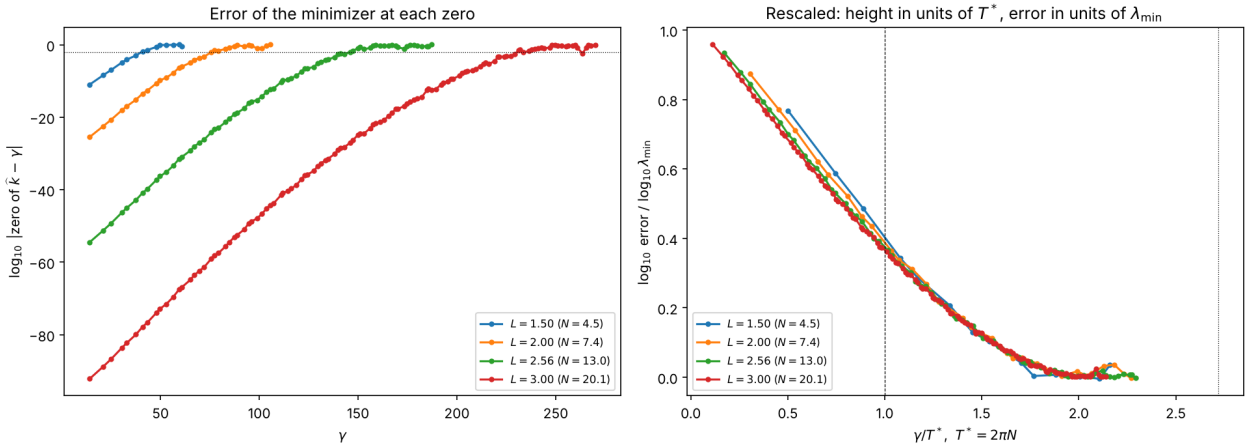


FIGURE 6. The minimizer of Weil's form in the windows $N = 4.5, 7.4, 13, 20.1$ [CE21]. Left: $\log_{10}$ of the distance from each zero of ζ to the nearest zero of the minimizer's transform. Right: the same data with the height in units of $T^* = 2\pi N$ and the error in units of $\lambda_{\min}$; dashed T^*, dotted eT^*.

7. Companion papers

The Caceres prime function [CE10]. The weight $a(n) = \sum_{p^k \| n} \log p\, p^{-k/2}$ has Dirichlet series $\zeta(s)\mathcal{K}(s)$. Its kernel $\mathcal{K}$ is $-\frac{\zeta'}{\zeta}(s+\frac{1}{2}) + \frac{\zeta'}{\zeta}(2s+\frac{1}{2}) - \frac{1}{12s}$ plus a function holomorphic in $\Re s > -\frac{1}{8}$, so it sees each zero twice, at $i\gamma$ and at $i\gamma/2$. In the language of this paper, the kernel's margin is $M(N) - 2M(\sqrt{N})$ under RH, and its slope is $\alpha(x) - \alpha(\sqrt{x})$ without hypothesis; its energy is $2(\log\xi)''(\frac{1}{2})$, less $2\sum\gamma^{-2}$ over the pairs of ordinates with $\gamma' = 2\gamma$. Since $a = 1 * b$ with b fixed by the primes, the kernel carries no information beyond Λ, and Theorem 23 applies to it verbatim.

The $6k \pm 1$ race [CE37]. The same construction for the character χ_{-3} gives a margin M_χ whose positivity is equivalent to GRH for $L(s, \chi_{-3})$. It holds with $0.102 \leq M_\chi \leq 0.355$ for $2 \leq N \leq 2^{27}$. In the margin's weighting, the prime powers $\equiv 2 \pmod 3$ lead those $\equiv 1 \pmod 3$ by exactly

$2\frac{L'}{L}(\frac{1}{2},\chi_{-3})\log N - \frac{1}{2}\psi'(\frac{3}{4}) + M_\chi(N)$, up to exponentially small terms. So under GRH the weighted $6k \pm 1$ race never reverses.

8. The margin at large N: a probabilistic projection

Proposition 11 says that under RH the margin comes arbitrarily close to 0. This section asks how fast. The answer is a projection, not a theorem: it assumes RH and the linear independence of the ordinates (LI), as in the work of Rubinstein and Sarnak [RuSa] and of Akbary, Ng and Shahabi [ANS] on limiting distributions.

The limiting distribution. Under RH, (3) writes $M(e^L) = \sum_{\gamma>0} a_\gamma(1-\cos\gamma L)$ with $a_\gamma = 4/\gamma^2$. Under LI, by the Kronecker–Weyl theorem the phases γL modulo 2π become equidistributed on the infinite torus as $L \to \infty$, so the proportion of $L \in [0,T]$ with $M(e^L) < x$ tends to $\mathbb{P}(X < x)$, where

$$X = \sum_{\gamma>0} a_\gamma(1-\cos\theta_\gamma), \qquad \mathbb{E}\, e^{itX} = \prod_{\gamma>0} e^{ita_\gamma} J_0(ta_\gamma), \tag{23}$$

with independent θ_γ uniform on $[0,2\pi)$. Its mean is $2(\log\xi)''(\frac{1}{2}) = 0.0924$ and its standard deviation $(8\sum\gamma^{-4})^{1/2} = 0.01724$. Fourier inversion of (23) with the 4520 certified zeros (the rest contribute their mean 0.00098 and a variance below $3\cdot10^{-11}$) gives its density, a smooth bump on $[0.04, 0.145]$ with two shallow maxima. For the lower tail we use the saddle point: with $K(s) = \log\mathbb{E}e^{sX} = \sum_\gamma \big(sa_\gamma + \log I_0(sa_\gamma)\big)$ and $K'(\hat s) = x$, the Lugannani–Rice formula gives

$$\mathbb{P}(X<0.04) = 3.7\cdot10^{-6}, \quad \mathbb{P}(X<0.03) = 1.0\cdot10^{-11},$$
$$\mathbb{P}(X<0.02) = 4.2\cdot10^{-27}, \quad \mathbb{P}(X<0.01) = 1.7\cdot10^{-97},$$

in agreement with Fourier inversion to 0.6% where both apply ($x \ge 0.03$). Numerically $\log\mathbb{P}(X < x) \approx -C/x^2$: halving x multiplies the exponent by about 3.7.

How often the margin dips. The rate of downcrossings of a level x, per unit of $L = \log N$, is given by Rice's formula $\nu(x) = f_X(x)\,\mathbb{E}\big[|X'|\,\big|\,X = x\big]/2$, with $X' = \sum_\gamma \gamma a_\gamma \sin\theta_\gamma$. Conditioning on $X = x$ by exponential tilting turns each θ_γ into a von Mises variable of concentration $\kappa_\gamma = |\hat s| a_\gamma$, for which $\mathbb{E}\sin^2\theta = I_1(\kappa)/(\kappa I_0(\kappa))$; taking X' conditionally Gaussian,

$$\nu(x) = \frac{f_X(x)}{\sqrt{2\pi}}\Big(\sum_{\gamma>0}\frac{16}{\gamma^2}\frac{I_1(\kappa_\gamma)}{\kappa_\gamma I_0(\kappa_\gamma)}\Big)^{1/2}, \qquad f_X(x) \approx \frac{e^{K(\hat s)-\hat s x}}{\sqrt{2\pi K''(\hat s)}}. \tag{24}$$

We tested the model against the function itself, computed from the zeros on $7 \le L \le 3000$: for $x = 0.060, 0.055, 0.050, 0.045$ the observed proportion of time below x is $2.12\cdot10^{-2}$, $6.1\cdot10^{-3}$, $1.17\cdot10^{-3}$, $1.2\cdot10^{-4}$ against $2.14\cdot10^{-2}$, $6.3\cdot10^{-3}$, $1.16\cdot10^{-3}$, $1.1\cdot10^{-4}$, and the observed number of excursions per unit of L on $[1000, 3000]$ is 0.686, 0.259, 0.068, 0.0085 against 0.683, 0.262, 0.065, 0.0083 from (24).

The record low. If dips below x arrive as a Poisson process of rate $\nu(x)$, then $\mathbb{P}\big(\min_{N'\le N} M(N') \ge x\big) \approx e^{-\nu(x)\log N}$. The median value of $\log N$ at which M first drops below x is $\log 2/\nu(x)$:

x	0.040	0.035	0.030	0.025	0.020	0.015	0.010
$1/x$	25	28.6	33.3	40	50	66.7	100
median $\log_{10} N$	$7.5\cdot10^2$	$7.9\cdot10^4$	$1.3\cdot10^8$	$2.7\cdot10^{13}$	$1.2\cdot10^{23}$	$5.8\cdot10^{42}$	$6.4\cdot10^{92}$

Since $\log\mathbb{P}(X<x) \approx -C/x^2$, the record low decreases like $(\log\log N)^{-1/2}$. Fitting the median gives

$$\max_{10^3\le N'\le N}\frac{1}{M(N')} \approx 10.4\sqrt{\log_{10}\log N} \qquad (N\to\infty), \tag{25}$$

which reproduces the median to within 3% for $\log_{10}\log N \ge 13$; for smaller N the median lies above the law (Figure 7(b)).

The actual record, under RH only. Under RH alone, $M(e^L) = W(L) + T(L)$, where W is the sum (3) over the 4520 certified zeros and $0 \le T \le 2\cdot0.00098$, T having mean 0.00098. We followed the

record low of $W + 0.00098$ for $7 \le L \le 10^5$: a coarse pass with the first 100 zeros, whose partial sum is a pointwise lower bound for W, selects the candidates, which are then refined with all 4520 zeros. The records are

$$0.0430\ (N = 3.4 \cdot 10^7), \quad 0.04269\ (1.06 \cdot 10^{11}), \quad 0.0399\ (2.2 \cdot 10^{29}),$$
$$0.0385\ (10^{1521}), \quad 0.0369\ (10^{8274}), \quad 0.0362\ (10^{24461}),$$

the first two agreeing with the primes (Proposition 8) to 10^{-5}. The last three fall inside the 10–90% band of the model; the first three come early (each had probability 2–6% of arriving so soon), so the region $N < 10^{30}$ is somewhat atypically low.

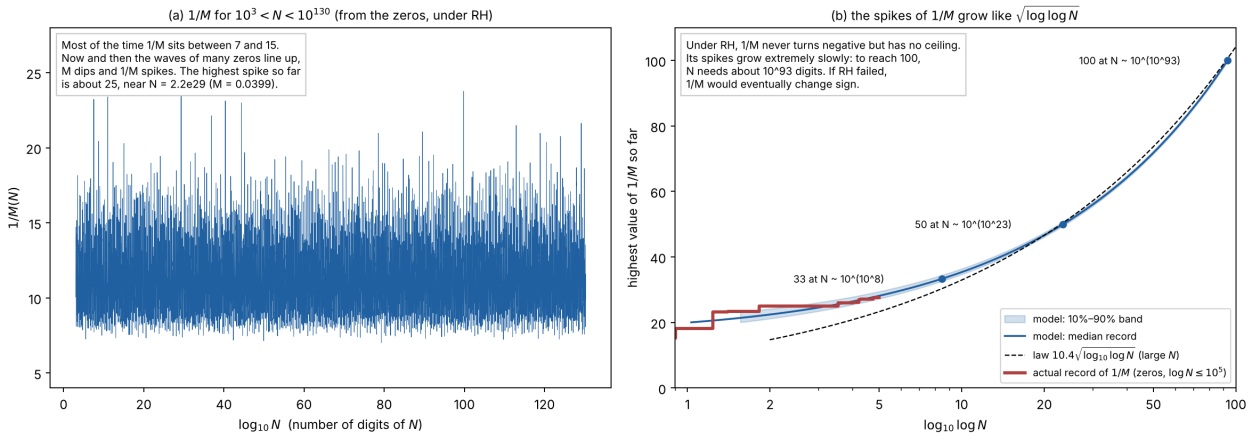


FIGURE 7. The reciprocal $1/M$. (a) $1/M(N)$ for $10^3 < N < 10^{130}$, from the certified zeros under RH: it stays between 7 and 15 most of the time and spikes when many waves are in phase. (b) The record of $1/M$ against $\log_{10} \log N$: actual records for $\log N \le 10^5$ (red), the median record and 10–90% band of the model (24) under RH and LI (blue), and the law (25) (dashed).

Remark 29 (status). Formula (23) for the limiting distribution is a theorem under RH and LI; the saddle-point values are numerical approximations checked against Fourier inversion; the rate (24), the Poisson approximation and the law (25) are heuristics, checked against the function itself where it can be computed. The projection describes how the margin behaves if RH holds. Read through the reciprocal, it says that $1/M$ is positive and unbounded, and that its spikes grow like $\sqrt{\log \log N}$, so that $1/M$ reaches 100 only for N with about 10^{93} digits. If RH failed, $1/M$ would eventually change sign.

9. The variance of the margin and the area of the quadruples

The mean of the margin in $\log N$ is half the window; this section shows that its variance is also a constant of ξ, the fourth logarithmic derivative at the centre, and it records, as a separate and elementary identity, how far the non-holomorphic sum $\sum |t_\rho|^{-4}$ is from that constant: by the squared areas of the rectangles formed by the zeros and their reflections. The two statements meet only under RH, where the areas vanish.

The quadruple and its area. A zero off the line comes with three others. With $\rho = \frac{1}{2} + \delta + i\gamma$, $\delta > 0$, $\gamma > 0$, the functional equation and the reality of ξ on the real axis give the four zeros $\frac{1}{2} \pm \delta \pm i\gamma$. They are the vertices of a rectangle centred at $s = \frac{1}{2}$, of width 2δ and height 2γ, so of area

$$A_\rho = 4\,\delta\,\gamma = 2\left|\operatorname{Im} t_\rho^2\right| = 2\left|\operatorname{Im} \lambda_\rho\right|, \qquad t_\rho = \gamma - i\delta, \quad \lambda_\rho = \rho(1-\rho) = \tfrac{1}{4} + t_\rho^2, \tag{26}$$

since $t_\rho^2=\gamma^2-\delta^2-2i\delta\gamma$ (Figure 8(a)). A zero on the line is a degenerate rectangle of area 0. In the variable t_ρ^2, or in λ_ρ, the quadruple becomes a pair of complex conjugates, and its area is twice their distance from the real axis. So RH is the statement that every quadruple has area zero, which is the statement that ξ, as a function of λ, has only real zeros. The margin sees the zeros through t_ρ^{-2} (formula (3) and Proposition 1), and the area is the imaginary part of the variable it uses.

Proposition 30 (the squared-area gap; the power sums over the zeros are classical [Leh])**.** *Without any hypothesis, the sums below converge absolutely, the sums being over all nontrivial zeros with multiplicity, and*

$$\mathcal{A}:=\frac12\sum_\rho\Big(\frac{1}{|t_\rho|^4}-\operatorname{Re}\frac{1}{t_\rho^4}\Big)=\sum_\rho\Big(\operatorname{Im}\frac{1}{t_\rho^2}\Big)^2=\sum_\rho\frac{A_\rho^2}{4\,|t_\rho|^8}\ \ge0,\qquad\sum_\rho\frac{1}{t_\rho^4}=-\frac16\,(\log\xi)^{(4)}(\tfrac12).\tag{27}$$

In particular $\mathcal{A}=0$ if and only if RH holds, and $\sum_\rho|t_\rho|^{-4}=-\frac16(\log\xi)^{(4)}(\frac12)+2\mathcal{A}$.

Proof. For a complex number u, $|u|^2-\operatorname{Re}u^2=2(\operatorname{Im}u)^2$; take $u=t_\rho^{-2}$, and $\operatorname{Im}t_\rho^{-2}=-\operatorname{Im}t_\rho^2/|t_\rho|^4$, which by (26) has modulus $A_\rho/(2|t_\rho|^4)$. The sums converge absolutely because $|t_\rho|\asymp|\rho|$ and $\sum|\rho|^{-2}<\infty$. By the Hadamard product, $\frac{d^k}{ds^k}\log\xi(s)=(-1)^{k-1}(k-1)!\sum_\rho(s-\rho)^{-k}$ for $k\ge2$, and at $s=\frac12$, $(s-\rho)^4=(it_\rho)^4=t_\rho^4$. Finally $\mathcal{A}=0$ if and only if every $A_\rho=0$, that is, $\delta=0$ for every zero. □

The two sums in (27) have different natures. The sum of t_ρ^{-4} is holomorphic in the zeros: it is a Taylor coefficient of $\log\xi$ at the centre and is computed without knowing a single zero. The sum of $|t_\rho|^{-4}$ is not holomorphic, and it is the kind of quantity that a mean square sees. This is the same structure as the gap $\mathcal{Q}$ of Part 5 [CE5], $\sum_\rho|\rho|^{-2}=2+\gamma_E-\log4\pi+\mathcal{Q}$, with $\mathcal{Q}$ a sum of δ^2 terms: here the defect is a sum of squared areas. The holomorphic sum is classical (the power sums of the zeros; see [Vo]), and the identity $|u|^2-\operatorname{Re}u^2=2(\operatorname{Im}u)^2$ is elementary. The next proposition says that, under RH, the mean square of the margin is that non-holomorphic sum.

Proposition 31 (the logarithmic mean and variance of the margin)**.** *Assume RH, and let m_γ be the multiplicity of the ordinate $\gamma>0$. Then, with $L=\log N$,*

$$\lim_{T\to\infty}\frac1T\int_0^T M(e^L)\,dL=2(\log\xi)''(\tfrac12),\qquad V_M:=\lim_{T\to\infty}\frac1T\int_0^T\big(M(e^L)-2(\log\xi)''(\tfrac12)\big)^2dL=8\sum_{\gamma>0}\frac{m_\gamma^2}{\gamma^4},$$

and

$$V_M\ \ge\ 4\sum_\rho\frac{1}{|\gamma_\rho|^4}=-\frac23\,(\log\xi)^{(4)}(\tfrac12),$$

with equality if and only if all zeros are simple. No hypothesis of linear independence of the ordinates is needed.

Proof. Under RH, (3) reads $M(e^L)=c-\sum_{\gamma>0}a_\gamma\cos(\gamma L)$ with $a_\gamma=4m_\gamma/\gamma^2$ over the distinct positive ordinates and $c=\sum a_\gamma=4\sum_{\gamma>0}m_\gamma\gamma^{-2}=2\sum_\rho\gamma_\rho^{-2}=2(\log\xi)''(\frac12)$, by the Hadamard product as above ($(\frac12-\rho)^2=-\gamma_\rho^2$). The series converges absolutely and uniformly, so $M(e^L)$ is a uniformly almost periodic function of L with distinct frequencies γ, and Parseval's identity for such functions [Bes] gives the mean c and the variance $\sum a_\gamma^2/2=8\sum m_\gamma^2\gamma^{-4}$. Since $m_\gamma^2\ge m_\gamma$, $V_M\ge8\sum_{\gamma>0}m_\gamma\gamma^{-4}=4\sum_\rho|\gamma_\rho|^{-4}$, with equality if and only if every $m_\gamma=1$, and under RH $\mathcal{A}=0$, so Proposition 30 gives the value $-\frac23(\log\xi)^{(4)}(\frac12)$. □

What happens without RH. If there is a quadruple of positive area, the unconditional formula $M(N)=2\sum_\rho(1-N^{it_\rho})/t_\rho^2$ contains, for each of its two zeros with $\delta>0$, a term of modulus about $2N^\delta/|t_\rho|^2$, since $|N^{it_\rho}|=N^\delta$; the margin leaves its window (Theorem 3) and its mean square over $[0,L]$ is unbounded as $L\to\infty$ (by the Cauchy–Schwarz and Laplace-transform argument of

Proposition 16); in this sense V_M is infinite. So the three quantities are ordered as follows. If RH holds, V_M is finite and exceeds the holomorphic constant $-\frac{2}{3}(\log\xi)^{(4)}(\frac12)$ only through multiple zeros. If RH fails, the holomorphic constant is still finite, $\sum|t_\rho|^{-4}$ exceeds it by twice the squared-area gap $\mathcal{A}>0$, and V_M is infinite. The variance does not measure the area; it is the area that decides whether the variance exists. This is the exponent form of the common wall of the book [CE0]: a quadruple of small width 2δ shows up only at $\log N \gtrsim 1/\delta$.

The constants. A least-squares fit of $\log\xi(\frac12+x)$, $0.01\le x\le 0.35$, by an even polynomial of degree 8, with ζ evaluated by Euler–Maclaurin, gives

$$(\log\xi)''(\tfrac12)=0.04620999,\qquad (\log\xi)^{(4)}(\tfrac12)=-4.46071\cdot10^{-4},$$

against $2\sum_{\gamma>0}\gamma^{-2}=0.04620999$ and $-12\sum_{\gamma>0}\gamma^{-4}=-4.460707\cdot10^{-4}$ from the 1517 zeros below 2000 (with the tail of $\sum\gamma^{-2}$ added; the tail of $\sum\gamma^{-4}$ is below 10^{-10}). So the prediction of Proposition 31 for simple zeros is

$$V_M=-\tfrac{2}{3}(\log\xi)^{(4)}(\tfrac12)=2.9738\cdot10^{-4},\qquad \sqrt{V_M}=0.017245,$$

the value used, under RH and LI, in Section 8; Proposition 31 shows that LI is not needed for it.

The measurement. We computed $M(N)$ from the 11 080 801 prime powers below $2\cdot10^8$, at $4\cdot10^5$ equally spaced values of $L=\log N$ in $[\log10,\log(2\cdot10^8)]$, and the right side of (3) from the 1517 zeros below 2000, with the zeros above 2000 replaced by their mean contribution. The two agree to $8.0\cdot10^{-5}$ for $N\ge10^3$, and their statistics are:

range of N	mean (primes)	mean (zeros)	variance (primes)	variance (zeros)
$[10,\,2\cdot10^8]$	0.09251	0.09248	$2.9929\cdot10^{-4}$	$2.9917\cdot10^{-4}$
$[10^3,\,2\cdot10^8]$	0.09242	0.09242	$3.0103\cdot10^{-4}$	$3.0103\cdot10^{-4}$
$[10^5,\,2\cdot10^8]$	0.09226	0.09226	$3.0228\cdot10^{-4}$	$3.0228\cdot10^{-4}$
$L\to\infty$ (Prop. 31)	0.092420		$2.9738\cdot10^{-4}$	

The primes and the zeros agree to three to four digits on every range, and the difference from the limiting values, between 0.6% and 1.7% in the variance, is the finite length of the window in L (about 38 periods of the first ordinate), which the zeros reproduce exactly. The running variance stays within 2% of $2.9738\cdot10^{-4}$ (Figure 8(b)), and the distribution of M is the non-Gaussian bump of Section 8 (Figure 8(c)), with least value 0.0430 and largest 0.1378 on this range. Status: Propositions 30 and 31 are proved; the constants and the table are numerical (`m_variance.py`, `area_fig.py`, `logxi_derivs.py`).

Two readings of the area. The same area appears in Part 3 [CE3]: at a zero off the line, the imaginary part of the reflected cross-energy tends to $\operatorname{Im}\frac{1}{\lambda_\rho}=\frac{A_\rho}{2|\lambda_\rho|^2}$, the only first-order view of δ in the book, and for the Davenport–Heilbronn zero $0.8085+85.699i$ the area is 105.75 while $A/(2|\lambda|^2)=9.8\cdot10^{-7}$. The Hankel–Jacobi criterion of Part 3 and the Hankel gauge of Part 27 ask that every λ_ρ be real, that is, that every area vanish. The margin asks the same question through $t_\rho^2=\lambda_\rho-\frac14$, and its variance is the sesquilinear (non-holomorphic) side of the gap (27).

10. The walk of the primes against the curve

Proposition 9 can be read as a geometric reformulation of the whole problem. Write $P_q=(A_q,B_q)$ for the pair of prime sums at the prime power q, and

$$D_q=2B_q+K_0-\Phi(u_q),\qquad \Phi(u)=8u(\log u-1),\qquad u_q=\frac{2A_q+c_0}{4}=\sqrt{q}+\frac{\alpha(q)}{2},\tag{28}$$

the last equality because $A_q=2\sqrt{q}+\alpha(q)-\frac{c_0}{2}$. By (9), D_q is, up to $|\epsilon_q|\le3\cdot10^{-6}$, the value of the local minimum of M attached to the block $[q,q')$ when that minimum is interior. Every local minimum of M on $N\ge2$ is of this form (Proposition 9(3)), and a sign change of the continuous function M, positive near $N=1$, would happen at a negative local minimum. So RH is the statement

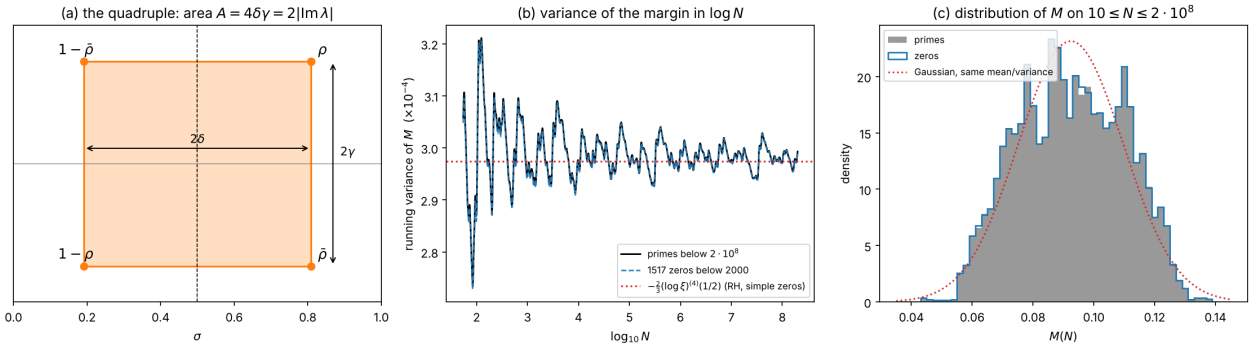


FIGURE 8. (a) The quadruple $\frac{1}{2} \pm \delta \pm i\gamma$ of an off-line zero: a rectangle of area $A = 4\delta\gamma = 2|\operatorname{Im}\lambda|$ (not to scale). (b) The running variance of $M(e^L)$ over $[\log 10, L]$, from the primes below $2 \cdot 10^8$ (black) and from the 1517 zeros below 2000 (dashed), against $-\frac{2}{3}(\log\xi)^{(4)}(\frac{1}{2}) = 2.9738 \cdot 10^{-4}$ (dotted). (c) The distribution of $M(N)$ on $10 \le N \le 2 \cdot 10^8$ in the measure dL, from the primes and from the zeros, and the Gaussian with the same mean and variance.

that *the walk $q \mapsto P_q$, pushed forward one prime power at a time, never crosses the fixed curve* $2B + K_0 = \Phi\big(\frac{2A+c_0}{4}\big)$ beyond some point; D_q is the height of the walk above the curve (Figure 9).

10.1. **The exact step.**

Proposition 32 (the step of the walk)**.** *Let $q < q'$ be consecutive prime powers, $\lambda = \Lambda(q')q'^{-1/2}$. Then, for some $\xi \in (u_q, u_{q'})$,*

$$D_{q'} - D_q = 2\lambda\Big(\log\frac{q'}{q} - 2\log\Big(1 + \frac{\alpha(q)}{2\sqrt{q}}\Big)\Big) - \frac{\lambda^2}{\xi}. \tag{29}$$

In particular the step is $2\lambda\,\log(q'/N_q^*) + O(\lambda^2/\sqrt{q})$: the prime power q' lowers *the floor when it arrives before the pending minimum N_q^*, and* raises *it when it arrives after; and no single prime power moves the floor appreciably: the first inequality below holds for every $q \ge 10^6$ with $|\alpha(q)| \le q^{1/4}$ (true throughout the computed range), and the second under the same hypothesis for $10^6 \le q < q' \le 4\cdot 10^{18}$,*

$$|D_{q'} - D_q| \le \frac{2\log q'}{\sqrt{q'}}\Big(\log\frac{q'}{q} + \frac{|\alpha(q)|+1}{\sqrt{q}}\Big) \le \frac{2\log q\,(|\alpha(q)|+2)}{q}.$$

Proof. $A_{q'} = A_q + \lambda$, $B_{q'} = B_q + \lambda\log q'$, $u_{q'} = u_q + \frac{\lambda}{2}$. So $D_{q'} - D_q = 2\lambda\log q' - \big(\Phi(u_q + \frac{\lambda}{2}) - \Phi(u_q)\big)$, and by Taylor with $\Phi'(u) = 8\log u$, $\Phi''(u) = 8/u$, the bracket is $4\lambda\log u_q + \lambda^2/\xi$. Now $2\log u_q = \log q + 2\log(1 + \frac{\alpha(q)}{2\sqrt{q}})$, which gives (29); and $2\log u_q = \log N_q^* + O(u_q^{-5})$ by Proposition 9(3). The first inequality follows from (29), $\Lambda(q') \le \log q'$, $\log(1+v) \le v$ for $v \ge 0$, $-\log(1-x) \le x + x^2$ for $0 \le x \le \frac{1}{2}$ (with $x = |\alpha|/2\sqrt{q}$ the excess in $2|\log(1 + \frac{\alpha}{2\sqrt{q}})|$ over $|\alpha|/\sqrt{q}$ is at most $\alpha^2/2q \le 1/(2\sqrt{q})$), and $\lambda^2/\xi \le 2\lambda \cdot 0.008/\sqrt{q}$. The second needs $\log(q'/q) \le q^{-1/2}$, that is a gap $q' - q \le \sqrt{q}$ between consecutive prime powers: the maximal prime gaps are below $1500 \le \sqrt{q}$ for $2.25 \cdot 10^6 \le q \le 4 \cdot 10^{18}$ and at most $148 < 10^3 \le \sqrt{q}$ for $10^6 \le q < 2.25 \cdot 10^6$ [OHP]. We also used that $\log q/\sqrt{q}$ decreases. Beyond $4 \cdot 10^{18}$ a gap below $\sqrt{q}$ is not known unconditionally, and we do not claim the second inequality there; between 10^{12} and $4 \cdot 10^{18}$ the hypothesis $|\alpha(q)| \le q^{1/4}$ is outside our computed range and remains a hypothesis of the statement (it would follow from an explicit bound for $\psi(t) - t$ of Büthe's type by partial summation, which we have not carried out). □

Identity (29) was checked against the exact D_q over 10^5-long stretches at $q \sim 10^6$, agreeing to $1.9 \cdot 10^{-11}$; the largest single step there is $1.6 \cdot 10^{-5}$, falling to $9 \cdot 10^{-8}$ at $q \sim 10^8$ and $3.3 \cdot 10^{-9}$

at $q \sim 10^9$. The two terms of the step are the two clocks of the problem: the gap $\log(q'/q)$, what the primes do, against $\alpha(q)/\sqrt{q}$, where the running account stands. Their accumulated difference is $-2\int \alpha\, d\log x$ (Section 3), so no individual prime matters and only the *area* under the slope does.

10.2. **What it takes to cross.**

Proposition 33 (local crossing bound)**.** *Let $x < y$ and suppose M changes sign in $[x, y]$, with $M(x) > 0$. Then*

$$\max_{t\in[x,y]} \alpha(t) \;\geq\; \frac{M(x)}{2\log(y/x)},$$

and the maximum is attained at a prime power of $[x, y]$ or at $t = x$. Moreover $\alpha(t) - \alpha(x)$ depends only on the prime powers in $(x, t]$: $\alpha(t) - \alpha(x) = \sum_{x<n\leq t} \Lambda(n)n^{-1/2} - 2(\sqrt{t} - \sqrt{x})$.

Proof. With $m(L) = M(e^L)$, $m' = -2\alpha(e^L) + \tau_1(L)$ and $\tau_1 > 0$ (Section 3), so for $L \in [\log x, \log y]$, $m(L) \geq m(\log x) - 2(L - \log x)\max_{[x,y]} \alpha \geq M(x) - 2\log(y/x)\max\alpha$. If M changes sign the left side is < 0 somewhere. Between prime powers $\alpha(t) = A - 2\sqrt{t} + \frac{c_0}{2}$ decreases, and it jumps up at each prime power, so its maximum on a closed block sits at a prime power. □

TABLE 2. What a first sign change in a dyadic block $[x, 2x]$ would require, from Proposition 33 with the verified lower bounds for $M(x)$: a surplus $\alpha \geq M(x)/(2\log 2)$, i.e. about $\alpha\sqrt{x}/\log x$ prime powers above the smooth count, present at some prime power of the block.

x	10^{12}	10^{20}	10^{27}	10^{29}
$M(x) \geq$	0.0426	0.0481	0.0481	0.0341
needs $\alpha \geq$	0.031	0.035	0.035	0.025
extra prime powers	$1.1 \cdot 10^3$	$7.5 \cdot 10^6$	$1.8 \cdot 10^{10}$	$1.2 \cdot 10^{11}$

Status: both propositions are proved, and they answer a natural question – can one added prime, or one added zero, flip the sign of M? – precisely. A zero on the line only adds the nonnegative window $4(1-\cos\gamma L)/\gamma^2$; an off-line zero flips the sign for large N, but that is RH itself; a single prime power beyond 10^6 in the computed range moves the floor by less than $7 \cdot 10^{-5}$ (Proposition 32; the largest step observed is $1.6 \cdot 10^{-5}$). Proposition 33 does not close the argument: the required surplus, about 0.15 standard deviations of α (whose mean square is $(\log\xi)''(\frac{1}{2})$ by Theorem 20), occurs in a positive proportion of blocks. The content is the exact bookkeeping: a sign change cannot start in a block where the primes run behind their smooth count, and the surplus must be present at a prime power, locally auditable through the second identity of Proposition 33. The obstruction to converting this into a proof is the one of Theorem 23: the bound uses $|\alpha|$, not the identity $\Lambda = \mu * \log$.

10.3. **Anatomy of the close approaches.** The walk has come closest to the curve at the record minima of Section 8. Two complementary mechanisms show up, one in the primes and one in the zeros.

In the primes. At the record minimum near $N = 34\,186\,454$ ($M = 0.042997$, between the prime powers 34 186 367 and 34 186 499), the surplus was assembled one octave earlier: on $(8.5 \cdot 10^6, 1.7 \cdot 10^7]$ the weighted count of prime powers exceeds $2\sqrt{x}$ by $+0.223$ in the slope – about 49 extra prime powers – while the neighbouring octaves run near balance ($+0.017$) or behind (-0.193, -0.203). The final descent is short: M crosses its mean 0.0924 for the last time at $N = 3.09 \cdot 10^7$ and reaches the record a factor 1.105 later, paying $2\int \alpha\, d\log x = 0.0494$ with α averaging $+0.25$ on the stretch; the bottom is reached exactly when the surplus is spent: $\alpha = 0.0149$ at the last prime power 34 186 367, and α vanishes at the minimum $N^* = 34\,186\,454.25$ (the minimum is the zero of α, Proposition 9). Figure 10(a) shows α from the primes against $\alpha = -2\sum_\gamma \sin(\gamma\log x)/\gamma$ from the 4520 certified

zeros: the two are indistinguishable at this scale. At the next record, $N = 1.056 \cdot 10^{11}$, the primes (Proposition 8(2)) give 0.042686 and the zeros give 0.042689: the same event seen from both sides, to $3 \cdot 10^{-6}$.

In the zeros. At every record the low-lying zeros conspire: of the first 20 phases $\cos(\gamma_k \log N)$, between 13 and 16 exceed $\frac{1}{2}$ simultaneously (Figure 10(b)), the first 10 zeros carry 74–92% of the depth $0.0924 - M$, and the first 100 carry 99%. Independent phases would align like this with probability of order $3^{-14} \approx 10^{-7}$ per sample, which is the rarity that the rates of Section 8 quantify; the next deeper conspiracies are predicted near $N \sim 10^{10^5}$.

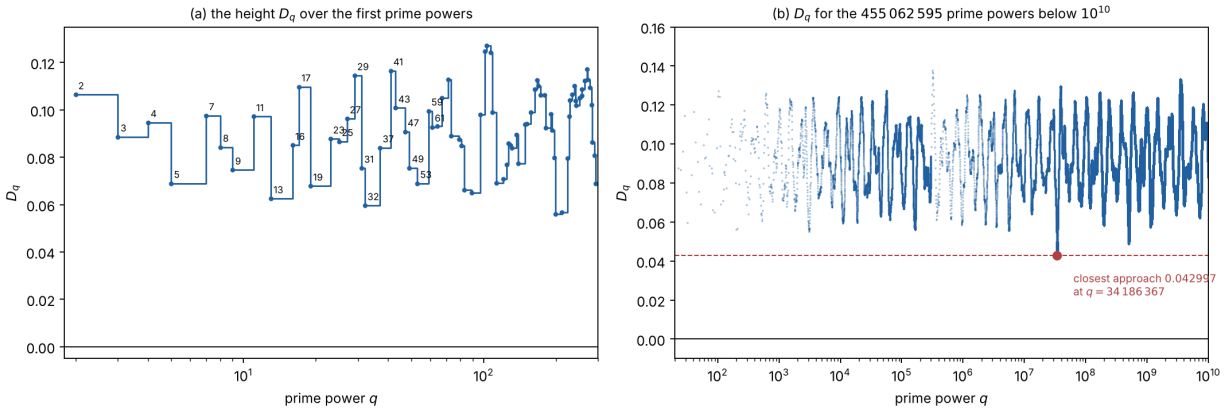


FIGURE 9. The walk against the curve. (a) The height D_q of (28) over the first prime powers: each prime power moves the point, down when it arrives before the pending minimum N_q^*, up when after (29). (b) D_q for all prime powers below 10^{10}; the closest approach 0.042997 at $q = 34\,186\,367$ is the minimum of M on $[10^3, 10^{10}]$ (beyond 10^{10} the verified lower bound is smaller, 0.0426 near $1.06 \cdot 10^{11}$).

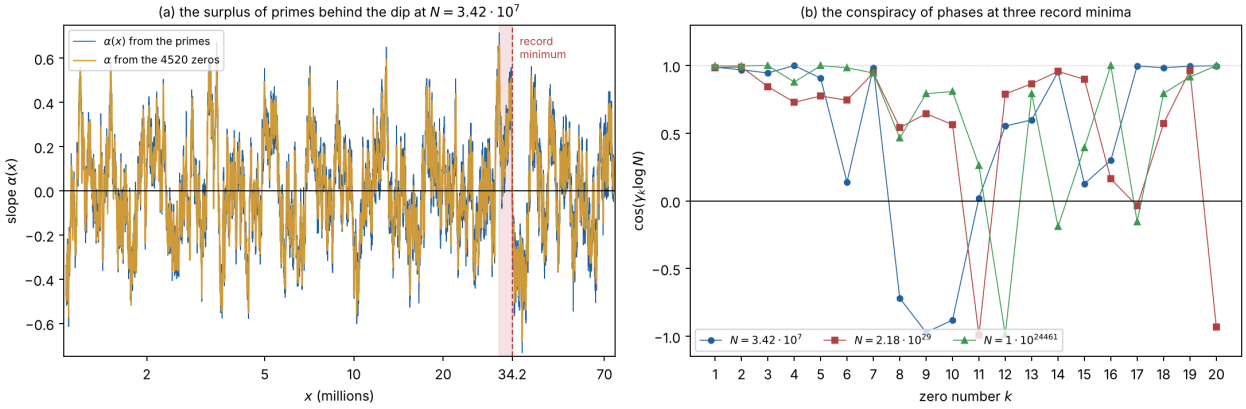


FIGURE 10. Anatomy of the record minima. (a) The slope α around the record at $N = 3.42 \cdot 10^7$, from the primes and from the zeros: a surplus of about 49 prime powers accumulates on $(8.5 \cdot 10^6, 1.7 \cdot 10^7]$, the margin pays for it on the shaded stretch, and the record sits where α returns to 0. (b) The phases $\cos(\gamma_k \log N)$ of the first 20 zeros at three records: the dips are conspiracies of the low zeros.

11. The margin as a one-sided mean of the prime-counting error

The attempt to push the verification beyond the computed range with explicit prime-counting bounds fails, and it fails for a reason worth recording exactly. Let

$$K_N(t) = t^{-3/2}\Big(1+\tfrac{1}{2}\log\tfrac{N}{t}\Big), \qquad E^*(t) = \psi(t)-t+\log 2\pi+\tfrac{1}{2}\log(1-t^{-2}), \qquad \widehat{B} = 1+\frac{\gamma_E}{2}-\frac{1}{2}\log 4\pi,$$

so that E^* is the remainder of ψ against the full smooth model of the explicit formula, $\widehat{B} = \sum_\rho \rho^{-1} = 0.0230957\ldots$ is the Hadamard constant of ξ, and $\int_1^N K_N(t)\,dt = \log N$ exactly, since $-t^{-1/2}\log(N/t)$ is an antiderivative of K_N.

Proposition 34 (the margin as a weighted mean)**.** *For every $N > 1$, exactly and unconditionally,*

$$M(N) = -2\int_1^N E^*(t)\,K_N(t)\,dt - 2\widehat{B}\log N. \tag{30}$$

Proof. All three steps are elementary. *First*, $C(N) = 2\int_1^N \psi(t)K_N(t)\,dt$: by the antiderivative above, $2\int_q^N K_N = 2q^{-1/2}\log(N/q)$, so the right side is $\sum_{q<N} 2\Lambda(q)q^{-1/2}\log(N/q) = C(N)$, with no boundary terms. *Second*, with $S(t) = t - \log 2\pi - \frac{1}{2}\log(1-t^{-2})$, termwise integration with $\int_1^N t^{a-1}(1+\frac{1}{2}\log(N/t))\,dt = \frac{N^a-1}{a}+\frac{N^a-1-a\log N}{2a^2}$, $a = -(2k+\frac{1}{2})$ for the k-th term of $-\frac{1}{2}\log(1-t^{-2}) = \frac{1}{2}\sum_k t^{-2k}/k$, gives

$$2\int_1^N S(t)K_N(t)\,dt = 8\sqrt{N} + \Big(2-5\log 2-2\log\pi-\frac{\pi}{2}\Big)\log N + K_0 - \tau(\log N):$$

the $N^{-2k-\frac{1}{2}}$ coefficients are $\frac{1}{k}\big(\frac{1}{a}+\frac{1}{2a^2}\big) = -\frac{2}{(2k+\frac{1}{2})^2}$, which is $-\tau$ term by term; the $\log N$ coefficient uses $\sum_{k\ge1}\frac{1}{k(4k+1)} = \psi(\frac{5}{4})+\gamma_E = 4-3\log 2-\frac{\pi}{2}$; and the constant uses $\sum_{k\ge1}\frac{1}{k(4k+1)^2} = 8-3\log 2-\frac{\pi}{2}-\frac{\pi^2}{4}-2G$, which follows from the trigamma value $\psi_1(\frac{1}{4}) = \pi^2+8G$, and comes out as exactly $K_0 = -16+\frac{\pi^2}{2}+4G$. *Third*, comparing with (1),

$$R(N)-2\int_1^N S(t)K_N(t)\,dt = \big(-c_0-2+5\log 2+2\log\pi+\tfrac{\pi}{2}\big)\log N = (\log 4\pi-\gamma_E-2)\log N = -2\widehat{B}\log N.$$

Since $\psi = S + E^*$, $M = R - C = -2\int E^*K_N - 2\widehat{B}\log N$. □

Identity (30) was checked to machine precision against (1) for $1.5 \le N \le 10^{15}$ (`ident.py`). The proof never mentions the zeros; the Hadamard constant $\widehat{B} = \sum_\rho \rho^{-1}$ materializes on the prime side out of the digamma values at $\frac{1}{4}$.

Corollary 35 (RH as one inequality for one mean)**.** *Let $\mu(N) = \frac{1}{\log N}\int_1^N E^*(t)K_N(t)\,dt$, the K_N-weighted mean of E^*, so that $M(N) = -2\log N\,(\mu(N)+\widehat{B})$. Then RH holds if and only if*

$$\mu(N) \le -\widehat{B} \qquad \textit{for all sufficiently large } N,$$

and in that case $-\widehat{B}-\frac{0.18484}{2\log N} \le \mu(N) \le -\widehat{B}$ for $N > 1$, so $\mu(N) \to -\widehat{B}$ one-sidedly. Unconditionally, $\mu(N) \le -\widehat{B} + E(N)/(2\log N)$ by Theorem 6, and the verification gives $\mu(N) \le -\widehat{B} - 0.0170/\log N$ for $10^3 \le N \le 10^{29}$ and $\mu(N) < -\widehat{B}$ up to $2.15\cdot 10^{29}$ (Figure 11).

Three remarks on what (30) says, and what it cannot do.

Where the weight sits. The measure $K_N(t)\,dt$ has total mass $\log N$, of which the interval $[1,2]$ – before the first prime – carries the share $1-2^{-1/2}+O(1/\log N)$, about 30%, and $[1,4]$ about 51%; the oscillation at height t enters with density $\sim t^{-1/2}$ per unit. The mean is dominated by the first integers.

The head start and the payback. On $[1,2]$, $\psi = 0$ and the integrand is forced: that interval alone contributes $d \log N + 0.115$ to $M(N)$, with $d = \int_1^2 S(t) t^{-3/2} dt = 0.06344$. Since the total drift in (30) is $-2\widehat{B} \log N = -0.04619 \log N$, the primes from 2 on must pay back at the exact asymptotic rate $d - 2\widehat{B} = 0.01725$ per unit of $\log N$ – they must run, in this weighted sense, slightly *ahead* of the smooth model forever – and RH is the statement that they never overpay: the accumulated account $\mu(N) + \widehat{B}$ reaches 0 from below at most.

Why explicit prime bounds are powerless here. The sharpest explicit estimates of the ψ-remainder, such as $|\psi(t) - t| \le 0.94\sqrt{t}$ for $11 < t \le 10^{19}$ [Bu], give only $|\mu(N)| \le 0.94\,(1 + \frac{1}{4}\log N) + O(1)$, which does not even keep μ bounded; the criterion needs the *signed* mean accurate to the absolute constant $\widehat{B} = 0.0231$. At $N = 10^{12}$ the gap is a factor of about 300. This is the quantitative form of the obstruction of Section 5: positivity of M is a statement about cancellation in E^* at the scale $\sqrt{t}$, which no absolute-value estimate can see, and Theorem 23 shows that the structural inputs of the energy arguments do not suffice. We record (30) because it compresses RH into a single one-sided inequality for a single weighted mean, with an explicit classical constant on the other side.

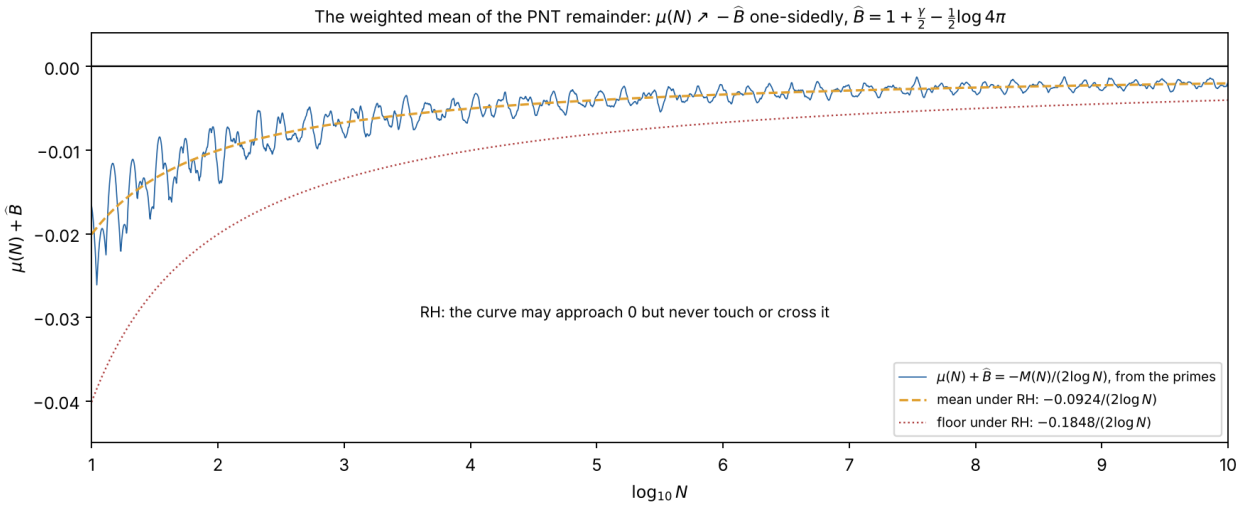


FIGURE 11. The one-sided convergence that is RH: $\mu(N) + \widehat{B} = -M(N)/(2\log N)$, computed from the primes up to 10^{10}, against the mean $-0.0924/(2\log N)$ and the extreme $-0.1848/(2\log N)$ allowed under RH. RH says the blue curve approaches the axis from below and never touches it.

12. Discussion

Every statement in this paper that is equivalent to RH is equivalent through a Landau-type argument, and each reduces to the holomorphy of $\frac{\zeta'}{\zeta}$, or of $\frac{1}{\zeta}$, to the right of the critical line. The unconditional results are of two kinds. Some are exact identities between the primes and the margin: the Fejér formula, the description of the minima, the energy identity and the Möbius representation. The others are statements of what cannot work: no uniform positive lower bound for M under RH (Proposition 11), and no argument that survives planted weights (Theorem 23). The second kind gives a test that any proposed argument can be run against: if it also works for the weights (19), or for a Beurling system [DMV], it cannot prove RH. What remains is an estimate that uses $N(x) = \lfloor x \rfloor$ exactly, through the sawtooth of Proposition 25. The best unconditional estimates of that kind save $\exp(-c(\log x)^{3/5}(\log\log x)^{-1/5})$, the Vinogradov–Korobov region. A saving of $x^{-1/2+\varepsilon}$ would be RH. We do not prove the Riemann hypothesis.

Remark 36 (Weil positivity for small support, and where this family sits). The margin evaluates the Weil functional on the one-parameter Fejér family: the autocorrelation underlying $M(N)$ has multiplicative support $[N^{-1}, N]$, so the primes enter only through the powers $q < N$. The regime in which Weil positivity is *proved* for all test functions is the archimedean one: Connes and Consani [CC] establish, through the prolate and Sonin operators, a positivity inequality for every g supported in $[2^{-1/2}, 2^{1/2}]$ (autocorrelation in $[\frac{1}{2}, 2]$), precisely the range in which no prime occurs; the semilocal programme [CCM] is to add the places $2, 3, \dots, p$ and the support (p^{-1}, p) one prime at a time. Specialized to our family, the proved range is $N \le 2$, where $M = R$ and positivity is the elementary estimate of Proposition 8(1); their theorem is of course much stronger there, being an operator bound valid for all g. The map between the two languages is Section 10: adding a place is admitting the next batch of prime powers into the walk, and the question both programmes face is the same uniformity – positivity must survive as the support, and with it every prime, enters. Theorem 23 constrains the operator route too: any semilocal positivity argument robust under the planted weights (19) cannot close RH, so the arithmetic identity $\Lambda = \mu * \log$ must enter the operator estimates at some stage. We do not see how to do this, and record the translation in the hope that the one-variable family, where every object is computable, is a useful test bench for that programme.

Data and code

The verification of Proposition 8 is `certificado_segmentado.py` (Python's `decimal` only). The minima and their recurrence (Section 3, Figure 3) come from `minimos.py`, `recurrencia.py`, `fig_fondos.py` and `fig_fondos_rec.py` in `anc/minimos`. The block energies to 10^{10} and the energy identity come from `blocks.py` (segmented sieve, extended precision); the planted systems from `fake.py`; Figure 4 from `fig_energy.py`, all in `anc/energy`. The Möbius representation and Figure 5 come from `compute.py`, `saw.py` and `figure.py` in `anc/mobius`. The wave (Figure 2) and the windows are computed by the ancillary files of [CE13, CE21]. Theorem 6, the extended certificate and Figure 1 come from `ebound.py`, `ebound_rig.py`, `cert_ext.py`, `cert_up.py`, `certify_zeros.py` and `fig_ext.py`, and part (2) of Proposition 8 from `verify_margin.py` in `anc/verification`. Section 8 and Figure 7 come from `dist.py`, `saddle.py`, `rice.py`, `sweep.py`, `project.py` and `fig_inv.py` in `anc/projection`. Section 10 comes from `traj.py`, `stepcheck.py`, `anatomy.py`, `phases.py` and `figs.py` in `anc/walk`. Section 11 and Figure 11 come from `ident.py`, `mu.py` and `fig_mu.py` in `anc/mean` in `anc/explicit`, with the ordinates in `zeros5000.npy`. All are included with this paper.

Use of AI assistance

The cross-energy, the one-variable formulation and the research programme are the author's. Claude (Anthropic) was used as a research and writing aid: to derive and check proofs, to write and run the computations, to test statements against the controls, to search the literature, and to draft and edit text and figures. The author checked all results and is responsible for the content.

References

[CC] A. Connes, C. Consani, Weil positivity and trace formula, the archimedean place, *Selecta Math.* 27 (2021), 77; arXiv:2006.13771.

[CCM] A. Connes, C. Consani, H. Moscovici, Zeta zeros and prolate wave operators, arXiv:2310.18423, 2023.

[Bu] J. Büthe, An analytic method for bounding $\psi(x)$, *Math. Comp.* 87 (2018), 1991–2009.

[Arb] F. Johansson, Arb: efficient arbitrary-precision midpoint-radius interval arithmetic, *IEEE Trans. Comput.* 66 (2017), 1281–1292.

[ANS] A. Akbary, N. Ng, M. Shahabi, Limiting distributions of the classical error terms of prime number theory, *Q. J. Math.* 65 (2014), 743–780.

[Bes] A. S. Besicovitch, *Almost Periodic Functions*, Cambridge University Press, 1932.

[Bo] E. Bombieri, Remarks on Weil's quadratic functional in the theory of prime numbers I, *Rend. Mat. Acc. Lincei* (9) 11 (2000), 183–233.

[Co] A. Connes, The Riemann hypothesis: past, present and a letter through time, arXiv:2602.04022, 2026.

[DH] H. Davenport, H. Heilbronn, On the zeros of certain Dirichlet series I, II, *J. London Math. Soc.* 11 (1936), 181–185, 307–312.

[Go] S. M. Gonek, On negative moments of the Riemann zeta-function, *Mathematika* 36 (1989), 71–88.

[DMV] H. G. Diamond, H. L. Montgomery, U. M. A. Vorhauer, Beurling primes with large oscillation, *Math. Ann.* 334 (2006), 1–36.

[IK] H. Iwaniec, E. Kowalski, *Analytic Number Theory*, AMS Colloquium Publications 53, 2004.

[In] A. E. Ingham, *The Distribution of Prime Numbers*, Cambridge Tracts 30, Cambridge University Press, 1932.

[Jn] D. R. Johnston, Improving bounds on prime counting functions by partial verification of the Riemann hypothesis, *Ramanujan J.* 59 (2022), 1307–1321.

[RuSa] M. Rubinstein, P. Sarnak, Chebyshev's bias, *Experiment. Math.* 3 (1994), 173–197.

[MV] H. L. Montgomery, R. C. Vaughan, *Multiplicative Number Theory I. Classical Theory*, Cambridge University Press, 2007.

[KLN] H. Kadiri, A. Lumley, N. Ng, Explicit zero density for the Riemann zeta function, *J. Math. Anal. Appl.* 465 (2018), 22–46; arXiv:2101.12263.

[Ng] N. Ng, The distribution of the summatory function of the Möbius function, *Proc. London Math. Soc.* (3) 89 (2004), 361–389.

[PT] D. Platt, T. Trudgian, The Riemann hypothesis is true up to $3 \cdot 10^{12}$, *Bull. Lond. Math. Soc.* 53 (2021), 792–797.

[Su1] M. Suzuki, Aspects of the screw function corresponding to the Riemann zeta-function, arXiv:2206.03682, 2022 (v4, 2023).

[Su2] M. Suzuki, On variants of Chebyshev's conjecture, *Ramanujan J.* (2025), doi:10.1007/s11139-025-01238-9, arXiv:2411.07436.

[T] E. C. Titchmarsh, *The Theory of the Riemann Zeta-Function*, 2nd ed., Oxford University Press, 1986, Section 10.25.

[Tr] T. S. Trudgian, An improved upper bound for the argument of the Riemann zeta-function on the critical line II, *J. Number Theory* 134 (2014), 280–292.

[OHP] T. Oliveira e Silva, S. Herzog, S. Pardi, Empirical verification of the even Goldbach conjecture and computation of prime gaps up to $4 \cdot 10^{18}$, *Math. Comp.* 83 (2014), 2033–2060.

[Vo] A. Voros, *Zeta Functions over Zeros of Zeta Functions*, Lecture Notes of the Unione Matematica Italiana 8, Springer, 2010.

[We52] A. Weil, Sur les "formules explicites" de la théorie des nombres premiers, *Comm. Sém. Math. Univ. Lund* (1952), tome supplémentaire, 252–265.

[Wi] D. V. Widder, *The Laplace Transform*, Princeton University Press, 1941.

[CE0] P. Caceres, Cross-Energy of Dirichlet Partial Sums. Part 0: An overview of the series, preprint, 2026.

[CE3] P. Caceres, Cross-Energy of Dirichlet Partial Sums. Part 3: The reflected cross-energy, preprint, 2026.

[CE5] P. Caceres, Cross-Energy of Dirichlet Partial Sums. Part 5: The energy of the primes, preprint, 2026.

[CE10] P. Caceres, Cross-Energy of Dirichlet Partial Sums. Part 10: The Caceres prime function: a square-root damped prime weight, its exact prime zeta structure, and a dual spectrum of the Riemann zeros, preprint, 2026.

[CE12] P. Caceres, Cross-Energy of Dirichlet Partial Sums. Part 12: One constant, three mirrors: $\widehat{B} = 1 + \frac{\gamma_E}{2} - \frac{1}{2}\log 4\pi$ from the zeros, from the primes, and from the Möbius sawtooths, preprint, 2026.

[CE13] P. Caceres, Cross-Energy of Dirichlet Partial Sums. Part 13: The Euler cross-energy, Weil positivity and a one-variable form of the Riemann hypothesis, preprint, 2026.

[CE14] P. Caceres, Cross-Energy of Dirichlet Partial Sums. Part 14: The margin as $W - E$, a Hodge index for the integers, preprint, 2026.

[CE21] P. Caceres, Cross-Energy of Dirichlet Partial Sums. Part 21: The minimizer of the Weil form in a window, and the scale $2\pi N$, preprint, 2026.

[CE23] P. Caceres, Cross-Energy of Dirichlet Partial Sums. Part 23: Ghost zeros and edge mass: how the minimizers of Weil's form in a window react to zeros off the critical line, preprint, 2026.

[CE29] P. Caceres, Cross-Energy of Dirichlet Partial Sums. Part 29: Weil windows as density operators: mixtures, coherences and negativity, preprint, 2026.

[CE36] P. Caceres, Cross-Energy of Dirichlet Partial Sums. Part 36: How random is the Möbius function? An exact variance identity, measured Gaussian statistics, and the wall before the Riemann hypothesis, preprint, 2026.

[CE37] P. Caceres, Cross-Energy of Dirichlet Partial Sums. Part 37: The $6k \pm 1$ race: sieve structure, Chebyshev's bias and a GRH-equivalent inequality, preprint, 2026.

[CE24] P. Caceres, Cross-Energy of Dirichlet Partial Sums. Part 24: The Weil window: a two-sided instrument on the cross-energy's test functions, preprint, 2026.

[CE27] P. Caceres, Cross-Energy of Dirichlet Partial Sums. Part 27: The Hankel gauge: Matiyasevich's moments, Lehmer's sums, and a graded positivity instrument for the Riemann hypothesis, preprint, 2026.

[CE30] P. Caceres, Cross-Energy of Dirichlet Partial Sums. Part 30: The antiphase lock: the mutual information of prime blocks and the zero-free band, preprint, 2026.

[Leh] D. H. Lehmer, The sum of like powers of the zeros of the Riemann zeta function, *Math. Comp.* 50 (1988), 265–273.

Universidad Europea de Valencia, Valencia, Spain
Email address: pedrojesus.caceres@universidadeuropea.es