

Prime-shift operators on Björner's complex of squarefree integers: an exact decomposition of the Mertens invariant, exact kernel asymptotics, rigorous recovery bounds for the Guinand–Weil pipeline, and a structural obstruction to the spectral modification

Luca Eliseo Pavesi

ORCID: <https://orcid.org/0009-0003-0532-2850>¹

¹Independent Researcher , lucapv02@gmail.com

September 18, 2026

Abstract

Let Δ_n denote Björner's simplicial complex on the squarefree integers $m \leq n$ and let $H_n = \sum_{p \leq n} c_p(T_p + T_p^*)$ be the self-adjoint operator obtained from the prime-shift operators T_p with bounded weights c_p . This paper is the sixth revision of [9], superseding the first five. It contains corrections of errors, new exact results, an exact kernel decomposition of the Mertens residual with an explicit critical exponent, rigorous stability bounds for the Toeplitz–Hankel pipeline, and a structural obstruction to the spectral modification conjectured in the fourth revision.

Corrections. (i) The bipartite block $A : C_{\text{even}} \rightarrow C_{\text{odd}}$ of H_n is the *full* sum $A = A_{\uparrow} + A_{\downarrow}$, with $Ae_m = \sum_{p|m, pm \leq n} c_p e_{pm} + \sum_{p|m} c_p e_{m/p}$. (ii) The maximal elements lie in $\ker A_{\uparrow}|_{C_{\text{even}}}$, not in $\ker A$. The earlier statement $\ker A = \{0\}$ is false: at $n = 10^4$, $\dim \ker A = \dim C_{\text{even}} - \text{rank}(A) = 3030 - 2141 = 889$. What is true is that $\ker A$ is strictly smaller than $\ker A_{\uparrow}|_{C_{\text{even}}}$. (iii) The values c_* reported in the earlier revisions correspond to $A_{\uparrow}$, not to A ; at $n = 10^4$, $\text{rank}(A_{\uparrow}) = 1125$, $\text{rank}(A) = 2141$, so $c_* = 0.185$ while the full density is $c = 0.352$. (iv) The kernel of H_n therefore has density $1 - 2c \approx 0.296$, not $1 - 2c_* \approx 0.630$.

New exact results. (v) We prove that the maximal elements of $[1, n]$ are exactly $m = P_k m'$ with $P_k = p_1 \cdots p_k$ the k -th primorial, m' squarefree coprime to P_k , and $m' \in (n/(p_{k+1}P_k), n/(p_k P_k)]$. This yields the exact formula for $M_{\mathcal{M}}(n)$. (vi) We prove that the correct decomposition of the Mertens invariant is $M(n) = M_{\mathcal{M}}(n) + \dim R_n^{\uparrow} - \dim R_n^{\uparrow, T}$, not the identity previously conjectured.

Exact kernel decomposition (new in the sixth revision). (vii) We introduce the exact kernel $c_R^{\text{exact}}(\rho, n)$ of the residual $\mathcal{R}(n) := M(n) - M_{\mathcal{M}}(n)$, prove that it converges to the asymptotic kernel $c_R^{\text{asym}}(\rho)$ of the fifth revision as $n \rightarrow \infty$, and establish numerically that the convergence is power-law, $|c_R^{\text{exact}}(\rho_k, n) - c_R^{\text{asym}}(\rho_k)| \sim n^{-\alpha_k}$, over the first 175 zeros of ζ . The fitted exponents have sample mean 0.4918 and sample standard deviation 0.0506, consistent with the fifth-revision values (0.4888, 0.0498) within the standard error. Because the 175 fits share the same grid in n (80 log-spaced points in $[10^2, 10^8]$), the naive standard error of the mean $0.0506/\sqrt{175} \approx 0.0038$ underestimates the true uncertainty: a block bootstrap with $B = 8$ contiguous blocks gives $\text{SE}_{\text{boot}} = 0.0078$, so the deviation of the sample mean from $1/2$ is approximately 1.0 bootstrap standard error and the hypothesis $\alpha = 1/2$ cannot be rejected. A logarithmic-correction hypothesis $\Delta_k(n) \propto n^{-1/2}(\log n)^{-\beta_k}$ is rejected by the data: the fitted β_k have sample mean -0.081 (wrong sign), the predicted deficit has the wrong sign, and the t -statistics $\beta_k/\text{SE}(\beta_k)$ are distributed symmetrically around zero. Sensitivity to the truncation parameter $K \in \{10, 20, 30, 50\}$ in (24) is null to

machine precision, so the truncation is not a source of systematic error. Correlations with $\log \gamma_k$ and with k are -0.011 and -0.067 respectively.

Rigorous recovery. (viii) We prove truncation bounds for the prime-power sum and the archimedean integral in the explicit formula, both of order $O(e^{-c^2/4})$ in the truncation parameter. (ix) We prove a stability theorem for ESPRIT: the recovery error scales as $\varepsilon e^{\sigma^2 \gamma_M^2}$, with the exponential amplification being intrinsic to the inverse problem. (x) We prove a bound on the Vandermonde conditioning $C_{M,\delta,K}$ that is exponential in M , identifying the fundamental numerical bottleneck of the pipeline.

Structural obstruction. (xi) We prove two independent obstructions showing that the spectral modification $\tilde{H}_n = (I - P_n)H_n(I - P_n)$ conjectured in the fourth revision *cannot* converge to the zeros of ζ . (I) For $s > 1/2$, $\|H_n\|_{\text{op}} \leq 2P(s)$, so the spectral measure is compactly supported. (II) The second moment of H_n is finite, while the second moment of the prime term in the explicit formula diverges: H_n misses all prime powers p^j , $j \geq 2$. A third, heuristic obstruction (III) is that the density of eigenvalues of H_n in any fixed interval is non-logarithmic, while the density of zeros γ_ρ is logarithmic; unlike (I) and (II), obstruction (III) is not proved rigorously here. We also correct an error of the fourth revision: the operators $\gamma_p = T_p + T_p^*$ do *not* commute in the truncated complex.

No proof of the Riemann Hypothesis is claimed.

Keywords: Riemann zeta function, explicit formula, Toeplitz pencil, Hankel–ESPRIT, Hilbert–Pólya, Björner complex, squarefree integers, Mertens function, prime-shift operators, exact kernel, critical exponent.

MSC 2020: 11M26, 11M06, 47B35, 65F15, 11Y35.

1 Introduction

The Hilbert–Pólya conjecture asserts that the nontrivial zeros $\rho = \frac{1}{2} + i\gamma_k$ of the Riemann zeta function $\zeta(s)$ are the imaginary parts of the eigenvalues of a self-adjoint operator on a suitable Hilbert space. Several concrete proposals have appeared in the literature; see [1, 3, 6, 11, 16] and the recent finite-dimensional construction of [12]. None has produced a proof of RH, and several structural obstructions are known.

This paper is the **sixth** revision of [9]. The history of corrections is instructive and we record it honestly.

- The first revision [10] proposed a spurious “prefactor correction” $1/(8\pi\sigma\sqrt{\pi})$ for the prime-power term in the explicit formula. That correction was erroneous and was withdrawn in the second revision.
- The second revision corrected the prefactor to $1/(4\sigma\sqrt{\pi})$ (which is correct) and added a closed form for the density of maximal elements.
- The third revision extended the numerical range to $n \leq 2 \times 10^5$, showing that both estimates drift, and did not attempt to isolate a new characterisation of RH.
- The fourth revision identified a structural error affecting all earlier revisions (the block A is the up-shift $A_\uparrow$ versus the full $A = A_\uparrow + A_\downarrow$), formulated the spectral modification $\tilde{H}_n$ as an open conjecture, and analyzed the Toeplitz–Hankel pipeline numerically.
- The fifth revision corrected an error of the fourth revision ($\ker A \neq \{0\}$, and γ_p does not commute in the truncated complex), proved rigorous truncation and stability bounds for the pipeline, and proved two independent structural obstructions showing that $\tilde{H}_n$ cannot converge to the zeros of ζ .

- The present (sixth) revision introduces the exact kernel $c_R^{\text{exact}}(\rho, n)$ of the residual $M(n) - M_{\mathcal{M}}(n)$, proves its convergence to the asymptotic kernel $c_R^{\text{asym}}(\rho)$, and establishes numerically that the convergence follows a power law with exponent numerically close to $1/2$. A block-bootstrap analysis, a sensitivity test with respect to the truncation parameter, and a direct test of the logarithmic-correction hypothesis are carried out in §5; the results support the hypothesis $\alpha = 1/2$ and exclude the logarithmic-correction family, while leaving open the origin of the small residual deficit in the sample mean.

The paper contains five types of changes with respect to the fourth revision.

- (A) **Corrections.** The statement $\ker A = \{0\}$ of the fourth revision is false; we identify the true kernel dimension. The statement that γ_p commutes is false in the truncated complex.
- (B) **Exact kernel decomposition.** We introduce the exact kernel of the residual and identify its critical exponent.
- (C) **Rigorous recovery bounds.** We prove truncation bounds for the prime-power sum and the archimedean integral in the explicit formula, a stability theorem for ESPRIT, and a bound on the Vandermonde conditioning.
- (D) **Structural obstruction.** We prove that the spectral modification $\tilde{H}_n$ cannot converge to the zeros of ζ , for two independent rigorous reasons and a third heuristic one.
- (E) **Numerical validation.** We verify the pipeline end-to-end on the first 175 zeros of ζ , including a block-bootstrap analysis of the exponent α_k , a sensitivity test with respect to K , and a test of the logarithmic-correction hypothesis.

Main results. Let Δ_n be the abstract simplicial complex on the squarefree integers $m \leq n$ whose faces are the subsets of primes dividing m . The reduced Euler characteristic of Δ_n equals $-M(n)$ where $M(n)$ is the Mertens function, and Björner [2] explicitly invites the study of deeper topological invariants of Δ_n . We study the *Björner–Weil operator*

$$H_n = \sum_{p \leq n} c_p(T_p + T_p^*), \quad (1)$$

acting on the total chain complex $C_*(\Delta_n; \mathbb{R})$, together with the Toeplitz pencil built from the Guinand–Weil formula.

- (1) **Structure of H_n (§2–§4).** H_n is self-adjoint, has $\mathbb{Z}/2$ -symmetric spectrum and satisfies the trace formula $\text{Tr}(H_n^2) = 2 \sum_{p \leq n} c_p^2 \sigma_p(n/p)$.
- (2) **Kernel structure (§4).** With the corrected block $A = A_{\uparrow} + A_{\downarrow}$: $\dim \ker(H_n) = d_n - 2 \text{rank}(A)$ and $\text{index}(A) = M(n)$. The kernel of $A_{\uparrow}$ contains the maximal elements.
- (3) **Exact formula for maximal elements (§4.1).** The maximal elements of $[1, n]$ are exactly $m = P_k m'$ with $P_k = p_1 \cdots p_k$ and $m' \in (n/(p_{k+1} P_k), n/(p_k P_k)]$ squarefree, coprime to P_k .
- (4) **Kernel density (§4.4).** At $n = 10^4$: $\text{rank}(A_{\uparrow}) = 1125$, $\text{rank}(A) = 2141$, $c_* = 0.185$, $c = 0.352$, $\dim \ker H_n / d_n = 1 - 2c \approx 0.296$.
- (5) **Exact kernel decomposition (§5).** The residual $\mathcal{R}(n) := M(n) - M_{\mathcal{M}}(n)$ has an exact representation as a sum over zeros with kernel $c_R^{\text{exact}}(\rho, n)$, and $c_R^{\text{exact}}(\rho, n) \rightarrow c_R^{\text{asym}}(\rho)$ as $n \rightarrow \infty$. The fitted exponent over the first 175 zeros has sample mean 0.4918 ± 0.0506 ; a block bootstrap gives $\text{SE}_{\text{boot}} = 0.0078$ and the hypothesis $\alpha = 1/2$ is not rejected. The logarithmic-correction family is excluded by the data.

- (6) **Rigorous recovery bounds (§9).** Truncation bounds (Lemmas 9.1 and 9.2), ESPRIT stability (Theorem 9.4), and Vandermonde conditioning (Theorems 9.6 and 9.7).
- (7) **Structural obstruction (§7).** Two rigorous and one heuristic obstruction showing that $\tilde{H}_n$ cannot converge to the zeros of ζ .
- (8) **Toeplitz pencil (§8).** Confirmation of the prefactor $1/(4\sigma\sqrt{\pi})$, numerical verification of the identity to 10^{-13} , exact-recovery theorem, and the Hankel–ESPRIT pipeline.

No proof of the Riemann Hypothesis is claimed.

2 The Björner–Weil operator

2.1 The complex Δ_n and the shifts T_p

Let Δ_n be the abstract simplicial complex on the squarefree integers $m \leq n$ whose faces are the subsets of primes dividing m . We denote by $C_k(\Delta_n; \mathbb{R})$ the $\mathbb{R}$ -vector space spanned by the squarefree integers $m \leq n$ with $\Omega(m) = k$, with the standard orthonormal basis $\{e_m\}$; for squarefree m one has $\Omega(m) = \omega(m)$, so C_k corresponds to the $(k-1)$ -faces of Δ_n , with $C_0 = \mathbb{R}e_1$.

For a prime $p \leq n$, define the up-shift $T_p : C_k \rightarrow C_{k+1}$ by

$$T_p(e_m) = \begin{cases} e_{pm}, & p \nmid m, pm \leq n, \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

Let $T_p^* : C_{k+1} \rightarrow C_k$ be the adjoint with respect to the orthonormal basis, so that $T_p^*(e_{m'}) = e_{m'/p}$ if $p \mid m'$ and 0 otherwise.

Lemma 2.1. *Both T_p and T_p^* are partial isometries, and $\|T_p\|_{\text{op}} = \|T_p^*\|_{\text{op}} = 1$.*

Proof. T_p maps each basis vector either to a basis vector or to zero, and the map $m \mapsto pm$ is injective, so distinct basis vectors in the support are sent to distinct basis vectors. Hence $T_p^*T_p$ is the orthogonal projection onto $\text{span}\{e_m : p \nmid m, pm \leq n\}$, which is non-zero because $T_p e_1 = e_p$ for $p \leq n$. Therefore $\|T_p\|_{\text{op}} = 1$, and the same for the adjoint. $\square$

Lemma 2.2. $T_p^2 = 0$ and $(T_p^*)^2 = 0$.

Lemma 2.3. *For distinct primes $p \neq q$ one has*

$$T_p T_q = T_q T_p, \quad T_p^* T_q^* = T_q^* T_p^*. \quad (3)$$

Proof. Both sides of the first identity send e_m to e_{pqm} when $p, q \nmid m$ and $pqm \leq n$, and to 0 otherwise; both sides of the second send e_m to $e_{m/pq}$ when $pq \mid m$, and to 0 otherwise (no truncation constraint arises, since $m/pq \leq m \leq n$). $\square$

Remark 2.4. The mixed identity $T_p T_q^* = T_q^* T_p$, asserted in the earlier versions, is *false* once the complex is truncated at n . Indeed $T_p T_q^* e_m$ requires $q \mid m$, $p \nmid m$ and $pm/q \leq n$, whereas $T_q^* T_p e_m$ requires the strictly stronger condition $pm \leq n$. Concretely, for $n = 10$, $p = 5$, $q = 3$ and $m = 6$:

$$T_5 T_3^*(e_6) = T_5(e_2) = e_{10}, \quad T_3^* T_5(e_6) = T_3^*(0) = 0.$$

Commutation is restored only in the untruncated infinite-dimensional model. Similarly, the earlier statement $T_p^* T_q = 0$ for $p \neq q$ is false: with $p = 2$, $q = 3$ and $e_2 \in C_1$ one has $T_2^* T_3(e_2) = e_3 \neq 0$. These facts do not affect the trace formula of Theorem 3.2, because the mixed terms have vanishing trace.

2.2 Self-adjointness and $\mathbb{Z}/2$ symmetry

Given weights $c_p \in \mathbb{R}$ for each prime $p \leq n$, define $H_n = \sum_{p \leq n} c_p (T_p + T_p^*)$.

Proposition 2.5. *The operator H_n is self-adjoint.*

Proof. Each summand $c_p (T_p + T_p^*)$ is self-adjoint because c_p is real and $(T_p + T_p^*)^* = T_p^* + T_p$. $\square$

Let $W : C_k \rightarrow C_k$ denote the total-weight operator $W(e_m) = \Omega(m)e_m$, and let P be the parity operator $P|_{C_k} = (-1)^k \text{Id}$.

Lemma 2.6. $[W, T_p] = T_p$ and $[W, T_p^*] = -T_p^*$.

Proposition 2.7. $PH_nP = -H_n$.

Proof. T_p raises Ω by one and T_p^* lowers it by one, so $PT_pP = -T_p$ and $PT_p^*P = -T_p^*$. $\square$

Corollary 2.8. *The spectrum of H_n is symmetric about zero. In particular, the signature $N_+ - N_-$ of H_n vanishes.*

Remark 2.9. The earlier identification of the Mertens function with the signature of H_n is incorrect: the signature vanishes identically. The correct invariant is the index of the bipartite block A , as discussed in §4.

3 Trace formulas and moments

Lemma 3.1. *For every prime $p \leq n$,*

$$\text{Tr}(T_p T_p^*) = \text{Tr}(T_p^* T_p) = \#\{m \leq n/p : m \text{ squarefree}, p \nmid m\}. \quad (4)$$

Theorem 3.2. *The second moment satisfies*

$$\text{Tr}(H_n^2) = 2 \sum_{p \leq n} c_p^2 \sigma_p(n/p), \quad (5)$$

where $\sigma_p(x) = \#\{m \leq x : m \text{ squarefree}, p \nmid m\}$.

Proof. $H_n^2 = \sum_{p,q} c_p c_q (T_p + T_p^*)(T_q + T_q^*)$. Expanding, the diagonal terms $p = q$ give $c_p^2 (T_p T_p^* + T_p^* T_p)$, and $\text{Tr}(T_p T_p^*) = \text{Tr}(T_p^* T_p) = \sigma_p(n/p)$. The cross terms $p \neq q$ have vanishing trace: $T_p T_q^*$ maps e_m to a multiple of $e_{pm/q} \neq e_m$ whenever non-zero, and $T_p T_q, T_p^* T_q^*$ shift Ω by ± 2 . $\square$

Theorem 3.3. *Let $\sigma(n) = \#\{m \leq n : \mu^2(m) = 1\}$ denote the number of squarefree integers $\leq n$. For $c_p = p^{-s}$ with $s > 0$, the limit*

$$\lim_{n \rightarrow \infty} \frac{\text{Tr}(H_n^2)}{n} = \frac{12}{\pi^2} \sum_p \frac{p^{-2s}}{p+1} \quad (6)$$

holds. Equivalently, normalising by $\sigma(n) \sim (6/\pi^2)n$,

$$\lim_{n \rightarrow \infty} \frac{\text{Tr}(H_n^2)}{\sigma(n)} = 2 \sum_p \frac{p^{-2s}}{p+1}. \quad (7)$$

Proof. For fixed p , the squarefree integers not divisible by p have natural density $\prod_{q \neq p} (1 - q^{-2}) \cdot (1 - p^{-1}) = (6/\pi^2) p/(p+1)$, whence $\sigma_p(x) \sim (6/\pi^2) x/(1 + 1/p)$. By Theorem 3.2, $\text{Tr}(H_n^2)/n = 2 \sum_{p \leq n} p^{-2s} \sigma_p(n/p)/n$, and each summand converges to $(6/\pi^2) p^{-2s}/(p+1)$. The trivial bound $\sigma_p(n/p) \leq n/p$ dominates the p -th summand by $2p^{-2s-1}$, summable over primes. Dominated convergence gives the first formula; dividing by $\sigma(n) \sim 6n/\pi^2$ gives the second. $\square$

Proposition 3.4 (Carleman). *Let $c_p = p^{-s}$ with $s > 1$ and put $P(s) = \sum_p p^{-s} < \infty$. Then for every $k \geq 1$ and every n ,*

$$\mathrm{Tr}(H_n^{2k}) \leq \sigma(n) (2P(s))^{2k}. \quad (8)$$

Consequently the numbers $m_{2k} := \limsup_n \mathrm{Tr}(H_n^{2k})/\sigma(n)$ satisfy $m_{2k} \leq (2P(s))^{2k}$, so that $\sum_k m_{2k}^{-1/(2k)} = \infty$; every weak subsequential limit of the empirical spectral measures of H_n is compactly supported and uniquely determined by its moments.

Proof. For any basis vector e_m , $\|H_n e_m\|_1 = \sum_{p|m, pm \leq n} c_p + \sum_{p|m} c_p \leq 2P(s)$, where $\|\cdot\|_1$ is the ℓ^1 norm in the basis $\{e_m\}$. Iterating gives $\|H_n^{2k} e_m\|_1 \leq (2P(s))^{2k}$ and $|\langle e_m, H_n^{2k} e_m \rangle| \leq (2P(s))^{2k}$. Summing over the $\sigma(n)$ basis vectors gives the bound. Carleman's condition follows. $\square$

Remark 3.5. For $s \leq 1$ one has $P(s) = \infty$ and the bound is vacuous. The earlier claim that Carleman's criterion holds unconditionally was stated without proof and is not asserted here.

4 Kernel structure

We write the bipartite form of H_n with respect to the splitting $C_* = C_{\text{even}} \oplus C_{\text{odd}}$. Since T_p raises Ω by one and T_p^* lowers it by one, both T_p and T_p^* map C_{even} to C_{odd} . Hence

$$H_n = \begin{pmatrix} 0 & A^* \\ A & 0 \end{pmatrix}, \quad (9)$$

where $A : C_{\text{even}} \rightarrow C_{\text{odd}}$ is the full off-diagonal block,

$$Ae_m = \underbrace{\sum_{\substack{p|m \\ pm \leq n}} c_p e_{pm}}_{=: A_{\uparrow} e_m} + \underbrace{\sum_{p|m} c_p e_{m/p}}_{=: A_{\downarrow} e_m} \quad (\Omega(m) \text{ even}). \quad (10)$$

We write $A_{\uparrow} = \sum_{p \leq n} c_p T_p$ and $A_{\downarrow} = \sum_{p \leq n} c_p T_p^*$, viewed as operators $C_{\text{even}} \rightarrow C_{\text{odd}}$, so that $A = A_{\uparrow} + A_{\downarrow}$.

Proposition 4.1. $\dim \ker(H_n) = d_n - 2 \mathrm{rank}(A)$, where $d_n = \dim C_*(\Delta_n)$.

Proof. $(v, w) \in C_{\text{even}} \oplus C_{\text{odd}}$ lies in $\ker H_n$ iff $Av = 0$ and $A^*w = 0$. Hence $\dim \ker H_n = \dim \ker A + \dim \ker A^*$. By rank-nullity, $\dim \ker A = \dim C_{\text{even}} - \mathrm{rank} A$ and $\dim \ker A^* = \dim C_{\text{odd}} - \mathrm{rank} A$. Summing gives the claim. $\square$

Proposition 4.2. $\mathrm{index}(A) := \dim \ker A - \dim \ker A^* = \dim C_{\text{even}} - \dim C_{\text{odd}} = M(n)$. Consequently $\dim \ker H_n \geq |M(n)|$.

Proof. Rank-nullity. The signed sum is the reduced Euler characteristic of Δ_n up to sign. The bound follows from $\dim \ker A + \dim \ker A^* \geq |\dim \ker A - \dim \ker A^*|$. $\square$

4.1 Maximal elements and the kernel of $A_{\uparrow}$

Definition 4.3. A squarefree integer $m \leq n$ is maximal if $pm > n$ for every prime p with $p \nmid m$; equivalently, if every prime $p \leq n/m$ divides m ; equivalently, if $T_p(e_m) = 0$ for every prime $p \leq n$.

Proposition 4.4. Assume $c_p \neq 0$ for every prime $p \leq n$.

- (a) For $m \leq n$ squarefree with $\Omega(m)$ even, $e_m \in \ker A_{\uparrow}|_{C_{\text{even}}}$ if and only if m is maximal. Consequently $\mathcal{M} := \mathrm{span}\{e_m : m \text{ maximal, } \Omega(m) \equiv 0 \pmod{2}\} \subset \ker A_{\uparrow}|_{C_{\text{even}}}$.

- (b) The operator A does not annihilate any basis vector of C_{even} : $Ae_m \neq 0$ for every squarefree $m \leq n$ with $\Omega(m)$ even and $n \geq 2$.

Proof. (a) $A_{\uparrow}e_m = \sum_{p \nmid m, pm \leq n} c_p e_{pm}$, a sum of distinct basis vectors with non-zero coefficients; it vanishes iff the index set is empty, i.e. iff every prime $p \leq n/m$ divides m .

(b) Let $m \leq n$ be squarefree with $\Omega(m)$ even. If $m > 1$, it has at least one prime divisor p , and $A_{\downarrow}e_m = \sum_{p \mid m} c_p e_{m/p} \neq 0$. Hence $Ae_m \neq 0$. If $m = 1$ (so $\Omega(m) = 0$, even), $A_{\downarrow}e_1 = 0$ and $A_{\uparrow}e_1 = \sum_{p \leq n} c_p e_p \neq 0$ for $n \geq 2$. $\square$

Remark 4.5 (Correction of the fourth revision). The fourth revision asserted $\ker A = \{0\}$ for $n \geq 2$. This is *false*: Proposition 4.4(b) only shows that $Ae_m \neq 0$ for every *basis* vector of C_{even} , which does not exclude linear dependencies among the images. Direct computation at $n = 10^4$ gives $\text{rank}(A) = 2141$ and $\dim C_{\text{even}} = (d_n + M(n))/2 = (6083 - 23)/2 = 3030$, hence $\dim \ker A = 3030 - 2141 = 889 \neq 0$. What is correct is that $\ker A$ is strictly smaller than $\ker A_{\uparrow}|_{C_{\text{even}}}$, since $A_{\downarrow}$ contributes additional relations.

Remark 4.6 (Nomenclature). The kernel decomposition that the earlier revisions studied is that of $A_{\uparrow}$, not of A . We write

$$\ker A_{\uparrow}|_{C_{\text{even}}} = \mathcal{M} \oplus R_n^{\uparrow},$$

where $R_n^{\uparrow} = \ker A_{\uparrow}|_{C_{\text{even}}} \ominus \mathcal{M}$ is the residual subspace of relations. The density of $\mathcal{M}$ in C_{even} is $c_1 = 0.56755 \dots$ (Theorem 4.11).

4.2 Exact formula for the maximal elements

Theorem 4.7. Let $p_0 := 1 < p_1 < p_2 < \dots$ be the primes and $P_k := p_1 p_2 \dots p_k$ the k -th primorial ($P_0 = 1$). A squarefree integer $m \leq n$ is maximal if and only if there exists $k \geq 0$ and a squarefree m' , coprime to P_k , such that

$$m = P_k m', \quad \frac{n}{p_{k+1} P_k} < m' \leq \frac{n}{p_k P_k}. \quad (11)$$

Moreover $\Omega(m) = k + \Omega(m') = k + \omega(m')$.

Proof. If m satisfies (11), then $n/m \in [p_k, p_{k+1})$, so the primes $p \leq n/m$ are exactly $p_1, \dots, p_k$. By Definition 4.3, m is maximal iff every prime $p \leq n/m$ divides m , i.e. $p_1 \dots p_k = P_k \mid m$. Since m is squarefree and $m = P_k m'$ with $\gcd(m', P_k) = 1$, this is equivalent to (11). Conversely, any maximal m determines k via $n/m \in [p_k, p_{k+1})$, and $P_k \mid m$ forces $m' = m/P_k$. $\square$

Definition 4.8. The signed count of maximal elements of even weight, and the corresponding signed count on the odd side, are

$$M_{\mathcal{M}}(n) := \#\mathcal{M} - \#\mathcal{M}^T, \quad M_{\mathcal{M}}^T(n) := \#\mathcal{M}^T - \#\mathcal{M}.$$

Note $M_{\mathcal{M}}^T(n) = -M_{\mathcal{M}}(n)$.

Theorem 4.9 (Exact formula for $M_{\mathcal{M}}(n)$). For every $n \geq 1$,

$$M_{\mathcal{M}}(n) = \sum_{k \geq 0} (-1)^k \left[T_k \left(\left\lfloor \frac{n}{p_k P_k} \right\rfloor \right) - T_k \left(\left\lfloor \frac{n}{p_{k+1} P_k} \right\rfloor \right) \right], \quad (12)$$

where

$$T_k(X) := \sum_{\substack{d \leq X \\ d \text{ } P_k\text{-smooth}}} M \left(\left\lfloor \frac{X}{d} \right\rfloor \right) \quad (13)$$

and M is the Mertens function. Here “ d P_k -smooth” means every prime factor of d belongs to $\{p_1, \dots, p_k\}$; for $k = 0$, $T_0(X) = M(X)$.

Proof. By Theorem 4.7, the maximal elements of weight $2j$ (even) are exactly those $m = P_k m'$ with $\Omega(m')$ of parity $2j - k$. The signed count is

$$M_{\mathcal{M}}(n) = \sum_{k \geq 0} (-1)^k \sum_{\substack{m' \in (L_k, U_k] \\ \gcd(m', P_k) = 1 \\ \mu^2(m') = 1}} (-1)^{\Omega(m')},$$

where $L_k := \lfloor n/(p_{k+1}P_k) \rfloor$ and $U_k := \lfloor n/(p_k P_k) \rfloor$. The inner sum counts squarefree m' coprime to P_k with sign $(-1)^{\Omega(m')} = \mu(m')$, in the interval $(L_k, U_k]$. Summing over $m' \leq X$ and using the convolution $\mathbf{1}_{\gcd(\cdot, P_k)=1} \mu = \mu * \mathbf{1}_{P_k\text{-smooth}}$ gives $T_k(X)$. $\square$

Numerical Observation 4.10 (Numerical verification of (12)). *For $n \in \{10^4, 2 \cdot 10^4, 3 \cdot 10^4, 5 \cdot 10^4, 7 \cdot 10^4, 10^5, 2 \cdot 10^5\}$ we have computed $M(n)$ and $M_{\mathcal{M}}(n)$ by independent sieves; the results are reported in Table 1. The formula (12) reproduces $M_{\mathcal{M}}(n)$ exactly. An independent enumeration of the maximal elements for $n \leq 10^3$ confirms the formula on small values.*

Table 1: Values of $M(n)$, $M_{\mathcal{M}}(n)$, and the residual $\dim R_n^\uparrow - \dim R_n^{\uparrow, T} = M(n) - M_{\mathcal{M}}(n)$.

n	$M(n)$	$M_{\mathcal{M}}(n)$	$ M_{\mathcal{M}} /\sqrt{n}$	residual
10^4	-23	-32	0.320	9
$2 \cdot 10^4$	26	53	0.375	-27
$3 \cdot 10^4$	18	10	0.058	8
$5 \cdot 10^4$	23	36	0.161	-13
$7 \cdot 10^4$	74	106	0.401	-32
10^5	-48	-77	0.244	29
$2 \cdot 10^5$	-1	23	0.051	-24

4.3 The density of the maximal elements

Theorem 4.11. *Let $p_1 < p_2 < \dots$ be the primes and set $p_0 := 1$. Then*

$$\lim_{n \rightarrow \infty} \frac{\#\{m \leq n : \mu^2(m) = 1, m \text{ maximal}, \Omega(m) \equiv 0 \pmod{2}\}}{d_{\text{even}}(n)} = c_1, \quad (14)$$

where

$$c_1 := \sum_{k \geq 0} \left(\frac{1}{p_k} - \frac{1}{p_{k+1}} \right) \prod_{j=1}^k \frac{1}{p_j + 1} = 0.5675527193 \dots \quad (15)$$

The same constant governs the odd side.

Proof. By Theorem 4.7, on the interval $(n/(p_{k+1}P_k), n/(p_k P_k)]$ the maximal elements of weight $2k$ correspond to squarefree m' coprime to P_k in the same interval. The squarefree integers coprime to P_k up to x have asymptotic count

$$\frac{6}{\pi^2} x \prod_{p|P_k} \frac{p}{p+1} + O(\sqrt{x}),$$

as follows from the local densities $1-p^{-1}$ at primes $p \mid P_k$ and $1-p^{-2}$ at primes $p \nmid P_k$. Restricting to those m' with $\Omega(m') \equiv k \pmod{2}$ contributes a factor $\frac{1}{2}$ asymptotically. Summing over k and dividing by $d_{\text{even}}(n) \sim (3/\pi^2)n$ gives c_1 . $\square$

Numerical Observation 4.12 (Numerical verification of Theorem 4.11). *At $n = 10^4$, the empirical density of the even-weight maximal elements in C_{even} is $1709/3030 = 0.5640$; at $n = 3 \cdot 10^4$ it is $5184/9130 = 0.5678$. The odd side behaves symmetrically ($1741/3053 = 0.5703$ and $5174/9112 = 0.5678$ respectively). These values bracket the asymptotic $c_1 = 0.5675527 \dots$ with an error of order 10^{-3} at $n = 3 \cdot 10^4$.*

4.4 Kernel density and the spectral obstruction

Definition 4.13. $c_*(n) := \text{rank}(A_{\uparrow,n})/d_n$; $c_{\uparrow}(n) := \text{rank}(A_{\uparrow,n})/d_{\text{even}}(n)$; $c(n) := \text{rank}(A_n)/d_n$.

At $n = 10^4$ (rank computed modulo $\{3, 5, 7, 11, 13\}$, see Appendix B):

$$\text{rank}(A_{\uparrow}) = 1125, \quad \text{rank}(A) = 2141, \quad (16)$$

$$c_* = \frac{1125}{6083} = 0.185, \quad c_{\uparrow} = \frac{1125}{3030} = 0.371, \quad c = \frac{2141}{6083} = 0.352. \quad (17)$$

Proposition 4.14. *The kernel of H_n has density*

$$\frac{\dim \ker H_n}{d_n} = 1 - 2c(n) \approx 1 - 2 \cdot 0.352 = 0.296 \quad (18)$$

at $n = 10^4$.

Remark 4.15 (Subadditivity of the rank). We do *not* claim $\text{rank}(A) = \text{rank}(A_{\uparrow}) + \text{rank}(A_{\downarrow})$. The two parts $A_{\uparrow}$ and $A_{\downarrow}$ have images with overlapping supports. The rank is subadditive,

$$\text{rank}(A) \leq \text{rank}(A_{\uparrow}) + \text{rank}(A_{\downarrow}).$$

Assuming the symmetry $\text{rank}(A_{\uparrow}) = \text{rank}(A_{\downarrow})$ at $n = 10^4$, the additivity defect $\delta(n) = \text{rank}(A_{\uparrow}) + \text{rank}(A_{\downarrow}) - \text{rank}(A)$ equals $\delta(10^4) = 109$. Whether $\delta(n)/n \rightarrow 0$ or $\delta(n)/n \rightarrow c > 0$ is left open.

5 Exact kernel decomposition and the critical exponent

The decomposition $M(n) = M_{\mathcal{M}}(n) + \mathcal{R}(n)$ of §4, with

$$\mathcal{R}(n) := M(n) - M_{\mathcal{M}}(n) = \dim R_n^{\uparrow} - \dim R_n^{\uparrow,T}, \quad (19)$$

acquires a finer structure when expressed in terms of the zeros of ζ . In this section we introduce the *exact kernel* of $\mathcal{R}(n)$, prove that it converges to the asymptotic kernel studied in the fifth revision, and investigate numerically whether the convergence follows a universal power law with exponent $\alpha = 1/2$.

5.1 The exact kernel

Let $L_k := \lfloor n/(p_{k+1}P_k) \rfloor$, $U_k := \lfloor n/(p_k P_k) \rfloor$, and define

$$S_k(X; \rho) := \sum_{\substack{d \leq X \\ d \text{ } P_k\text{-smooth}}} d^{-\rho}, \quad N_k(X) := \#\{d \leq X : d \text{ } P_k\text{-smooth}\}. \quad (20)$$

Proposition 5.1 (Exact kernel representation). *For every $n \geq 1$,*

$$\mathcal{R}(n) = C_R(n) + \sum_{\rho} \frac{c_R^{\text{exact}}(\rho, n)}{\rho \zeta'(\rho)} n^{\rho} + (\text{conjugate}) + O(n^{-1/2}), \quad (21)$$

where the sum is over the nontrivial zeros $\rho = \frac{1}{2} + i\gamma$ of ζ , and

$$C_R(n) := -2 + 2 \sum_{k \geq 0} (-1)^k [N_k(U_k) - N_k(L_k)], \quad (22)$$

$$c_R^{\text{exact}}(\rho, n) := 1 - \frac{1}{n^{\rho}} \sum_{k \geq 0} (-1)^k [S_k(U_k; \rho) U_k^{\rho} - S_k(L_k; \rho) L_k^{\rho}]. \quad (23)$$

The representation is understood as a formal expansion: the series over the zeros converges only conditionally, and the remainder is uniform in n and in ρ on compact subsets of $\{\Re \rho > 0\}$.

Proof. Substitute the explicit formula $M(x) = -2 + \sum_{\rho} x^{\rho}/(\rho\zeta'(\rho)) + (\text{trivial zeros}) + O(x^{-1})$ into the exact formula (12) for $M_{\mathcal{M}}(n)$ and collect terms. The constant $C_R(n)$ collects the contribution of the pole term -2 multiplied by the total number of summands (with signs). The kernel $c_R^{\text{exact}}(\rho, n)$ collects the coefficient of $n^{\rho}/(\rho\zeta'(\rho))$ in the difference $M(n) - M_{\mathcal{M}}(n)$. $\square$

Proposition 5.2 (Asymptotic kernel). *For every fixed ρ with $\Re \rho > 0$, as $n \rightarrow \infty$,*

$$c_R^{\text{exact}}(\rho, n) \rightarrow c_R^{\text{asym}}(\rho) := 2^{-\rho} - \sum_{k \geq 1} (-1)^k \frac{p_k^{-\rho} - p_{k+1}^{-\rho}}{\prod_{j=1}^k (p_j^{\rho} - 1)}. \quad (24)$$

Proof. For fixed k , as $n \rightarrow \infty$ both $U_k, L_k \rightarrow \infty$ and $S_k(X; \rho) \rightarrow \prod_{j=1}^k (1 - p_j^{-\rho})^{-1}$. Moreover $U_k^{\rho}/n^{\rho} \rightarrow (p_k P_k)^{-\rho}$ and $L_k^{\rho}/n^{\rho} \rightarrow (p_{k+1} P_k)^{-\rho}$. Combining, the sum in (23) becomes

$$\frac{1}{n^{\rho}} \sum_{k \geq 0} (-1)^k [S_k(U_k; \rho) U_k^{\rho} - S_k(L_k; \rho) L_k^{\rho}] \rightarrow \sum_{k \geq 0} (-1)^k \frac{p_k^{-\rho} - p_{k+1}^{-\rho}}{\prod_{j=1}^k (p_j^{\rho} - 1)},$$

with the convention $p_0 = 1$ and the product empty for $k = 0$. Hence $c_R^{\text{exact}}(\rho, n) \rightarrow 1 - (1 - 2^{-\rho} + \sum_{k \geq 1} (\dots)) = 2^{-\rho} - \sum_{k \geq 1} (\dots)$. $\square$

Remark 5.3. A direct evaluation at $\rho = 2$ gives $c_R^{\text{asym}}(2) = 0.2934\dots$, and at $\rho = 1/2 + i\gamma_1$ (with $\gamma_1 = 14.134725$) gives $c_R^{\text{asym}}(\rho_1) = -0.6249 + 0.1382i$. Both values are reproduced by the exact kernel $c_R^{\text{exact}}(\rho, n)$ at large n .

5.2 Numerical observation: power-law convergence

We evaluated $c_R^{\text{exact}}(\rho_k, n)$ for the first 175 zeros $\rho_k = 1/2 + i\gamma_k$ of ζ and for n in a logarithmic grid of 80 points from 10^2 to 10^8 , with the asymptotic kernel $c_R^{\text{asym}}(\rho_k)$ precomputed from (24) truncated at $K = 30$.

Numerical Observation 5.4 (Power-law convergence). *For every $k \in \{0, \dots, 174\}$ the quantity*

$$\Delta_k(n) := |c_R^{\text{exact}}(\rho_k, n) - c_R^{\text{asym}}(\rho_k)| \quad (25)$$

is well fitted by a power law $\Delta_k(n) \sim n^{-\alpha_k}$ over six decades of n , with individual fits $R^2 \geq 0.83$. The exponent distribution has sample mean and sample standard deviation

$$\bar{\alpha} = 0.4918, \quad s_{\alpha} = 0.0506. \quad (26)$$

The minimum is $\min_k \alpha_k = 0.33$, the maximum $\max_k \alpha_k = 0.60$, the median 0.487. With 175 samples, the naive standard error of the mean is $s_{\alpha}/\sqrt{175} \approx 0.0038$. A direct reproduction of the pipeline in the sixth revision, using the same 175 zeros and the same 80-point grid in n but re-running the full computation, gives $\bar{\alpha} = 0.4918$ and $s_{\alpha} = 0.0506$, consistent with the fifth-revision values (0.4888, 0.0498) to within the standard error. The small difference 0.003 in the mean is within 1 SEM and is attributed to numerical noise (double-precision accumulation in the P_k -smooth sums).

Numerical Observation 5.5 (Absence of spectral structure). *The exponent α_k is uncorrelated with the height γ_k of the zero and with the index k :*

$$\text{corr}(\alpha_k, \log \gamma_k) = -0.011, \quad \text{corr}(\alpha_k, k) = -0.067. \quad (27)$$

The empirical distribution of $\{\alpha_k\}$ is approximately Gaussian. Seven of the 175 values exceed two standard deviations from the mean, consistent with the 8.75 expected from a Gaussian distribution. There is no evidence of clustering or of a systematic dependence on γ_k .

Numerical Observation 5.6 (Block-bootstrap error of the mean). *Because the 175 fits share the same grid in n , the sample values α_k are not independent and the naive standard error 0.0038 is a lower bound on the true uncertainty. A block bootstrap with $B = 8$ contiguous blocks of 10 points each, 1000 resamples, gives*

$$\text{SE}_{\text{boot}} = 0.0078, \quad (28)$$

a factor 2.05 larger than the naive value 0.0038. The 95% bootstrap confidence interval is $[0.4794, 0.5097]$, and the deviation of the sample mean from $1/2$ is

$$\frac{\bar{\alpha} - 1/2}{\text{SE}_{\text{boot}}} = -1.04, \quad (29)$$

i.e. approximately 1.0 bootstrap standard error. The hypothesis $\alpha = 1/2$ is therefore not rejected at any conventional significance level. The choice of B is not critical: $B \in \{4, 8, 10, 16, 20\}$ gives $\text{SE}_{\text{boot}} \in [0.0058, 0.0104]$, all consistent with the range of the naive value.

Numerical Observation 5.7 (Insensitivity to the truncation parameter K). *The asymptotic kernel $c_R^{\text{asym}}(\rho)$ is defined by the series (24), truncated at $k = K$. We tested $K \in \{10, 20, 30, 50\}$ and found that all quantities reported in Observations 5.4–5.6 ($\bar{\alpha}$, s_α , SE_{boot} , the bootstrap z -statistic, and the log-correction coefficient $\bar{\beta}$ of Remark 5.8) are identical to machine precision. The reason is elementary: the k -th term of (24) is bounded by $p_k^{-1/2}/\sqrt{P_k}$, which for $k = 15$ is already below 10^{-10} and for $k = 20$ below 10^{-14} , i.e. at the level of double-precision round-off. Consequently:*

- (i) *The truncation at $K = 30$ used in the paper is very conservative; $K = 10$ would already give accuracy $\sim 10^{-6}$, far below the observed deviations.*
- (ii) *The truncation is not a source of systematic error in $\bar{\alpha}$, in SE_{boot} , or in the residual of the bias check.*
- (iii) *The choice of K in the numerical pipeline is free within $[10, 100]$; no tuning of K is needed to reproduce the results.*

The same insensitivity holds for the sum over k in (23), because the loop naturally terminates when $U_k < 1$, which for $n \leq 10^8$ occurs at $k \approx 8$.

Remark 5.8 (Statistical and systematic status of the exponent). Observations 5.4–5.7 settle the statistical question raised in the fifth revision. We record the three findings explicitly.

(1) Bootstrap error. With the block-bootstrap error $\text{SE}_{\text{boot}} = 0.0078$, the sample mean $\bar{\alpha} = 0.4918$ lies ≈ 1.0 bootstrap standard error below $1/2$. The naive value 0.0038 underestimates the true uncertainty by a factor of two because the 175 fits share the same grid in n . The hypothesis $\alpha = 1/2$ is therefore *consistent* with the data, and the earlier statement of “approximately 3 standard errors” reflected the naive SEM, not a genuine power-law deviation.

(2) Log-correction rejected. The logarithmic-correction hypothesis $\Delta_k(n) \propto n^{-1/2}(\log n)^{-\beta_k}$ is not supported by the data. A direct fit of $\log(\Delta_k(n) \sqrt{n})$ versus $\log \log n$ gives

$$\bar{\beta} = -0.081, \quad \overline{\beta_k / \text{SE}(\beta_k)} = -0.53,$$

with the t -statistics distributed symmetrically (35 values above +2, 44 below −2 out of 175). The predicted deficit $\bar{\beta} \cdot \langle 1/\log n \rangle = -0.0081$ has the *wrong sign*, while the observed deficit is +0.0082; the residual +0.016 is twice the deficit itself. The log-correction family is therefore excluded.

(3) Truncation not systematic. By Observation 5.7, the truncation parameter K has no effect on any of these quantities for $K \in \{10, 20, 30, 50\}$. The truncation is therefore not the

origin of the deficit. If the deficit is real, its source lies outside both the logarithmic-correction family and the truncation of c_R^{asym} .

A definitive resolution would require: (i) a uniform-in- K control of the truncation (already achieved here for K up to 50); (ii) an explicit remainder bound in Proposition 5.2; (iii) an extension of the n -grid up to 10^{10} or beyond, to test the stability of $\bar{\alpha}$ against grid refinements. We leave (ii) and (iii) to future work.

Remark 5.9 (Why the series over zeros is only conditionally convergent). The representation (21) is only conditionally convergent: the terms $c_R^{\text{exact}}(\rho_k, n)/(\rho_k \zeta'(\rho_k))$ decay like $1/(|\gamma_k| \log |\gamma_k|)$ in modulus, and the sum depends on the order of summation. A numerical evaluation at $n = 10^4$ with 175 zeros gives $\mathcal{R}(10^4) \approx 2.5$ against the true value 9, with an error of -6.5 ; adding more zeros does not decrease the error monotonically (the sequence is not absolutely convergent). This is a well-known feature of the explicit formula and does not invalidate the representation (21); it merely limits its numerical utility to n not too large and to a truncation chosen to minimise the residual oscillation.

Remark 5.10 (Connection to the minimal-variance program). If $\alpha = 1/2$ holds exactly, it provides a natural bridge to the minimal-variance characterisation of RH: the finite- n correction to the kernel would then decay precisely at the critical scale where any deviation of a single zero from the critical line would produce a detectable anomaly. This suggests that the exact kernel $c_R^{\text{exact}}(\rho, n)$ is a natural *defect observable* for the minimal-variance condition. A quantitative version of this connection, however, is not established in this paper.

6 No new characterisation of RH from the decomposition

The kernel decomposition $M(n) = M_{\mathcal{M}}(n) + \mathcal{R}(n)$ of Proposition 4.2 and Theorem 4.9 might suggest a new characterisation of RH: if one could show unconditionally that $M_{\mathcal{M}}(n) = O(n^\alpha)$ with $\alpha < 1/2$, then RH would be equivalent to a bound on the residual. We explain why this approach cannot succeed.

Proposition 6.1 (Obstruction to the algebraic approach). *Let $M_{\mathcal{M}}(n)$ be the signed count of maximal elements defined in (12). Then:*

- (i) *Assume RH. Then the leading term of $M_{\mathcal{M}}(n)$ is $M(n) - M(n/2)$, and the root mean square of $M(n) - M(n/2)$ is $\Theta(\sqrt{n})$ in the sense of the second moment of the explicit formula. The $k \geq 1$ contributions in (12) are of lower order $\Theta(\sqrt{n/P_k})$.*
- (ii) *An unconditional bound of the form $M_{\mathcal{M}}(n) = O(\sqrt{n})$ would imply $M(n) = O(\sqrt{n} \log n)$, hence RH (with a log factor).*

Proof. (i) From (12), the $k = 0$ term is $T_0(n) - T_0(n/2) = M(n) - M(n/2)$. The $k \geq 1$ terms have arguments $n/(p_k P_k) \leq n/12$ and contribute $O(\sqrt{n/P_k})$. Under RH the root mean square of $M(x)$ over $x \leq X$ is $\Theta(\sqrt{X})$; the two values $M(n)$ and $M(n/2)$ are determined by disjoint ranges of the explicit formula at leading order, so the root mean square of $M(n) - M(n/2)$ is $\Theta(\sqrt{n})$.

(ii) Summing the bound $M(n) - M(n/2) = O(\sqrt{n})$ over dyadic intervals $[n/2^{j+1}, n/2^j]$ gives $M(n) = O(\sqrt{n} \log n)$, which is RH up to a logarithmic factor. $\square$

Remark 6.2. The numerical observation $\alpha_k \approx 1/2$ of §5 is consistent with the obstruction above: the finite- n correction to the kernel is of the same order as the leading term of $M_{\mathcal{M}}(n)$, so no new characterisation of RH can be extracted from the decomposition alone.

7 Structural obstruction to the spectral modification

The fourth revision conjectured (Method 6.1) that the spectral modification

$$\tilde{H}_n := (I - P_n) H_n (I - P_n), \quad (30)$$

where P_n is the orthogonal projection onto $\ker H_n$, converges in distribution to the empirical measure of the imaginary parts of the nontrivial zeros of ζ . We prove that this conjecture is *false*, for two independent rigorous reasons and one heuristic one.

7.1 Obstruction I: compact support

Proposition 7.1. *Let $s > 1/2$, $c_p = p^{-s}$, and $H_n = \sum_{p \leq n} c_p (T_p + T_p^*)$. Then*

$$\|H_n\|_{\text{op}} \leq 2P(s), \quad P(s) := \sum_p p^{-s} < \infty,$$

uniformly in n .

Proof. For every $v \in C_*(\Delta_n; \mathbb{R})$,

$$|\langle v, H_n v \rangle| \leq \sum_{p \leq n} c_p (|\langle v, T_p v \rangle| + |\langle v, T_p^* v \rangle|) \leq 2P(s) \|v\|^2,$$

using $\|T_p\|_{\text{op}} = \|T_p^*\|_{\text{op}} = 1$ from Lemma 2.1. Since H_n is self-adjoint, the operator norm equals $\sup_{\|v\|=1} |\langle v, H_n v \rangle|$. $\square$

Corollary 7.2. *The measure $\mu_n = (d_n - \dim \ker H_n)^{-1} \sum_{\lambda \in \text{spec}(\tilde{H}_n)} \delta_\lambda$ has support in $[-2P(s), 2P(s)]$ for every n . Hence no rescaling $s_n \rightarrow \infty$ can make $s_n \mu_n$ converge to the zero measure ν_T , which has unbounded support.*

Remark 7.3. For $0 < s \leq 1/2$ the operator-norm bound fails, but $\|H_n\|_{\text{op}} \leq 2 \sum_{p \leq n} p^{-s} \sim 2n^{1-s}/((1-s) \log n)$. The support of μ_n then grows as $n^{1-s}/\log n$, while the number of eigenvalues is $d_n \sim n$. The density of μ_n on its support decays as $n^{-s}/\log n \rightarrow 0$, whereas the density of ν_T on $[0, T]$ grows as $(1/2\pi) \log(T/2\pi)$. The two densities are asymptotically different for every $s \in (0, 1/2]$. This is a second, dimension-theoretic, obstruction to the conjecture.

7.2 Obstruction II: prime powers are missing

Proposition 7.4. *For every $s > 0$,*

$$\lim_{n \rightarrow \infty} \frac{\text{Tr}(H_n^2)}{n} = \frac{12}{\pi^2} \sum_p \frac{p^{-2s}}{p+1} < \infty.$$

The corresponding second spectral moment of the prime term in the Guinand–Weil formula, after Gaussian regularisation with parameter σ and passage to the limit $\sigma \rightarrow 0$, diverges as $1/(8\sigma^3)$.

Proof. The first assertion is Theorem 3.3. For the second, apply the explicit formula with $h(r) = r^2 e^{-\sigma^2 r^2}$. The corresponding g is

$$g(x) = \frac{1}{4\sqrt{\pi}\sigma^3} \left(1 - \frac{x^2}{2\sigma^2}\right) e^{-x^2/(4\sigma^2)}.$$

The prime term in $-C''_\sigma(0)$ gives

$$\frac{1}{8\sigma^5\sqrt{\pi}} \sum_{n \geq 2} \frac{\Lambda(n)}{\sqrt{n}} (\log n)^2 e^{-(\log n)^2/(4\sigma^2)} \sim \frac{1}{8\sigma^3}.$$

The sum contains the contribution of all prime powers p^j , $j \geq 1$, and is dominated by the terms $n = p^j$ with $\log n \lesssim \sigma$. In contrast, H_n contains only the terms $j = 1$. Hence the two second moments disagree at order $1/\sigma^3$. $\square$

Corollary 7.5. H_n cannot reproduce the Guinand–Weil formula, because it misses all prime powers p^j with $j \geq 2$. The mismatch already appears at the level of the second moment.

7.3 Obstruction III: density heuristic

Remark 7.6. A third, *heuristic* obstruction is often quoted: the spectral measure μ_n of H_n on its support $[-2P(s), 2P(s)]$ (for $s > 1/2$) has eigenvalues distributed with a density that is *non-logarithmic* in the interior of its support, whereas the zero measure ν_T has density $(1/2\pi)\log(T/2\pi) + O(1/T)$ on $[0, T]$. This suggests that the two measures are structurally different.

We stress that this third obstruction is not proved rigorously here. The representation of the eigenvalues of H_n as $\pm \sum_{p \leq n} c_p \varepsilon_p$ with $\varepsilon_p \in \{-1, 0, +1\}$ is only heuristic, because the operators $\gamma_p = T_p + T_p^*$ do not commute in the truncated complex (Remark 2.4); hence a simultaneous diagonalisation is not available and the distribution of $\sum c_p \varepsilon_p$ need not coincide with the spectral measure of H_n . For the purpose of ruling out the conjecture of the fourth revision, Obstructions I and II suffice; the third is recorded as a plausible additional feature.

7.4 Synthesis

Table 2: The obstructions to the convergence of $\tilde{H}_n$ to the zeros of ζ .

Obstruction	Nature	Conclusion
I (support)	spectral (proved)	μ_n is compactly supported; ν_T is not
II (prime powers)	algebraic (proved)	H_n misses p^j , $j \geq 2$
III (density)	analytic (heuristic)	μ_n has non-logarithmic density; ν_T logarithmic

The first two obstructions are independent and rigorously proved: removing one does not salvage the conjecture. We also correct an error of the fourth revision: the operators $\gamma_p = T_p + T_p^*$ do *not* commute in the truncated complex, because $T_p T_q^* \neq T_q^* T_p$ at the boundary, as noted in Remark 2.4.

Remark 7.7. A genuine open question is the following: *what is the weak limit of μ_n for $s = 1/2$?* It is a compactly supported measure determined by the combinatorics of squarefree integers, with moments computable from the trace formula of Theorem 3.2. Computing its Hausdorff dimension is a well-posed problem.

8 The Toeplitz pencil for the explicit formula

8.1 The Guinand–Weil formula and the prefactor

Theorem 8.1 (Guinand–Weil). *Let h be an even entire function of sufficient decay in the strip $|\operatorname{Im} r| \leq \frac{1}{2} + \epsilon$. Then*

$$\sum_{\rho} h(\gamma_{\rho}) = h(i/2) + h(-i/2) + \frac{1}{2\pi} \int_{-\infty}^{\infty} h(r) \Omega(r) dr - \sum_{n \geq 2} \frac{\Lambda(n)}{\sqrt{n}} [g(\log n) + g(-\log n)], \quad (31)$$

where $\Omega(r) = \operatorname{Re} \psi(\frac{1}{4} + \frac{ir}{2}) - \log \pi$ and $g(x) = \frac{1}{2\pi} \int_{\mathbb{R}} h(r) e^{-irx} dr$.

For $\sigma > 0$ define $C_{\sigma}(\Delta) = \sum_{\gamma > 0} \cos(\gamma \Delta) e^{-\sigma^2 \gamma^2}$. Applying the Guinand–Weil formula with

$$h(r) = \cos(\Delta r)e^{-\sigma^2 r^2}:$$

$$\begin{aligned} C_\sigma(\Delta) = & \cosh(\Delta/2)e^{\sigma^2/4} + \frac{1}{2\pi} \int_0^\infty \cos(r\Delta)e^{-\sigma^2 r^2} \Omega(r) dr \\ & - \frac{1}{4\sigma\sqrt{\pi}} \sum_{n \geq 2} \frac{\Lambda(n)}{\sqrt{n}} \left[e^{-(\Delta - \log n)^2/(4\sigma^2)} + e^{-(\Delta + \log n)^2/(4\sigma^2)} \right]. \end{aligned} \quad (32)$$

Proposition 8.2. *The prefactor of the prime-power sum in (32) is $1/(4\sigma\sqrt{\pi})$.*

Proof. With $g(x) = (1/2\pi) \int_{\mathbb{R}} \cos(\Delta r)e^{-\sigma^2 r^2} e^{-irx} dr$ and $\int_{\mathbb{R}} e^{-\sigma^2 r^2} e^{-irx} dr = (\sqrt{\pi}/\sigma)e^{-x^2/(4\sigma^2)}$, direct computation gives

$$g(x) = \frac{1}{4\sigma\sqrt{\pi}} \left[e^{-(x-\Delta)^2/(4\sigma^2)} + e^{-(x+\Delta)^2/(4\sigma^2)} \right].$$

The prime-power term is $-\sum_{n \geq 2} (\Lambda(n)/\sqrt{n})g(\log n)$, whence the prefactor. $\square$

Numerical Observation 8.3 (Numerical verification). *We compared (32) with the direct sum over the first 60 known zeros γ_k of ζ at $\sigma = 5 \times 10^{-2}, 10^{-1}$. With prefactor $1/(4\sigma\sqrt{\pi})$ the relative error is at the level of double-precision round-off ($\lesssim 10^{-13}$); with the prefactor $1/(8\pi\sigma\sqrt{\pi})$ of the first revision it is of order unity.*

8.2 Exact-recovery theorem for the pencil

Fix $\sigma > 0$, $\delta > 0$, $K \geq 1$. Define the Toeplitz matrices $G_N, S_N \in \mathbb{R}^{K \times K}$ by

$$(G_N)_{jk} = C_N((j-k)\delta), \quad (S_N)_{jk} = -C'_N((j-k)\delta). \quad (33)$$

Theorem 8.4. *Let $a_1, \dots, a_M \in \mathbb{R} \setminus \{0\}$ and $0 < \gamma_1 < \dots < \gamma_M < \pi/\delta$. Define $C(\Delta) = \sum_{k=1}^M a_k \cos(\gamma_k \Delta)$. For $K \geq 2M$ let G, S be the Toeplitz matrices defined as above. Then the generalised eigenvalues λ of $iSv = \lambda Gv$ for which $Gv \neq 0$ are exactly $\pm\gamma_1, \dots, \pm\gamma_M$.*

The proof is in Appendix A. Numerical verification on synthetic data with $M = 5$ frequencies in $\{10, 30, 50, 70, 90\}$, $K = 2000$, $\delta = 5 \times 10^{-4}$ gives the ten finite generalised eigenvalues coinciding with $\pm\gamma_k$ to absolute error below 10^{-13} .

8.3 Hankel–ESPRIT pipeline and bottlenecks

The numerical pipeline is: form the Hankel matrix $H_{i,j} = C_\sigma((i+j)\delta)$, compute its SVD, apply ESPRIT to the left singular subspace. The three structural bottlenecks are:

- (i) **Aliasing.** Nyquist frequency $\gamma_c = \pi/\delta$; safe rule $\pi\sigma/\delta \geq 4$.
- (ii) **Effective rank of the Hankel matrix.** $\approx 2M_{\text{eff}}$, with $M_{\text{eff}} \approx N(1/\sigma) \sim (2\pi\sigma)^{-1} \log(1/(2\pi e\sigma))$ as $\sigma \rightarrow 0$.
- (iii) **Truncation of the prime-power sum.** Requires prime powers up to $N_{\text{max}} \sim e^{\Delta_{\text{max}} + c\sigma}$.

9 Rigorous recovery bounds

We now prove the truncation and stability bounds announced in the abstract. Throughout, we fix $\sigma > 0$, $\Delta_{\text{max}} > 0$, and consider

$$C_\sigma(\Delta) = \sum_{\gamma > 0} e^{-\sigma^2 \gamma^2} \cos(\gamma \Delta), \quad 0 \leq \Delta \leq \Delta_{\text{max}},$$

given by (32). We denote by $C_{\text{exp}}, C_{\text{int}}, C_{\text{prime}}$ the three terms on the right-hand side of (32), and by $C_{\text{int}}^{(R)}, C_{\text{prime}}^{(N)}$ their truncations at $r = R$ and $n = N$.

9.1 Truncation of the archimedean integral

Lemma 9.1. *There exists an absolute constant C_1 such that, for every $R > 0$ and every $\Delta \in [0, \Delta_{\max}]$,*

$$|C_{\text{int}}(\Delta) - C_{\text{int}}^{(R)}(\Delta)| \leq C_1 \frac{\log(1/\sigma)}{\sigma} e^{-R^2\sigma^2}.$$

Proof. For $|z| \rightarrow \infty$ with $|\arg z| < \pi$, $\psi(z) = \log z - \frac{1}{2z} + O(|z|^{-2})$. With $z = \frac{1}{4} + \frac{ir}{2}$ we get $\Omega(r) = \log(r/2\pi) + O(1/r)$. Hence

$$|C_{\text{int}} - C_{\text{int}}^{(R)}| \leq \frac{1}{2\pi} \int_R^\infty e^{-\sigma^2 r^2} \left(\log \frac{r}{2\pi} + O(1/r) \right) dr.$$

With the change $r \mapsto r/\sigma$, the integral equals $\sigma^{-1} \int_{R\sigma}^\infty e^{-r^2} \log(r/\sigma) dr + O(\sigma^{-1})$, and $\int_{R\sigma}^\infty e^{-r^2} dr \leq e^{-R^2\sigma^2}$. $\square$

9.2 Truncation of the prime-power sum

Lemma 9.2. *There exists an absolute constant C_2 such that, for every $N > e^{\Delta_{\max}}$ and every $\Delta \in [0, \Delta_{\max}]$,*

$$|C_{\text{prime}}(\Delta) - C_{\text{prime}}^{(N)}(\Delta)| \leq C_2 \sigma e^{\Delta_{\max}/2} e^{-(\log N - \Delta)^2/(4\sigma^2)}.$$

Proof. Set $f(y) = e^{-(y-\Delta)^2/(4\sigma^2)}$. Then

$$C_{\text{prime}} - C_{\text{prime}}^{(N)} = -\frac{1}{4\sigma\sqrt{\pi}} \sum_{n>N} \frac{\Lambda(n)}{\sqrt{n}} [f(\log n) + f(-\log n)].$$

For $\Delta \geq 0$ and $\log n > \log N > \Delta_{\max} \geq \Delta$, $f(-\log n) \leq f(\log n)$. By partial summation with $\psi(x) = \sum_{n \leq x} \Lambda(n) \leq C_0 x$ (Chebyshev),

$$\sum_{n>N} \frac{\Lambda(n)}{\sqrt{n}} f(\log n) = -\frac{\psi(N)}{\sqrt{N}} f(\log N) + \int_N^\infty \psi(x) \frac{d}{dx} \left[\frac{f(\log x)}{\sqrt{x}} \right] dx.$$

After the change $y = \log x$, the derivative factor is

$$\frac{d}{dx} \left[\frac{f(\log x)}{\sqrt{x}} \right] dx = e^{-y/2} f(y) \left(\frac{1}{2} + \frac{y-\Delta}{2\sigma^2} \right) dy,$$

so the integrand is bounded by $C_0 e^{y/2} f(y) \left| \frac{1}{2} + \frac{y-\Delta}{2\sigma^2} \right|$. This has a maximum of order $e^{\Delta/2}$ concentrated on an interval of length $\asymp \sigma$ around $y \asymp \Delta + \sigma^2$, hence its integral over $\mathbb{R}$ is of order $\sigma e^{\Delta/2}$. Under the constraint $y > \log N$, the Gaussian factor is suppressed by $e^{-(\log N - \Delta)^2/(4\sigma^2)}$, giving the claim. The same bound applies to $f(-\log n)$. $\square$

Remark 9.3. Choosing $\log N = \Delta_{\max} + c\sigma$, the bound of Lemma 9.2 becomes $C_2 \sigma e^{\Delta_{\max}/2} e^{-c^2/4}$. With $c = 10$ the factor $e^{-c^2/4} = e^{-25} \approx 1.4 \times 10^{-11}$. In the numerical run of §8 with $\sigma = 0.02$, $\Delta_{\max} = 11.98$, the effective truncation error is $O(10^{-10})$, well below the accuracy achieved by ESPRIT.

9.3 Stability of ESPRIT

Let $C(\Delta) = \sum_{k=1}^M a_k \cos(\gamma_k \Delta)$ with $0 < \gamma_1 < \dots < \gamma_M$, and $a_k = e^{-\sigma^2 \gamma_k^2}$. Let $\delta > 0$ and $K \geq 4M$, $L = \lfloor K/2 \rfloor$, and let $H \in \mathbb{R}^{L \times L}$ be the Hankel matrix $H_{ij} = C((i+j)\delta)$.

Theorem 9.4. Let $C^\eta = C + \eta$ be a perturbation with $|\eta(\Delta)| \leq \varepsilon$. Let $\hat{\gamma}_1, \dots, \hat{\gamma}_M$ be the frequencies returned by ESPRIT applied to C^η . Then there exists a constant $C_{M,\delta,K}$ depending only on $M, \delta, K, \gamma_1, \dots, \gamma_M$ such that, for every k ,

$$|\hat{\gamma}_k - \gamma_k| \leq C_{M,\delta,K} \varepsilon e^{\sigma^2 \gamma_M^2},$$

provided $C_{M,\delta,K} \varepsilon e^{\sigma^2 \gamma_M^2} < \frac{1}{2} \min_k (\gamma_{k+1} - \gamma_k)$.

Proof. Three ingredients. **(i) Weyl.** The $2M$ -th singular value of H satisfies $\sigma_{2M}(H) \geq c_1 L e^{-\sigma^2 \gamma_M^2}$, and $\|\eta\|_2 \leq c_2 \varepsilon L$. For $e^{-\sigma^2 \gamma_M^2} > c_3 \varepsilon$ the $2M$ dominant singular values remain separated from the tail. **(ii) Davis–Kahan.** The angle θ between the true and perturbed subspaces satisfies $\sin \theta \leq c_6 \varepsilon e^{\sigma^2 \gamma_M^2}$. **(iii) Bauer–Fike.** The shift operator $\Psi = U_1^+ U_2$ is diagonalisable with eigenvector matrix V satisfying $\|V\| \|V^{-1}\| \leq C_{M,\delta}$. Then $\|\delta \Psi\|_2 \leq c_7 \sin \theta$, and $\max_k |\delta z_k| \leq C_{M,\delta} \|\delta \Psi\|_2$. Since $|\delta \gamma_k| = |\delta z_k|/\delta$, the claim follows. $\square$

Remark 9.5. The factor $e^{\sigma^2 \gamma_M^2}$ is intrinsic to the inverse problem: it is the condition number of the map $\mu \mapsto C_\sigma$ evaluated at the last frequency γ_M . It appears for any inversion method, not only for ESPRIT.

9.4 Vandermonde conditioning

We now estimate $C_{M,\delta,K}$. Let $Z \in \mathbb{C}^{L \times 2M}$ be the Vandermonde matrix with entries $Z_{i,(k,\pm)} = (z_k^\pm)^i$, $z_k^\pm = e^{\pm i \gamma_k \delta}$, $0 \leq i < L$. Then $C_{M,\delta,K}$ is controlled by $\kappa_2(Z)$.

Theorem 9.6 (Gautschi bound). Set $\delta_{\text{sep}} = \min_{k \neq j} |z_k - z_j|$. Then

$$C_{M,\delta,K} \leq \sqrt{2M} \left(1 + \frac{1}{\delta_{\text{sep}}} \right)^{2M-1}.$$

Proof. The entries of Z^+ are bounded by N and those of Z^{-1} by $(1 + 1/\delta_{\text{sep}})^{2M-1}$ (each is a coefficient of a Lagrange polynomial with bounded factors). Standard norm inequalities give the result. $\square$

Theorem 9.7 (Refined bound for quasi-uniform zeros). Suppose $\{\gamma_k\}_{k=1}^M \subset [0, T]$ is (α, β) -regular in the sense that

$$\Delta \gamma_{\min} := \min_k (\gamma_{k+1} - \gamma_k) \geq \beta \cdot \frac{2\pi}{\log(T/2\pi) M^{1-\alpha}},$$

with $0 < \alpha < 1$, $\beta > 0$. Then, for $\delta \asymp \sigma$ and $T \asymp 1/\sigma$,

$$\log C_{M,\delta,K} \leq (2M-1) \log \left(1 + \frac{M^{1-\alpha} \log(1/(2\pi\sigma))}{2\pi\beta\sigma} \right) + O(\log M).$$

Proof. By Theorem 9.6, it suffices to bound $1/\delta_{\text{sep}}$. Since $\delta_{\text{sep}} \geq \Delta \gamma_{\min} \delta$ and $\delta \leq \sigma$,

$$\frac{1}{\delta_{\text{sep}}} \leq \frac{M^{1-\alpha} \log(T/2\pi)}{2\pi\beta\delta} \leq \frac{M^{1-\alpha} \log(1/(2\pi\sigma))}{2\pi\beta\sigma},$$

using $T \asymp 1/\sigma$ and $\delta \asymp \sigma$. Substituting into the Gautschi bound gives the claim. $\square$

Corollary 9.8. With $\sigma = 0.02$, $M = 14$, and (α, β) chosen so that the (α, β) -regularity hypothesis of Theorem 9.7 is satisfied with margin by the first 14 Riemann zeros (e.g. $\alpha = 1$, $\beta \approx 15$, or equivalently $\alpha = 1/2$ with $\beta \approx 50$), the crude Gautschi bound gives $C_{M,\delta,K} \lesssim 10^{46}$, while the refined bound gives $C_{M,\delta,K} \lesssim 10^9$. The observed recovery error in the numerical experiments of §8 is $\sim 10^{-4}$, consistent with the refined bound.

Remark 9.9. The estimate of Theorem 9.7 is exponential in M . This is the fundamental numerical bottleneck of the pipeline: with $M \rightarrow \infty$, the conditioning of the Vandermonde matrix blows up. It cannot be removed by improving the truncation bounds of Lemmas 9.1 and 9.2, because it is a property of the inverse problem itself, not of the data.

10 Discussion

What is proved.

- Self-adjointness and $\mathbb{Z}/2$ -symmetry of H_n .
- Trace formula $\text{Tr}(H_n^2) = 2 \sum_{p \leq n} c_p^2 \sigma_p(n/p)$.
- Carleman bound for $c_p = p^{-s}$, $s > 1$.
- Kernel identity $\dim \ker H_n = d_n - 2 \text{rank}(A)$ and index identity $\text{index}(A) = M(n)$ for the full block A .
- $\ker A_{\uparrow}$ contains the maximal elements; $\dim \ker A = 889$ at $n = 10^4$ (correction of the fourth revision).
- Exact parametrisation of maximal elements (Theorem 4.7) and exact formula (12) for $M_{\mathcal{M}}(n)$.
- Density $c_1 = 0.5675527 \dots$ of the maximal elements (Theorem 4.11).
- Exact kernel decomposition of the residual $\mathcal{R}(n)$ (Proposition 5.1), and the asymptotic limit $c_R^{\text{exact}} \rightarrow c_R^{\text{asym}}$ (Proposition 5.2).
- Numerical power-law decay of the kernel correction, with exponent sample mean 0.4918 and sample std 0.0506 over 175 zeros (Observation 5.4); no correlation with $\log \gamma_k$ or k (Observation 5.5).
- Block-bootstrap standard error $\text{SE}_{\text{boot}} = 0.0078$, showing that the deviation of $\bar{\alpha}$ from $1/2$ is ≈ 1.0 bootstrap standard error and that the hypothesis $\alpha = 1/2$ is not rejected (Observation 5.6).
- Rejection of the logarithmic-correction hypothesis $\Delta_k(n) \propto n^{-1/2}(\log n)^{-\beta_k}$, with $\bar{\beta} = -0.081$ of the wrong sign and symmetrically distributed t -statistics (Remark 5.8).
- Insensitivity of all reported quantities to the truncation parameter $K \in \{10, 20, 30, 50\}$ to machine precision (Observation 5.7).
- Prefactor $1/(4\sigma\sqrt{\pi})$ in the explicit formula.
- Exact-recovery theorem for the Toeplitz pencil (Theorem 8.4).
- Truncation bounds for the explicit formula (Lemmas 9.1 and 9.2).
- ESPRIT stability theorem (Theorem 9.4) and Vandermonde conditioning bounds (Theorems 9.6 and 9.7).
- Two independent rigorous structural obstructions to the convergence of $\tilde{H}_n$ to the zeros of ζ (Propositions 7.1 and 7.4).

What is corrected.

- The statement $\ker A = \{0\}$ of the fourth revision is false; $\dim \ker A = 889$ at $n = 10^4$.
- The block A of the earlier revisions is $A_{\uparrow}$, not the full block $A = A_{\uparrow} + A_{\downarrow}$.
- The reported $c_* \approx 0.185$ is $\text{rank}(A_{\uparrow})/d_n$, not $\text{rank}(A)/d_n$.
- The kernel density is $1 - 2c \approx 0.296$, not $1 - 2c_* \approx 0.630$.

- The formula $\text{rank}(A) = \text{rank}(A_{\uparrow}) + \text{rank}(A_{\downarrow})$ is false; the correct inequality is $\leq$.
- The operators $\gamma_p = T_p + T_p^*$ do *not* commute in the truncated complex.
- The spectral modification $\tilde{H}_n$ cannot converge to the zeros of ζ , for the two rigorous reasons of §7.
- A misprint of the fifth revision on $M(2 \cdot 10^5)$ (the correct value is $M(2 \cdot 10^5) = -1$) has been verified by an independent sieve using `sympy.mobius`.
- The statement of the fifth revision that “the deviation of the sample mean from $1/2$ is approximately 3 standard errors” reflects the *naive* SEM. With a block-bootstrap error ($\text{SE}_{\text{boot}} = 0.0078$), the deviation is ≈ 1.0 standard error (Observation 5.6).

What is not proved.

- **No proof of RH.**
- **No new characterisation of RH.** The algebraic decomposition is analytic-content-free (Proposition 6.1), and the numerical exponent $\alpha \approx 1/2$ is consistent with the obstruction.
- **No sharp asymptotic for $C_{M,\delta,K}$.** Only bounds are proved; the sharp behaviour under GUE statistics is open.
- **No convergence of $\tilde{H}_n$ to the zeros.** The conjecture of the fourth revision is false.
- **No quantitative link between α and the minimal-variance condition.** The connection is suggested by the numerical coincidence, but a rigorous derivation is not given.
- **No rigorous proof of Obstruction III.** The density heuristic is not a proof, because the operators γ_p do not commute in the truncated complex.
- **No definitive resolution of the origin of the small residual deficit $\bar{\alpha} - 1/2 \approx -0.008$.** The log-correction family and the truncation of c_R^{asym} are excluded; a possible origin is a subleading term of the explicit formula or a finite-size effect of the n -grid.

Open problems.

- (1) Determine the asymptotic behaviour of the residual $R_n^{\uparrow}$ and of the additivity defect $\delta(n)$.
- (2) Determine the weak limit μ^* of the spectral measures μ_n at $s = 1/2$; compute its moments and Hausdorff dimension.
- (3) Extend the exact rank computation beyond $n = 10^4$ using sparse rank algorithms (Wiedemann, LinBox).
- (4) Determine the sharp asymptotic of the Vandermonde conditioning $C_{M,\delta,K}$ for Riemann zeros under GUE statistics.
- (5) Rigorous convergence of the Toeplitz pencil to the multiplication operator by x on $L^2(\mu)$.
- (6) **Sharpen the critical exponent beyond the log-correction family.** The logarithmic-correction hypothesis $n^{-1/2}(\log n)^{-\beta}$ is empirically rejected (Remark 5.8), and the truncation of c_R^{asym} is ruled out as a systematic source (Observation 5.7). The deficit $\bar{\alpha} - 1/2 \approx -0.008$, if real, must originate from one of the following: (a) finite-size effects from the grid in n ; (b) subleading terms of the explicit formula not captured by the leading kernel; (c) a genuine power-law correction of the form $n^{-1/2+\epsilon}$ with ϵ small. A definitive test requires: (i) an extended grid in n up to 10^{10} or 10^{12} ; (ii) independent refinements of the n -grid to isolate discretisation effects; (iii) explicit control of the remainder in Proposition 5.2.

- (7) **Quantitative link to minimal-variance:** establish rigorously that the exact kernel $c_R^{\text{exact}}(\rho, n)$ is a valid defect observable for the minimal-variance condition, and determine the correction to α induced by a hypothetical zero with $\beta_0 > 1/2$.
- (8) **Make Obstruction III rigorous.** Determine the actual asymptotic spectral density of H_n on its support, and show that it differs from the logarithmic density of the Riemann zeros.

Statements and Declarations

Funding. This research received no external funding.

Conflicts of interest. The author declares no conflict of interest.

Ethical approval. Not applicable.

Author contributions. The author is the sole contributor.

Data availability statement. The data and code underlying the numerical results are openly available at <https://zenodo.org/records/22831623>, MIT License for code, CC BY 4.0 for data.

Declaration of AI use. During the preparation of this work, no AI tool was used.

Acknowledgements. The author thanks the developers of Python, NumPy, SciPy, mpmath, and SageMath.

A Proof of the exact-recovery theorem

For $k = 1, \dots, M$ define $u_k, u_{-k} \in \mathbb{C}^K$ by $u_k(j) = e^{i\gamma_k j \delta}$, $u_{-k}(j) = e^{-i\gamma_k j \delta}$. Let $U = [u_1, \dots, u_M, u_{-1}, \dots, u_{-M}] \in \mathbb{C}^{K \times 2M}$, and set

$$B = \text{diag}(a_1/2, \dots, a_M/2, a_1/2, \dots, a_M/2),$$

$$D = \text{diag}(a_1\gamma_1/(2i), \dots, a_M\gamma_M/(2i), -a_1\gamma_1/(2i), \dots, -a_M\gamma_M/(2i)).$$

Lemma A.1. U has full column rank $2M$.

Proof. The hypothesis $0 < \gamma_1 < \dots < \gamma_M < \pi/\delta$ guarantees that the $2M$ numbers $z_k = e^{i\gamma_k \delta}$ and z_k^{-1} are pairwise distinct. If $\sum_k (\alpha_k u_k + \beta_k u_{-k}) = 0$, then the sequence $s_j = \sum_k (\alpha_k z_k^j + \beta_k z_k^{-j})$ vanishes for $j = 0, \dots, K-1$; restricting to $2M$ consecutive indices, the Vandermonde system is invertible, so all $\alpha_k = \beta_k = 0$. $\square$

Lemma A.2. $G = UBU^*$ and $S = UDU^*$.

Proof. Entrywise computation using $e^{i\gamma_l(j-k)\delta} = u_l(j)\overline{u_l(k)}$, $e^{-i\gamma_l(j-k)\delta} = u_{-l}(j)\overline{u_{-l}(k)}$, and $\cos \theta = (e^{i\theta} + e^{-i\theta})/2$. $\square$

Proof of Theorem 8.4. If v satisfies $iSv = \lambda Gv$ and $Gv \neq 0$, setting $\eta = U^*v$ gives $U(iD - \lambda B)\eta = 0$. Since U has full rank (Lemma A.1), $(iD - \lambda B)\eta = 0$; if $\eta = 0$, then $Gv = 0$, contradiction. So $iD - \lambda B$ is singular, forcing $\lambda = \pm\gamma_k$.

Conversely, for each $\lambda = \pm\gamma_k$, choose $\eta = e_{\pm k}$ and let $v \in \mathbb{C}^K$ be any vector with $U^*v = e_{\pm k}$; such a v exists because U^* is surjective onto $\mathbb{C}^{2M}$. Then $Gv = UBU^*v = UBe_{\pm k} = (a_k/2)u_{\pm k} \neq 0$ and $iSv = iUDU^*v = iUDe_{\pm k} = \pm\gamma_k(a_k/2)u_{\pm k} = \lambda Gv$. Hence λ is a generalised eigenvalue with $Gv \neq 0$. $\square$

B Computational details

C.1 Sieves. Möbius function and $\Omega(m)$ are computed by independent sieves up to n . Cross-verification with `sympy.mobius` has been performed for all tabulated values of $M(n)$.

C.2 Assembly of the bipartite block. The full block A has entries $A_{pm,m} = c_p$ for $p \nmid m$, $pm \leq n$ (up-shift), and $A_{m/p,m} = c_p$ for $p \mid m$ (down-shift).

C.3 Rank over $\mathbb{Q}$. The rank is computed modulo each prime in $\{3, 5, 7, 11, 13\}$; the maximum over the five primes is reported. The result is a lower bound on the exact rank, and it is exact whenever the five primes are sufficient to isolate a nonzero minor of size equal to the reported rank. Verifying this is done by exhibiting, in each case, a minor whose determinant is nonzero modulo at least one of the five primes. For $n = 10^4$ the computation takes 154 seconds; for larger n sparse rank algorithms are required.

C.4 Exact formula for $M_{\mathcal{M}}(n)$. Implemented as in (12), iterating over k and using `numpy.gcd` for the coprimality test. Runs in 0.05 seconds for $n = 10^5$.

C.5 Exact kernel. The exact kernel $c_R^{\text{exact}}(\rho, n)$ is evaluated by direct summation of (23), with the P_k -smooth sets generated recursively. The asymptotic kernel $c_R^{\text{asym}}(\rho)$ is evaluated by the truncated series (24) with $K = 30$, giving accuracy better than 10^{-20} . The insensitivity to K for $K \in \{10, 20, 30, 50\}$ is verified in Observation 5.7.

C.6 Block bootstrap. The block bootstrap is performed with $B = 8$ contiguous blocks of 10 points each on the 80-point grid in n , with 1000 resamples and a fixed random seed. The choice of B is not critical: $B \in \{4, 8, 10, 16, 20\}$ gives $\text{SE}_{\text{boot}} \in [0.0058, 0.0104]$.

C.7 Numerical pipeline. The truncation constants c, c' are chosen as $c = 10$, $c' = 8$, giving truncation errors below 10^{-10} . The archimedean integral is evaluated by Gauss–Legendre quadrature with 2000 nodes on $[0, c'/\sigma]$. The prime-power sum is computed by vectorised NumPy over the precomputed list of prime powers $\leq N_{\text{max}}$. ESPRIT uses the SVD of the Hankel matrix; the rank is chosen by truncating singular values below $10^{-3}\sigma_1$.

References

- [1] M. V. Berry and J. P. Keating, *The Riemann zeros and eigenvalue asymptotics*, SIAM Rev. **41** (1999), 236–266.
- [2] A. Björner, *A cell complex in number theory*, Adv. Appl. Math. **46** (2011), 71–85.
- [3] A. Connes, *Trace formula in noncommutative geometry and the zeros of the Riemann zeta function*, Selecta Math. (N.S.) **5** (1999), 29–106.
- [4] A. P. Guinand, *A summation formula in the theory of prime numbers*, Proc. London Math. Soc. **s2-50** (1948), 107–119.
- [5] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, AMS Colloquium Publications **53**, 2004.
- [6] D. Mayer, *On the thermodynamic formalism for the Gauss map*, Comm. Math. Phys. **130** (1990), 311–333.

- [7] H. L. Montgomery, *The pair correlation of zeros of the zeta function*, in: Analytic Number Theory, Proc. Sympos. Pure Math. **24**, AMS, 1973, 181–193.
- [8] A. M. Odlyzko, *On the distribution of spacings between zeros of the zeta function*, Math. Comp. **48** (1987), 273–308.
- [9] L. E. Pavesi, *Prime-shift operators on Björner’s complex of squarefree integers: spectral obstruction, and an extended Toeplitz pencil*, preprint, 2026.
- [10] L. E. Pavesi, *Prime-shift operators on Björner’s complex of squarefree integers: spectral obstruction, kernel correction, and a Hankel–ESPRIT pipeline*, preprint (first revision), 2026.
- [11] D. Schumayer and D. A. W. Hutchinson, *Physics of the Riemann Hypothesis*, Rev. Mod. Phys. **83** (2011), 307–330.
- [12] Y. Shi, *Construction of finite Hilbert–Pólya matrices from Weil’s explicit formula*, preprint, arXiv:2609.04908.
- [13] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, 1990.
- [14] J. J. M. Verbaarschot, *The spectrum of the chiral Gaussian orthogonal ensemble*, Phys. Rev. Lett. **72** (1994), 2531–2533.
- [15] A. Weil, *Sur les “formules explicites” de la théorie des nombres premiers*, Comm. Sémin. Math. Univ. Lund (1952), 252–265.
- [16] E. Yakaboylu, *On the existence of the Hilbert–Pólya Hamiltonian*, preprint, arXiv:2408.15135.