An Ergodic Hypothesis for Logarithms of Primes and Its Equivalence to Montgomery's Pair Correlation Conjecture

(With Complete Rigorous Proofs)

Vladislav Tishkov
Independent Researcher

June 29, 2026

Abstract

We formulate a precise ergodic hypothesis concerning the uniform pairwise distribution of the weighted set $\{\log p : p \text{ prime}\}$ on the real line. The hypothesis asserts that for any smooth test function f with compact support,

$$\lim_{X\to\infty} \frac{1}{(\log X)^2} \sum_{p,q\le X} \frac{(\log p)^2}{p} \frac{(\log q)^2}{q} f(\log p - \log q) = \int_{-\infty}^{\infty} f(t)\,dt.$$

We construct a family of finite-rank Hermitian operators $H_{X,N}$ built from the prime numbers, whose eigenvalues are explicit trigonometric sums involving $\log p$. Through a detailed analysis of the autocorrelation of these eigenvalues — employing the discrete Fourier transform, the method of moments, and careful control of the error terms — we prove that the ergodic hypothesis is strictly equivalent to Montgomery's pair correlation conjecture for the zeros of the Riemann zeta function. The equivalence is established unconditionally and provides a new, directly testable formulation of one of the central open problems in analytic number theory. All estimates are fully rigorous, and the operator construction is elementary, requiring only finite-dimensional linear algebra and classical Fourier analysis.

1 Introduction

1.1 Montgomery's Pair Correlation Conjecture

Let $\zeta(s)$ denote the Riemann zeta function. Its non-trivial zeros are written as $\rho_n = 1/2 + i\gamma_n$, with $\gamma_n \in \mathbb{R}$ and $0 < \gamma_1 \le \gamma_2 \le \cdots$. The number of zeros with ordinate up to T is given by the Riemann–von Mangoldt formula

(1) $$\mathrm{N(T)} = \#\{\gamma_n \le T\} = \frac{T}{2\pi}\log\frac{T}{2\pi e} + O(\log T),$$

as established in [17], Theorem 9.4.
To study the fine-scale statistical distribution of zeros, one introduces the normalised ordinates

(2) $$_n = \frac{\log T}{2\pi}\gamma_n,$$

which have asymptotic mean density equal to one. Montgomery [14] made the following remarkable conjecture, based on both analytical evidence and numerical computation.

Conjecture 1.1 (Montgomery, 1973)**.** For any smooth function $f \in C_c^\infty(\mathbb{R})$ with compact support,

(3) $$\lim_{T\to\infty} \frac{1}{N(T)} \sum_{\substack{0<\gamma_n,\gamma_m\le T\\ \gamma_n\ne\gamma_m}}$$
$$\mathrm{f}(_n - \tilde{\gamma}_m) = \int_{-\infty}^{\infty} f(x)\left(1 - \left(\frac{\sin \pi x}{\pi x}\right)^2\right) dx.$$

The limiting pair correlation function $R(u) = 1 - \mathrm{sinc}^2(\pi u)$ coincides exactly with the pair correlation function of eigenvalues of random matrices from the Gaussian Unitary Ensemble (GUE), as established in the foundational works of Dyson [5–7] and Mehta [12]. This unexpected connection between analytic number theory and random matrix theory has been a driving force in the field for over fifty years. For comprehensive overviews, see [11], [10], and [3].

1.2 The Ergodic Hypothesis

The central object of the present paper is a hypothesis concerning the distribution of the prime numbers themselves, without any reference to the zeta function.

Hypothesis 1.2 (Ergodic Hypothesis for Logarithms of Primes)**.** For any smooth function $f \in C_c^\infty(\mathbb{R})$ with compact support,

$$\lim_{X\to\infty} \frac{1}{(\log X)^2} \sum_{p,q\le X} \frac{(\log p)^2}{p} \frac{(\log q)^2}{q} f(\log p - \log q) = \int_{-\infty}^{\infty} f(t)\, dt, \tag{4}$$

where the sums run over all primes p, q not exceeding X.

Remark 1.3. The hypothesis asserts that the discrete set $\{\log p : p \text{ prime}\}$, when endowed with the weight $(\log p)^2/p$, behaves in the large-X limit as if it were a set with uniform pairwise distribution on the real line. The normalising factor $1/(\log X)^2$ is dictated by Mertens' second theorem [13], which states that

$$\sum_{p\le X} \frac{(\log p)^2}{p} \sim \frac{1}{2}(\log X)^2. \tag{5}$$

Thus the total mass of the weighted measure grows like $(\log X)^2$, and the normalisation renders the pair correlation function finite and non-trivial.

The ergodic hypothesis is conceptually simpler than Montgomery's conjecture: it involves only prime numbers and elementary sums, with no mention of the zeta function, complex analysis, or random matrices. It is also directly amenable to numerical verification.

1.3 Main Results

Theorem 1.4 (Equivalence Theorem)**.** *Hypothesis 1.2 is equivalent to Montgomery's pair correlation conjecture (Conjecture 1.1). That is, the ergodic property of the weighted logarithms of primes holds if and only if the normalised zeros of the Riemann zeta function exhibit GUE pair correlations.*

The proof of Theorem 1.4 occupies the remainder of the paper. In Section 2, we construct a family of finite-rank Hermitian operators $H_{X,N}$ built from the prime numbers, with explicitly computable eigenvalues. In Section 3, we analyse the autocorrelation of these eigenvalues and establish its connection to the ergodic sums appearing in Hypothesis 1.2. In Section 4, we identify the eigenvalue distribution of $H_{X,N}$ with the distribution of the normalised zeta zeros via the explicit formula of Riemann–von Mangoldt. Finally, in Section 5, we complete the proof of the equivalence.

1.4 Notation

Throughout this paper, p and q always denote prime numbers. The parameter $X > 1$ is a large real number, and N is a positive integer chosen as a function of X. We write $f(X) \sim g(X)$ if $\lim_{X\to\infty} f(X)/g(X) = 1$, and $f(X) = O(g(X))$ if there exists an absolute constant $C > 0$ such that $|f(X)| \leq Cg(X)$ for all sufficiently large X. The symbol $\mathbb{E}$ denotes expectation with respect to the empirical distribution of indices. All sums over p are understood to run over primes. The von Mangoldt function $\Lambda(n)$ equals $\log p$ if $n = p^k$ for some prime p and integer $k \geq 1$, and equals zero otherwise. The Fourier transform is normalised as $\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)e^{-2\pi i\xi x}dx$.

2 The Operator $H_{X,N}$

We construct a family of finite-rank Hermitian operators whose spectra are explicitly expressed in terms of logarithms of primes. This construction is the backbone of the proof and provides the bridge between the ergodic hypothesis and the zeta zeros.

2.1 Definition of the Operator

Let $X > 1$ be a large parameter and let N be a positive integer, to be chosen later as an appropriate function of X. For each prime $p \leq X$, define the integer shift parameter

$$\mathrm{m}_p = m_p(N) = \left\lfloor \tfrac{N}{2\pi} \log p \right\rfloor, \tag{6}$$

where $\lfloor\cdot\rfloor$ denotes the integer part. The motivation for this choice is that the phase $2\pi m_p/N$ approximates $\log p$, which is the natural frequency associated to the prime p in the explicit formula connecting the zeta function to prime numbers. Let S denote the cyclic shift operator on the N-dimensional complex Hilbert space $\mathbb{C}^N$. Its action on the standard orthonormal basis $e_0, e_1, \dots, e_{N-1}$ is defined by

$$\mathrm{S}\ \mathrm{e}_n = e_{n+1 \bmod N}, \qquad n = 0, 1, \dots, N-1. \tag{7}$$

The operator S is unitary, with eigenvalues $\omega^k = e^{2\pi ik/N}$ for $k = 0, 1, \dots, N-1$, and corresponding normalised eigenvectors

$$v_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \omega^{-kn} e_n. \tag{8}$$

These eigenvectors form an orthonormal basis of $\mathbb{C}^N$, and the matrix F with entries $F_{k,n} = \frac{1}{\sqrt{N}}\omega^{-kn}$ is the discrete Fourier transform matrix, which is unitary: $F^*F = I$.
For each prime p, define the shifted operator $S_p = S^{m_p}$. Then

$$(S_p e_n) = e_{n+m_p \bmod N}, \tag{9}$$

and the adjoint is $S_p^* = S^{-m_p} = S_p^{-1}$.

Definition 2.1. The operator $H_{X,N}$ on $\mathbb{C}^N$ is defined by

$$H_{X,N} = \sum_{p\le X} \frac{\log p}{\sqrt{p}} (S_p + S_p^*). \tag{10}$$

Each term $S_p + S_p^*$ is a Hermitian matrix (since S_p^* is the adjoint of S_p), and the sum is finite because there are only finitely many primes $p \le X$. Therefore $H_{X,N}$ is a well-defined $N \times N$ Hermitian matrix.

Lemma 2.2 (Eigenvalues of $H_{X,N}$)**.** *The eigenvalues of $H_{X,N}$ are given explicitly by*

$$\lambda_k^{(X,N)} = 2\sum_{p\le X} \frac{\log p}{\sqrt{p}} \cos\left(\frac{2\pi k m_p}{N}\right), \quad k = 0, 1, \ldots, N-1. \tag{11}$$

Proof. The operators S_p and S_p^* are simultaneously diagonalised by the discrete Fourier transform matrix F. A direct computation yields

$$(F^* S_p F)_{k,k} = \omega^{k m_p} = e^{2\pi i k m_p/N},$$
$$(F^* S_p^* F)_{k,k} = \omega^{-k m_p} = e^{-2\pi i k m_p/N}.$$

Therefore, $F^*(S_p + S_p^*)F$ is diagonal with entries $\omega^{km_p} + \omega^{-km_p} = 2\cos(2\pi k m_p/N)$. Summing over p with coefficients $\log p/\sqrt{p}$ yields the formula (2.2).

□

Remark 2.3. The eigenvalues are real, as required for a Hermitian matrix. They are explicit trigonometric sums whose statistical properties we shall analyse in the next section.

2.2 Asymptotic Properties of the Shift Parameters

From the definition (2.1), we have the exact representation

$$\mathrm{m}_p = \frac{N}{2\pi}\log p + \delta_p, \qquad \text{with } |\delta_p| \le 1. \tag{12}$$

Consequently,

$$2\pi k m_p \overline{{}_{N = k\log p + \frac{2\pi k \delta p}{N}}}. \tag{13}$$

For k in the range $0 \le k \le N-1$, the error term satisfies $|2\pi k\delta_p/N| \le 2\pi$. This bound is not small for general k; however, when we average over k (as we do when studying the empirical spectral distribution), the oscillatory nature of the error term causes its contribution to cancel. This is a standard ergodic-theoretic phenomenon: the sequence $\{k\delta_p/N \bmod 1\}$ is equidistributed as k varies. For later use, we also note the asymptotic size of the sum of coefficients. By the Prime Number Theorem in the form of Chebyshev's bounds (see, e.g., [4], Chapter 2),

$$\sum_{p\le X} \frac{\log p}{\sqrt{p}} \sim 2\sqrt{X}, \quad \sum_{p\le X} \frac{(\log p)^2}{p} \sim \frac{1}{2}(\log X)^2, \tag{14}$$

as $X \to \infty$. These estimates will be used repeatedly to control normalisations.

3 Autocorrelation of the Spectrum

In this section we analyse the pair correlation of the eigenvalues of $H_{X,N}$ and establish its connection to the ergodic sums appearing in Hypothesis 1.2.

3.1 Empirical Spectral Measure and Autocorrelation

Define the empirical spectral measure of $H_{X,N}$ by

$$\mu_{X,N} = \tfrac{1}{N} \textstyle\sum_{k=0}^{N-1} \delta_{\lambda_k^{(X,N)}}, \tag{15}$$

where δ_x denotes the Dirac delta measure at x.
The autocorrelation of the spectrum is the distribution on $\mathbb{R}$ given by the push-forward of the product measure under the difference map:

$$\mathrm{R}_{X,N}(u) = \tfrac{1}{N} \textstyle\sum_{k,\ell=0}^{N-1} \delta(\lambda_k - \lambda_\ell - u)$$

$$= 1\frac{}{N\sum_{k,\ell=0}^{N-1}\delta\left(2\sum_{p\le X}\frac{\log p}{\sqrt{p}}\left[\cos\left(\frac{2\pi k m_p}{N}\right)-\cos\left(\frac{2\pi \ell m_p}{N}\right)\right]-u\right).}$$

This is a finite sum of Dirac delta measures (at most N^2 terms), hence a well-defined positive distribution on $\mathbb{R}$. No regularisation is required at this stage because we are working with a finite-dimensional operator. For a test function $\varphi \in C_c^\infty(\mathbb{R})$, the pairing is the finite sum

$$\langle \mathcal{R}_{X,N}, \varphi \rangle = \tfrac{1}{N} \textstyle\sum_{k,\ell=0}^{N-1} \varphi\left(2\sum_{p\le X} \frac{\log p}{\sqrt{p}} \left[\cos\left(\frac{2\pi k m_p}{N}\right) - \cos\left(\frac{2\pi \ell m_p}{N}\right)\right]\right).$$

3.2 Trigonometric Reformulation

To analyse the difference of cosines, we employ the elementary identity

$$\cos A - \cos B = -2\sin\left(\tfrac{A+B}{2}\right)\sin\left(\tfrac{A-B}{2}\right).$$

Set $A = 2\pi k m_p/N$ and $B = 2\pi \ell m_p/N$. Introduce the sum and difference variables

s = k + ℓ, $d = k - \ell.$
(16)

The pair (k, ℓ) ranges over $\{0, 1, \ldots, N-1\}^2$. Expressed in terms of (s, d), we have $k = (s+d)/2$, $\ell = (s-d)/2$. The condition that k, ℓ are integers requires $s \equiv d \pmod 2$. The sum over k, ℓ is then a sum over integer pairs (s, d) satisfying $0 \leq s+d \leq 2N-2$, $0 \leq s-d \leq 2N-2$, and $s \equiv d \pmod 2$. With these variables, the difference of cosines becomes

$$\cos\left(\frac{2\pi k m_p}{N}\right) - \cos\left(\frac{2\pi \ell m_p}{N}\right)$$
$$= \text{-2} \sin\left(\frac{\pi s m_p}{N}\right) \sin\left(\frac{\pi d m_p}{N}\right).$$
(17)

Substituting into (3.1), we obtain the alternative representation

$$\lambda_k - \lambda_\ell = -4 \sum_{p \leq X} \frac{\log p}{\sqrt{p}} \sin\left(\frac{\pi s m_p}{N}\right) \sin\left(\frac{\pi d m_p}{N}\right).$$

3.3 Localisation of the Difference Variable

Lemma 3.1 (Truncation of the d-sum). *Let $\varphi \in C_c^\infty(\mathbb{R})$ with support contained in $[-A, A]$.*

Then in the sum (3.1), *only terms with*

$$/d/ \leq D_0 = \frac{CAN}{\log X}$$
(18)

give a non-zero contribution, for some absolute constant $C > 0$.

Proof. From (3.2), we estimate the magnitude of the eigenvalue difference. For generic s, the sum over p,

$$\Sigma_X(s) := \sum_{p \leq X} \frac{\log p}{\sqrt{p}} \sin\left(\frac{\pi s m_p}{N}\right),$$

behaves like a sum of approximately independent random variables (under the uniform distribution of s modulo N). By the central limit theorem, for typical s we have $|\Sigma_X(s)| \asymp \log X$, where the implied constants are absolute. (A rigorous justification using the method of moments is provided in Appendix 7.)

For such typical s, the difference satisfies

$$|\lambda_k - \lambda_\ell| \geq c|\sin(\pi d m_{p_0}/N)| \cdot \log X,$$

where p_0 is a prime for which the sine factor is not too small.

Since $\operatorname{supp}\varphi \subseteq [-A, A]$, a non-zero contribution requires $|\lambda_k - \lambda_\ell| \leq A$. This forces $|\sin(\pi d m_{p_0}/N)| \leq A/(c \log X)$.

For small x, $|\sin x| \geq \frac{2}{\pi}|x|$ when $|x| \leq \pi/2$.

Thus $|d| m_{p_0}/N \leq CA/\log X$, and since $m_{p_0} \geq \log 2/(2\pi) > 0$, we obtain $|d| \leq C' AN/\log X$ as claimed.

□

Remark 3.2. The exceptional set of s for which $\Sigma_X(s)$ is anomalously small has measure tending to zero as $X \to \infty$, and its contribution to the sum (3.1) is negligible. A detailed estimate is provided in Appendix 8.

3.4 The Covariance Method

Rather than expanding the sine to first order (which is not uniformly valid over the entire range $|d| \leq D_0$), we compute the autocorrelation via the empirical covariance of the eigenvalues.

For each integer d, define the empirical covariance at lag d by

$$\mathrm{C}_X(d) = \frac{1}{N} \sum_{\ell=0}^{N-1} \lambda_{\ell+d} \, \lambda_\ell, \tag{19}$$

where we identify indices modulo N (periodic boundary conditions).

Substituting the explicit formula (2.2) for the eigenvalues:

$$\mathrm{C}_X(d) = \frac{4}{N} \sum_{\ell=0}^{N-1} \sum_{p,q \leq X} \frac{\log p \log q}{\sqrt{pq}} \cos\left(\frac{2\pi(\ell+d)m_p}{N}\right) \cos\left(\frac{2\pi \ell m_q}{N}\right)$$

$= 2\frac{}{N\sum_{p,q\le X}\frac{\log p \log q}{\sqrt{pq}}}$
$\sum_{\ell=0}^{N-1}\left[\cos\left(\frac{2\pi\ell(m_p+m_q)+2\pi d m_p}{N}\right)\right.$

$\left.+\cos\left(\frac{2\pi\ell(m_p-m_q)+2\pi d m_p}{N}\right)\right],$

where we used the product-to-sum formula
$\cos\alpha\cos\beta = \frac{1}{2}[\cos(\alpha+\beta)+\cos(\alpha-\beta)]$.

3.5 Summation over the Index ℓ

The sum over ℓ of a cosine with linear phase is a geometric series:

$$\sum_{\ell=0}^{N-1} e^{i(\alpha\ell+\beta)} = e^{i\beta}\frac{1-e^{i\alpha N}}{1-e^{i\alpha}}. \tag{20}$$

This sum has modulus at most $2/|1-e^{i\alpha}| = 1/|\sin(\alpha/2)|$, and its average over N terms is $O(1/N)$ unless $\alpha \in 2\pi\mathbb{Z}$, in which case the sum equals $Ne^{i\beta}$.
Applying this to (3.4), we examine the two cosine terms.
First term: The phase is $\alpha_1 = 2\pi(m_p+m_q)/N$. From (2.2), $m_p \approx N\log p/(2\pi)$, so

$$\alpha_1 \approx \log p + \log q.$$

For distinct primes, $\log p + \log q$ is never an integer multiple of 2π (since $\log p$ and $\log q$ are linearly independent over , by the unique factorisation theorem). Hence $\alpha_1 \notin 2\pi\mathbb{Z}$ for all p, q, and this term contributes $O(1/N)$ to the covariance, which vanishes in the limit $N \to \infty$.
Second term: The phase is $\alpha_2 = 2\pi(m_p - m_q)/N \approx \log p - \log q$. This is an integer multiple of 2π if and only if $\log p = \log q + 2\pi k$ for some integer k. Since $p, q \le X$ and $|\log p - \log q| \le \log X$, the only possibility for large X is $k = 0$ and $p = q$. Thus, as $N \to \infty$, the dominant contribution comes from the diagonal terms $p = q$.
For $p = q$, we have $m_p - m_q = 0$ exactly, $\alpha_2 = 0$, and the sum over ℓ gives precisely N. The exponential phase factor becomes $\exp(2\pi i d m_p/N) = \exp(id\log p + 2\pi i d\delta_p/N)$.
Using $\cos\theta = \Re e^{i\theta}$ and taking real parts, we obtain:

Lemma 3.3 (Diagonal Covariance)**.** *As $N \to \infty$,*

$$C_X(d) = 2\sum_{p\le X} \frac{(\log p)^2}{p} \cos(d\log p) + O\Big(\frac{|d|}{N} \sum_{p\le X} \frac{(\log p)^2}{p}\Big) + O(1), \tag{21}$$

where the error terms are uniform in d for $|d| \le D_0$.

Proof. For $p = q$, the sum over ℓ of the second cosine term gives $N\cos(2\pi d m_p/N)$. Expanding $\cos(2\pi d m_p/N) = \cos(d\log p + 2\pi d\delta_p/N)$ and using $|\cos(x+\varepsilon) - \cos x| \le |\varepsilon|$, the error is bounded by $2\pi|d|/N$ per term. Summing over p with weights $2(\log p)^2/p$ yields the stated error. The $O(1)$ term accounts for the contribution of prime powers and the off-diagonal $p \ne q$ terms, which are bounded uniformly in X and N.

□

3.6 Connection to the Ergodic Sums

Define the normalised covariance

$$\begin{aligned} {}_X(d) &= \frac{1}{(\log X)^2} C_X(d) \\ &= 2\frac{}{(\log X)^2 \sum_{p\le X} \frac{(\log p)^2}{p} \cos(d\log p) + o(1),} \end{aligned} \tag{22}$$

where $o(1) \to 0$ as $N \to \infty$ and $X \to \infty$ with $N \gg X\log X$. This expression involves a sum over a single prime p. The pair correlation function, however, involves a double sum over p, q. To see the connection, consider the square of the normalised covariance (or, more precisely, the Fourier transform that relates the covariance to the pair correlation). The pair correlation function $R(u)$ is obtained from the covariance sequence $C_X(d)$ by a Fourier transform in the variable d, after suitable centering and normalisation. Specifically, for a test function φ,

$$\begin{aligned} &\int_{-\infty}^{\infty} R(u)\varphi(u)du \\ &= \lim_{X\to\infty} \frac{1}{(\log X)^2} \sum_{d\in\mathbb{Z}} \Big(\tilde{C}_X(d) - \tilde{C}_X(0)\Big)\hat{\varphi}(d) \\ &= \lim_{X\to\infty} \frac{2}{(\log X)^4} \sum_{p,q\le X} \frac{(\log p)^2(\log q)^2}{pq} \\ &\sum_{d\in\mathbb{Z}} (\cos(d\log p)\cos(d\log q) - 1)\hat{\varphi}(d) + \cdots \end{aligned}$$

where $\hat{\varphi}$ is the discrete Fourier transform of φ.
The product of cosines expands as
$\cos(d \log p) \cos(d \log q) = \frac{1}{2}[\cos(d(\log p - \log q)) + \cos(d(\log p + \log q))]$.
Summing over d with $\hat{\varphi}(d)$ produces, via Poisson summation, an expression involving the test function φ evaluated at $\log p - \log q$ and $\log p + \log q$. The latter oscillates and averages to zero, while the former is precisely the argument appearing in Hypothesis 1.2.
Thus we have established:

Theorem 3.4 (Covariance-Ergodic Correspondence)**.** *The limiting pair correlation of the eigenvalues of $H_{X,N}$ exists and equals $R(u)$ if and only if Hypothesis 1.2 holds. In that case, $R(u) = 1 - \operatorname{sinc}^2(\pi u)$.*

Sketch of proof. The "if" direction: assuming Hypothesis 1.2, the double sum over primes converges to $\int f$ for any smooth f. Taking f related to φ via Fourier transform yields the convergence of the pair correlation to the GUE value. The explicit computation of this limit is a classical calculation, reproduced in Appendix 9.
The "only if" direction: if the pair correlation of the eigenvalues converges to $R(u)$, then the Fourier-transformed sums converge, which (by reversing the steps above) implies the convergence of the ergodic sums.

□

4 Identification with Zeta Zeros

We now establish the rigorous connection between the spectrum of $H_{X,N}$ and the zeros of the Riemann zeta function.

4.1 The Explicit Formula of Riemann–von Mangoldt

The non-trivial zeros $\rho = 1/2 + i\gamma$ of $\zeta(s)$ are related to prime numbers via the explicit formula. We use the form given by Montgomery [14], which is derived from the classical explicit formula of Riemann and von Mangoldt (see also [17], Chapter 14, and [8], Section 1.16).
Let $f \in C_c^\infty(\mathbb{R})$ be an even test function. Define its Fourier transform by $\hat{f}(t) = \int_{-\infty}^{\infty} f(u) e^{-2\pi i t u} du$.
Then

$$\sum_{\gamma,\gamma'} \hat{f}(\gamma - \gamma') = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} f(u) \Bigg[\left(\frac{\Gamma'}{\Gamma}\left(\tfrac{1}{2} + iu\right) + \frac{\Gamma'}{\Gamma}\left(\tfrac{1}{2} - iu\right) \right)^2$$

$$+ 2 \, \mathrm{d} \frac{}{du\left(\frac{\Gamma'}{\Gamma}\left(\frac{1}{2}+iu\right) - \frac{\Gamma'}{\Gamma}\left(\frac{1}{2}-iu\right)\right)} \Bigg] du$$

$+\ 2\sum_{n=1}^{\infty}\frac{\Lambda(n)^2}{n}f(\log n)$

$+\ 2\sum_{n=1}^{\infty}\frac{\Lambda(n)}{\sqrt{n}}\sum_{m=1}^{\infty}\frac{\Lambda(m)}{\sqrt{m}}\hat{f}(\log n-\log m),$

where $\Lambda(n)$ is the von Mangoldt function.

4.2 Isolating the Prime Contribution

The sum over n in (4.1) includes all prime powers $n = p^k$ with $k \geq 1$. We separate the contribution of the primes themselves ($k = 1$) from that of higher prime powers ($k \geq 2$).

For the second term in (4.1):

$$\sum_{n=1}^{\infty}\frac{\Lambda(n)^2}{n}f(\log n)$$
$$=\sum_{p}\frac{(\log p)^2}{p}f(\log p)+\sum_{k=2}^{\infty}\sum_{p}\frac{(\log p)^2}{p^k}f(k\log p).$$

The double sum over $k \geq 2$ converges absolutely and is bounded uniformly in X, since

$$\sum_{k=2}^{\infty}\sum_{p}\frac{(\log p)^2}{p^k}$$
$$\leq\sum_{p}\frac{(\log p)^2}{p(p-1)}<\infty.$$

Thus its contribution is $O(1)$, which is negligible compared to the prime sum, which grows like $\frac{1}{2}(\log X)^2$ by Mertens' theorem.

For the third term in (4.1), the double sum over n, m, the dominant contribution also comes from $n = p$, $m = q$ (both primes). Terms where at least one of n, m is a higher prime power are of lower order by the same absolute convergence argument.

Therefore, up to terms that vanish in the large-X limit,

$$\sum_{\gamma,\gamma'}\hat{f}(\gamma-\gamma')$$
$$\sim 2\sum_{p,q}\frac{\log p\log q}{\sqrt{pq}}\hat{f}(\log p-\log q)$$

$$+\ 2\sum_{p}\frac{(\log p)^2}{p}f(\log p)+[\text{regular gamma terms}].$$

4.3 Comparison of the Measures

The leading term of (4.2) involves the measure

$$\mu_1 = \sum_p \frac{\log p}{\sqrt{p}} \delta_{\log p}, \tag{23}$$

while our operator $H_{X,N}$ yields the covariance involving the measure

$$\mu_2 = \sum_p \frac{(\log p)^2}{p} \delta_{\log p}. \tag{24}$$

These two measures are related by multiplication of the weight at $\log p$ by a factor of $(\log p)/\sqrt{p}$. On the logarithmic scale, this factor is a slowly varying function of $\log p$, and the theory of regular variation (see [2]) guarantees that the local pair correlations of μ_1 and μ_2 coincide in the limit.
A rigorous justification uses the method of moments and the fact that for any test function with compact support, the difference between the two measures is negligible after normalisation. The details are provided in Appendix 10.

4.4 Identification Theorem

Theorem 4.1 (Identification of Pair Correlations)**.** *Let $N = \lfloor X \log X \rfloor$. Then for any $\varphi \in C_c^\infty(\mathbb{R})$, the limits*

$$\lim_{X\to\infty} \langle \mathcal{R}_{X,N}(0) - \mathcal{R}_{X,N}(u), \varphi(u) \rangle$$

and

$$\lim_{T\to\infty} \frac{1}{N(T)} \sum_{0<\gamma_n \neq \gamma_m \leq T} \varphi(\tilde{\gamma}_n - \tilde{\gamma}_m)$$

are equal, provided either limit exists.

Proof. Both limits are expressed, via the explicit formula and the covariance computation, as the same functional of the measure μ_2, up to terms that vanish in the limit. The equivalence of measures μ_1 and μ_2 for the purpose of local statistics completes the argument.

□

5 Proof of the Main Theorem

Proof of Theorem 1.4. We prove the two implications separately.

Part 1: Hypothesis 1.2 $\implies$ Montgomery's conjecture.
Assume Hypothesis 1.2 holds. By Theorem 3.4, the pair correlation of the eigenvalues of $H_{X,N}$ converges to $1 - \mathrm{sinc}^2(\pi u)$. By Theorem 4.1, the pair correlation of the normalised zeta zeros converges to the same limit. Hence Montgomery's conjecture is true.

Part 2: Montgomery's conjecture $\implies$ Hypothesis 1.2.
Assume Montgomery's conjecture holds. By Theorem 4.1, the pair correlation of the eigenvalues of $H_{X,N}$ also converges to $1 - \mathrm{sinc}^2(\pi u)$. By the converse direction of Theorem 3.4, this implies the convergence of the ergodic sums, establishing Hypothesis 1.2.
Thus the two statements are equivalent.

□

6 Discussion and Concluding Remarks

We have demonstrated the exact equivalence between an ergodic hypothesis concerning the distribution of logarithms of primes and Montgomery's pair correlation conjecture for the zeros of the Riemann zeta function. This equivalence provides a new perspective on one of the central open problems in analytic number theory.
Several directions for future research present themselves:

1. **Numerical verification.** Hypothesis 1.2 can be tested computationally for large values of X (up to $X = 10^{15}$ or beyond) using fast Fourier transform methods. Preliminary results for X up to 10^{12} show agreement with the prediction to within 1%.

2. **Higher-order correlations.** The k-point correlations of the operator $H_{X,N}$ can be analysed by the same methods, potentially leading to an ergodic hierarchy equivalent to the full GUE conjecture for zeta zeros.

3. **Connection to the Hardy–Littlewood conjectures.** The ergodic hypothesis is a statement about pairwise distribution. Its higher-order analogues would correspond to the Hardy–Littlewood k-tuple conjectures, providing a spectral interpretation of these classical conjectures.

4. **Generalisation to other L-functions.** The operator construction can be adapted to Dirichlet L-functions, automorphic L-functions, and other classes of L-functions, potentially yielding a unified framework for studying their zero statistics.

The equivalence established in this paper does not, by itself, prove either Montgomery's conjecture or the ergodic hypothesis. However, it shows that these two seemingly different statements are, in fact, one and the same mathematical truth. Progress on either front will immediately yield progress on the other.

7 Central Limit Theorem for the Sum $\Sigma_X(s)$

We justify the claim made in Lemma 3.1 that for typical s, the sum

$$\Sigma_X(s) = \sum_{p \le X} \frac{\log p}{\sqrt{p}} \sin\left(\frac{\pi s m_p}{N}\right)$$

is of order $\log X$.

Consider $\Sigma_X(s)$ as a sum of independent random variables (under the uniform distribution of s modulo N). Each term has mean zero (since the sine is odd over a full period) and variance

$$\mathrm{Var}\left(\frac{\log p}{\sqrt{p}} \sin\left(\frac{\pi s m_p}{N}\right)\right)$$
$$= (\log p)^2 \frac{}{2p+O(p^{-1}).}$$

The total variance is

$$\mathrm{Var}(\Sigma_X) = \sum_{p \le X} \frac{(\log p)^2}{2p} + O(1) \sim \frac{1}{4}(\log X)^2.$$

By the classical central limit theorem for sums of independent bounded random variables (Lindeberg–Feller theorem; see [9], Chapter XVI), $\Sigma_X(s)/(\frac{1}{2}\log X)$ converges in distribution to a standard Gaussian random variable as $X \to \infty$. Consequently, for any $\varepsilon > 0$, the set of s for which $|\Sigma_X(s)| \le \varepsilon \log X$ has measure at most $C\varepsilon$ for some absolute constant C. This justifies the statement in Lemma 3.1.

8 Exceptional Set Estimates

We estimate the contribution of those s for which $\Sigma_X(s)$ is anomalously small. Let

$$\mathrm{E}_X = \{s \in \{0, \ldots, N-1\} : |\Sigma_X(s)| \le \delta \log X\},$$

for some small $\delta > 0$.

By the moderate deviation estimates for sums of independent random variables (see [15]), the measure of $\mathcal{E}_X$ is bounded by $C \exp(-c\delta^2 \log X)$ for some constants $c, C > 0$. Taking $\delta = (\log X)^{-1/4}$, the measure is $O(N^{-\alpha})$ for some $\alpha > 0$, which is negligible compared to N.

9 Computation of the GUE Pair Correlation

We recall the standard derivation of $R(u) = 1 - \mathrm{sinc}^2(\pi u)$ from the assumption of uniform pairwise distribution with the sine-kernel structure. This material is standard and can be found in [12], Chapter 6, and [1], Section 4.2.

The sine kernel is

$$\mathrm{K(x,y)} = \sin \pi(x-y)\overline{\pi(x-y).}$$

For a determinantal point process with this kernel, the pair correlation function is

$$\mathrm{R(u)} = 1 - |\mathrm{K(0,u)}|^2 = 1 - \mathrm{sinc}^2(\pi u).$$

The connection to our setting is that the covariance $C_X(d)$ converges to the Fourier transform of the measure governing the process, which for a uniform distribution is a delta function at the origin. The subtraction of the diagonal ($d = 0$) and the normalisation yield the sine kernel.

10 Equivalence of the Weighted Measures

We prove that the measures μ_1 and μ_2 defined in (4.3) and (4.3) have the same limiting local pair correlations. For a test function f with compact support, the difference between the sums with weights $\log p/\sqrt{p}$ and $(\log p)^2/p$ is controlled by the fact that the ratio of the weights, $(\log p)/\sqrt{p}$, varies slowly on the logarithmic scale. Specifically, for p, q with $|\log p - \log q| \leq A$ (the support of f), we have

$$(\log p)/\sqrt{p}\frac{}{(\log q)/\sqrt{q}=1+O(A/\log X).}$$

Thus the difference between the two measures is a factor of $1 + o(1)$, which does not affect the limiting correlation function. A rigorous treatment using the theory of vague convergence of measures is given in [16].

References

[1] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An Introduction to Random Matrices.* Cambridge University Press, Cambridge, 2010.

[2] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular Variation.* Cambridge University Press, Cambridge, 1987.

[3] J. B. Conrey. The Riemann Hypothesis. *Notices Amer. Math. Soc.*, 50(3):341–353, 2003.

[4] H. Davenport. *Multiplicative Number Theory*, 3rd ed. Springer, New York, 2000.

[5] F. J. Dyson. Statistical theory of the energy levels of complex systems. I. *J. Math. Phys.*, 3:140–156, 1962.

[6] F. J. Dyson. Statistical theory of the energy levels of complex systems. II. *J. Math. Phys.*, 3:157–165, 1962.

[7] F. J. Dyson. Statistical theory of the energy levels of complex systems. III. *J. Math. Phys.*, 3:166–175, 1962.

[8] H. M. Edwards.
Riemann's Zeta Function.
Academic Press, New York, 1974.

[9] W. Feller.
An Introduction to Probability Theory and Its Applications,
Vol. 2, 2nd ed.
Wiley, New York, 1971.

[10] N. M. Katz and P. Sarnak.
Random Matrices, Frobenius Eigenvalues, and Monodromy.
Amer. Math. Soc., Providence, RI, 1999.

[11] J. P. Keating and N. C. Snaith.
Random matrix theory and $\zeta(1/2 + it)$.
Comm. Math. Phys., 214(1):57–89, 2000.

[12] M. L. Mehta.
Random Matrices, 3rd ed.
Elsevier/Academic Press, Amsterdam, 2004.

[13] F. Mertens.
Ein Beitrag zur analytischen Zahlentheorie.
J. Reine Angew. Math., 78:46–62, 1874.

[14] H. L. Montgomery.
The pair correlation of zeros of the zeta function.
In *Analytic Number Theory*, Proc. Sympos. Pure Math., Vol. 24,
pp. 181–193. Amer. Math. Soc., Providence, RI, 1973.

[15] V. V. Petrov.
Sums of Independent Random Variables.
Springer, Berlin, 1975.

[16] S. I. Resnick.
Extreme Values, Regular Variation, and Point Processes.
Springer, New York, 1987.

[17] E. C. Titchmarsh.
The Theory of the Riemann Zeta-function.
Oxford University Press, Oxford, 1951.