A Proposed Solution to Erdős Problem 1002

Shouqiao Wang

Columbia University Multiscalar Intelligence

Abstract

Let
$$S_N(\alpha) = \sum_{k=1}^{N} \left(\frac{1}{2} - \{k\alpha\}\right), \qquad \alpha \sim \mathrm{Unif}(0,1).$$
We prove that
$$\frac{S_N(\alpha)}{\log N} \Longrightarrow \mathrm{Cauchy}\left(0, \frac{1}{2\pi}\right).$$
Thus the distribution functions converge, at every real c, to
$$\frac{1}{2} + \frac{1}{\pi}\arctan(2\pi c).$$
The starting point of the rotation is fixed; no averaging in a second spatial coordinate or in time is used. The proof first reconstructs the sum in L^2, up to $o(\log N)$, as a primitive rational shot sum. A Ramanujan-sum square-function argument removes all nonresonant shots. The remaining signed, marked resonances converge by a continued-fraction rare-event theorem proved here. A cylinder oscillation argument supplies the growing mark without averaging the starting point. The resulting Poisson integral has the stated Cauchy scale. This proposed solution was found by GPT-5.6.

1 Introduction

Throughout, all logarithms are natural, and Leb denotes Lebesgue measure. Erdős asked [4, p. 63] whether the random variables
$$S_N(\alpha) := \sum_{k=1}^{N} \left(\frac{1}{2} - \{k\alpha\}\right), \qquad f_N(\alpha) := \frac{S_N(\alpha)}{\log N}, \qquad \alpha \sim \mathrm{Unif}(0,1),$$
possess an asymptotic distribution function along the full sequence of integers N, with convergence at every real threshold. We give an affirmative answer and identify the limit explicitly as a centered Cauchy law.

Theorem 1.1 (Main theorem)**.** *For every $c \in \mathbb{R}$,*
$$\lim_{N\to\infty} \mathrm{Leb}\left\{\alpha \in (0,1) : \frac{S_N(\alpha)}{\log N} \le c\right\} = g(c) := \frac{1}{2} + \frac{1}{\pi}\arctan(2\pi c).$$
Equivalently,
$$\frac{S_N}{\log N} \Longrightarrow X, \qquad \mathbb{E}e^{\mathrm{i}tX} = \exp\left(-\frac{|t|}{2\pi}\right).$$

Write

$$e(x) = \exp(2\pi \mathrm{i} x), \qquad \psi(x) = \frac{1}{2} - \{x\}, \qquad V(x) = \frac{\{x\}(1-\{x\})}{2}.$$

We write

$$\|u\|_{\mathbb{T}} := \min_{j \in \mathbb{Z}} |u - j|, \qquad \mathrm{Si}(x) := \int_0^x \frac{\sin t}{t} \, \mathrm{d}t,$$

with the integrand at zero defined by continuity. The symbol $\mathrm{nint}(t)$ denotes a nearest integer to t, with the smaller one chosen at a tie. Values at the finitely many rational discontinuities will never affect an integral or a distribution. We identify $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ with $[0, 1)$.

The distinction between Theorem 1.1 and the classical two-variable result is important. Kesten's spatial theorem averages both the rotation parameter and the starting point [7, 8]; here the starting point is exactly zero.

For $p \geq 1$, let

$$q_p = q_p(\alpha) := \mathrm{nint}(p\alpha), \qquad \delta_p = \delta_p(\alpha) := p\alpha - q_p \in (-1/2, 1/2],$$

put $L = L_N := \log N$, and impose $(p, q_p) = 1$. The central object is

$$Y_N(\alpha) = \sum_{\substack{1 \leq p \leq N \\ (p, q_p) = 1}} \frac{V(N\delta_p)}{p\delta_p}. \tag{1.1}$$

At the null set where $\delta_p = 0$, this quotient is assigned the value zero. This convention will be used for every shot sum below. The proof has three stages. The final transition-band step in stage (ii) invokes the marked Poisson theorem from stage (iii); that theorem is proved independently of stage (ii), so there is no circular dependence.

(i) In Section 2 we prove

$$\|S_N - Y_N\|_2 = o(\log N). \tag{1.2}$$

The proof is an exact Fourier–Ramanujan reconstruction, including its principal-value and endpoint conventions.

(ii) In Section 3 we prove the small-jump estimate: for every $\eta > 0$,

$$\lim_{A \to \infty} \limsup_{N \to \infty} \mathrm{Leb} \left\{ \frac{1}{\log N} \left| \sum_{\substack{p \leq N, \ (p, q_p) = 1 \\ (\log N)|p\delta_p| > A}} \frac{V(N\delta_p)}{p\delta_p} \right| > \eta \right\} = 0. \tag{1.3}$$

This is the fixed-start step which is absent from the usual two-variable argument.

(iii) In Section 4, the complementary finite process

$$\{((\log N)p\delta_p, \ N\delta_p \bmod 1) : p \leq N, \ (p, q_p) = 1\}$$

is identified with rare regular continued-fraction convergents. Exponential mixing proves the unmarked Poisson law; a two-scale cylinder argument proves that the marks become independent Haar variables. The limiting marked process has intensity $1/\zeta(2)$, and sending the compact cutoff to infinity gives the characteristic exponent $-|t|/(2\pi)$.

We use the Fourier convention

$$\widehat{f}(n) = \int_0^1 f(\alpha) e(-n\alpha) \, \mathrm{d}\alpha.$$

For an integrable function on $\mathbb{R}$, the corresponding real-line convention is

$$\widehat{f}(t) = \int_{\mathbb{R}} f(v) e(-tv) \, \mathrm{d}v.$$

All L^2-norms without a displayed measure are over $\mathbb{T}$. We use the standard arithmetic functions μ (Möbius), φ (Euler), τ (the divisor-counting function), and $\sigma_s(n) = \sum_{d|n} d^s$.

2 Primitive-shot reconstruction

2.1 Periodization and its Fourier transform

The absolutely convergent Fourier series

$$V(x)=\frac{1}{12}-\sum_{\ell\neq 0}\frac{e(\ell x)}{4\pi^2\ell^2}=\sum_{r\geq 1}\frac{1-\cos(2\pi r x)}{2\pi^2 r^2} \tag{2.1}$$

will provide the mark. For $P\geq 1$, define the nearest-cell sum

$$Y_{N,P}(\alpha)=\sum_{\substack{p\leq P\\(p,q_p)=1}}\frac{V(N\delta_p)}{p\delta_p}$$

and its full periodization

$$Z_{N,P}(\alpha)=\sum_{p\leq P}\frac{1}{p}\,\mathrm{PV}\sum_{\substack{q\in\mathbb{Z}\\(p,q)=1}}\frac{V(N(p\alpha-q))}{p\alpha-q}. \tag{2.2}$$

The symmetric principal value is taken in q.

Lemma 2.1 (Fourier transform)**.** *For $s>0$,*

$$I_N(s):=\mathrm{PV}\int_{\mathbb{R}}\frac{V(Nx)}{x}e(-sx)\,\mathrm{d}x=-\frac{\mathrm{i}}{2\pi}\sum_{k>s/N}\frac{1}{k^2}, \tag{2.3}$$

where the term $k=s/N$ has weight $1/2$ when s/N is an integer. Consequently, for $n>0$,

$$\widehat{Z}_{N,P}(n)=-\frac{\mathrm{i}}{2\pi}\sum_{p\leq P}\frac{c_p(n)}{p^2}H_*\left(\frac{n}{pN}\right),\qquad H_*(u)=\sum_{k>u}\frac{1}{k^2}+\frac{\mathbf{1}_{\{u\in\mathbb{N}\}}}{2u^2}, \tag{2.4}$$

where the last term is understood as zero at $u=0$, and

$$c_p(n)=\sum_{\substack{a \bmod p\\(a,p)=1}}e(an/p)$$

is the Ramanujan sum. Moreover,

$$\widehat{Z}_{N,P}(0)=0,\qquad \widehat{Z}_{N,P}(-n)=\overline{\widehat{Z}_{N,P}(n)}.$$

Proof. Put $h_N(x)=V(Nx)/x$, with an arbitrary value at $x=0$. Before computing, we make the periodization limit explicit. Let $\mathcal{A}_p$ be the reduced residues in $\{0,\dots,p-1\}$, and for $R\geq 1$ define the finite sum

$$Z^{(R)}_{N,P}(\alpha)=\sum_{p\leq P}\frac{1}{p}\sum_{a\in\mathcal{A}_p}\sum_{|j|\leq R}h_N(p\alpha-a-pj).$$

For $z=p\alpha-a\in[-p,p]$, periodicity of V gives, when $j\geq 2$,

$$h_N(z-pj)+h_N(z+pj)=V(Nz)\left(\frac{1}{z-pj}+\frac{1}{z+pj}\right),$$

and hence

$$\sup_{|z|\leq p}\left|\sum_{j>R}\left(h_N(z-pj)+h_N(z+pj)\right)\right|\ll\sum_{j>R}\frac{|z|}{p^2j^2}\ll\frac{1}{R}.$$

The finitely many apparent quotient singularities are canceled by the zeros of the numerator and leave only bounded one-sided jumps, because $V(Nx) = O_N(|x|)$ at an integer. Thus $Z_{N,P}^{(R)}$ converges in $L^2(\mathbb{T})$, in fact uniformly away from those finitely many jump points, to $Z_{N,P}$. Compare this residue-wise cutoff $|j| \le R$ with the literal symmetric cutoff $|q| \le pR$. In a fixed residue class $q = a + pj$, the two index intervals differ by at most two endpoint indices, and the corresponding summands are $O_p(R^{-1})$. Thus their difference is $O_p(R^{-1})$. For an arbitrary literal cutoff radius S, take $R = \lfloor S/p \rfloor$. The cutoffs $|q| \le S$ and $|q| \le pR$ differ by $O_p(1)$ endpoints, each of size $O_p(S^{-1})$. Hence the residue-wise convention and the symmetric principal value in q used in (2.2) have the same limit (with $p \le P$ fixed here).

To justify the principal-value calculation, first integrate over $\eta < |x| < R$, insert the absolutely and uniformly convergent series (2.1), and then let $\eta \downarrow 0$ symmetrically and $R \to \infty$. The first limit can be passed through the series because, uniformly near zero,

$$\sum_{k \ge 1} \frac{|1 - \cos(2\pi k N x)|}{k^2 |x|} \ll N;$$

split the sum at $k = (N|x|)^{-1}$. For the second limit, the symmetric truncated integrals of each exponential divided by x are bounded uniformly in its frequency and in R, by the elementary boundedness of the sine integral. Thus $\sum k^{-2} < \infty$ supplies domination. Equivalently one may insert $e^{-\epsilon|x|}$, retain the symmetric principal value at zero, and let $\epsilon \downarrow 0$; for every real a,

$$\lim_{\epsilon \downarrow 0} \mathrm{PV} \int_{\mathbb{R}} \frac{e(-ax)e^{-\epsilon|x|}}{x} \, \mathrm{d}x = -\mathrm{i}\pi \, \mathrm{sgn}(a),$$

with value zero at $a = 0$. The distributional identity

$$\mathrm{PV} \int_{\mathbb{R}} \frac{e(-sx)}{x} \, \mathrm{d}x = -\mathrm{i}\pi \, \mathrm{sgn}(s)$$

and the second series in (2.1) show that the k-th summand contributes $-\mathrm{i}/(2\pi k^2)$ precisely when $kN > s$, zero when $kN < s$, and half at equality. This proves (2.3).

We may now compute with the finite periodization. The changes of variables $t = p\alpha$ and $x = t - a - pj$ give, for every integer $n \ge 0$,

$$\widehat{Z}_{N,P}^{(R)}(n) = \sum_{p \le P} \frac{1}{p^2} \sum_{a \in \mathcal{A}_p} e(-na/p) \int_{-a-pR}^{p-a+pR} h_N(x) e(-nx/p) \, \mathrm{d}x.$$

Indeed, as j ranges from $-R$ to R, the corresponding intervals of length p are adjacent and tile the displayed interval. Replacing its endpoints by $\pm pR$ changes the integral by $O_p(R^{-1})$, since $|h_N(x)| \ll |x|^{-1}$ there. Hence, for $n > 0$, the integral tends to $I_N(n/p)$; for $n = 0$, it tends to zero because h_N is odd and the same endpoint estimate reduces the limit to a symmetric integral. Since

$$\sum_{a \in \mathcal{A}_p} e(-na/p) = c_p(-n) = c_p(n),$$

the L^2-limit $Z_{N,P}^{(R)} \to Z_{N,P}$ and continuity of Fourier coefficients on L^2 prove (2.4) and $\widehat{Z}_{N,P}(0) = 0$. Finally $Z_{N,P}$ is real, which gives the negative coefficients by conjugacy. □

2.2 The exact all-denominator identity

Define $Z_{N,\infty}$ to be the mean-zero $L^2(\mathbb{T})$ function whose positive Fourier coefficients are given by the right side of (2.4) with the sum over all $p \ge 1$, and whose negative coefficients are their conjugates. The series over p is absolutely convergent for each coefficient, since the divisor formula for Ramanujan sums gives

$$\sum_{p\geq 1}\frac{|c_p(n)|}{p^2}\leq \sum_{d|n}\frac{1}{d}\sum_{r\geq 1}\frac{1}{r^2}\ll \sigma_{-1}(n).$$

Lemma 2.2 below shows that the resulting positive coefficient is

$$O_N(n^{-1}),$$

and hence that the coefficient sequence is square summable. This also fixes the symmetric principal-value convention in the notation $Z_{N,\infty}$.

Lemma 2.2 (Möbius collapse)**.** *For every $n>0$,*

$$\sum_{p\geq 1}\frac{c_p(n)}{p^2}H_*\left(\frac{n}{pN}\right)=\frac{1}{n}\left(\sum_{\substack{d|n\\ d<N}} d+\frac{N}{2}\mathbf{1}_{\{N|n\}}\right). \tag{2.5}$$

Hence, in $L^2(\mathbb{T})$,

$$Z_{N,\infty}=S_N-\frac{1}{2}\psi(N\alpha). \tag{2.6}$$

Proof. Use

$$c_p(n)=\sum_{a|(p,n)} a\,\mu(p/a)$$

and write $p=ar$. The left side of (2.5) becomes

$$\sum_{a|n}\frac{1}{a}\sum_{\substack{r,k\geq 1\\ rk>n/(aN)}}\frac{\mu(r)}{r^2k^2}.$$

On grouping by $m=rk$,

$$\sum_{rk>X}\frac{\mu(r)}{r^2k^2}=\sum_{m>X}\frac{1}{m^2}\sum_{r|m}\mu(r)=\begin{cases}1, & X<1,\\ 0, & X>1,\end{cases}$$

with midpoint value $1/2$ at $X=1$. Replacing a by $d=n/a$ gives (2.5).

Since

$$\psi(x)=\sum_{m\neq 0}\frac{e(mx)}{2\pi \mathrm{i} m}\quad \text{in } L^2(\mathbb{T}),$$

the positive n-th coefficient of S_N is

$$-\frac{\mathrm{i}}{2\pi n}\sum_{\substack{d|n\\ d\leq N}} d.$$

Combining this with Lemma 2.1 and (2.5) yields exactly the coefficient of $S_N-\frac{1}{2}\psi(N\alpha)$. Negative coefficients are conjugate, and both sides have mean zero. □

2.3 Nearest cell versus full periodization

Proposition 2.3 (Window estimate)**.** *Uniformly for $N, P \geq 2$,*

$$\|Y_{N,P} - Z_{N,P}\|_2^2 \ll \log(2P)(1 + \log\log(e^e + NP)). \tag{2.7}$$

Proof. Let $q(t) = \mathrm{nint}(t)$. Möbius inversion and the cotangent principal-value identity give

$$\mathrm{PV}\sum_{\substack{q\in\mathbb{Z}\\(p,q)=1}} \frac{1}{t-q} = \sum_{d|p} \mu(d)\frac{\pi}{d}\cot\left(\frac{\pi t}{d}\right). \tag{2.8}$$

Put

$$L_d(u) = \frac{\mathbf{1}_{\{\|u\|_{\mathbb{T}}<1/(2d)\}}}{u - \mathrm{nint}(u)} - \pi\cot(\pi u),$$

where the first quotient is interpreted in the unique local lift. The poles cancel. Since $V(N(t - q)) = V(Nt)$, writing $p = dr$ in the difference between the nearest coprime pole and (2.8) yields

$$Y_{N,P} - Z_{N,P} = \sum_{dr\leq P} \frac{\mu(d)}{d^2 r} F_{N,d}(r\alpha), \qquad F_{N,d}(u) = V(Ndu)L_d(u). \tag{2.9}$$

For $m \neq 0$, direct integration on $(-1/(2d), 1/(2d))$ gives

$$\widehat{L}_d(m) = \mathrm{i}\pi\,\mathrm{sgn}(m) - 2\mathrm{i}\,\mathrm{Si}(\pi m/d), \qquad \left|\widehat{L}_d(m)\right| \ll \min\left(1, \frac{d}{|m|}\right), \tag{2.10}$$

and $\widehat{L}_d(0) = 0$. If v_ℓ are the coefficients in (2.1), then

$$\widehat{F}_{N,d}(m) = \sum_{\ell\in\mathbb{Z}} v_\ell \widehat{L}_d(m - \ell N d).$$

For fixed ℓ, the contribution to the n-th Fourier coefficient of (2.9) is

$$C_\ell(n) = \sum_{\substack{dr\leq P\\ r|n}} \frac{\mu(d)}{d^2 r}\widehat{L}_d(n/r - \ell N d). \tag{2.11}$$

We use the elementary estimate

$$\sum_{m\geq 1} \sigma_{-1}(m)\left|\widehat{L}_d(m - M)\right|^2 \ll d(1 + \log\log(e^e + |M| + d)). \tag{2.12}$$

Indeed, on $|m - M| \leq d$, the first bound in (2.10) and $\sigma_{-1}(m) \ll 1 + \log\log(e^e + m)$ give

$$\sum_{\substack{m\geq 1\\ |m-M|\leq d}} \sigma_{-1}(m)|\widehat{L}_d(m - M)|^2 \ll d(1 + \log\log(e^e + |M| + d)).$$

On the annulus $2^j d < |m - M| \leq 2^{j+1}d$, the second bound in (2.10) gives

$$\sum_{\substack{m\geq 1\\ 2^j d<|m-M|\leq 2^{j+1}d}} \sigma_{-1}(m)|\widehat{L}_d(m - M)|^2 \ll 2^{-j}d(1 + \log\log(e^e + |M| + 2^{j+1}d)).$$

Here

$$\log\log(e^e + |M| + 2^{j+1}d) \ll \log\log(e^e + |M| + d) + \log(j + 2),$$

and both $\sum_{j\geq 0} 2^{-j}$ and $\sum_{j\geq 0} 2^{-j}\log(j + 2)$ converge. Summing the annuli proves (2.12).

Apply Cauchy–Schwarz to (2.11) with the fixed weights $d^{-3/2}r^{-1}$. Since

$$\sum_{\substack{dr\le P\\ r|n}} d^{-3/2}r^{-1} \ll \sigma_{-1}(n),$$

this gives

$$|C_\ell(n)|^2 \ll \sigma_{-1}(n) \sum_{\substack{dr\le P\\ r|n}} d^{-5/2}r^{-1} \left|\widehat{L}_d(n/r-\ell Nd)\right|^2.$$

Sum over $n = rm$, use $\sigma_{-1}(rm) \le \sigma_{-1}(r)\sigma_{-1}(m)$, and then (2.12) with $M = \ell Nd$. We obtain

$$\sum_{n\ge 1} |C_\ell(n)|^2 \ll \log(2P)\big(1+\log\log(e^e+|\ell|NP)\big),$$

because $\sum_{r\le R}\sigma_{-1}(r)/r \ll \log(2R)$ and $\sum_d d^{-3/2} < \infty$. The negative frequencies are conjugate, while the zero coefficient vanishes because $V(Ndu)$ is even and $L_d(u)$ is odd. Finally,

$$\sum_{\ell\in\mathbb{Z}} |v_\ell|\sqrt{1+\log\log(e^e+|\ell|NP)} \ll \sqrt{1+\log\log(e^e+NP)}$$

by $|v_\ell| \ll (1+\ell^2)^{-1}$. Minkowski and Plancherel prove (2.7). □

2.4 The superlinear tail and the natural cutoff

Lemma 2.4 (Crude all-p tail)**.** *For $P \ge N$,*

$$\|Z_{N,\infty} - Z_{N,P}\|_2^2 \ll \frac{N}{P}\big(1+\log(NP)\big)^3. \tag{2.13}$$

Proof. Expanding $c_p(n)$, the positive n-th coefficient, apart from the fixed factor $-\mathrm{i}/(2\pi)$, is

$$A_P(n) = \sum_{d|n}\frac{1}{d}\sum_{r>P/d}\frac{\mu(r)}{r^2}H_*\left(\frac{n}{drN}\right).$$

Since $H_*(u) \ll \min(1,u^{-1})$, dropping the Möbius signs gives

$$|A_P(n)| \ll \begin{cases} \tau(n)/P, & n \le NP,\\ \dfrac{N\tau(n)}{n}\left(1+\log\dfrac{n}{NP}\right), & n > NP.\end{cases}$$

The first line follows by summing r^{-2} above P/d. For the second, split the r-sum at $n/(dN)$. Now use

$$\sum_{n\le X}\tau(n)^2 \ll X(1+\log X)^3$$

to obtain

$$\sum_{n\le NP}|A_P(n)|^2 \ll \frac{1}{P^2}\sum_{n\le NP}\tau(n)^2 \ll \frac{N}{P}\big(1+\log(NP)\big)^3.$$

For $2^jNP < n \le 2^{j+1}NP$, $j \ge 0$, the second pointwise bound similarly gives

$$\sum_{2^jNP<n\le 2^{j+1}NP}|A_P(n)|^2 \ll \frac{N}{P}2^{-j}(j+2)^2\big(1+\log(2^{j+1}NP)\big)^3.$$

The last expression is summable in j, with total $O((N/P)(1+\log(NP))^3)$. Summing the dyadic blocks proves (2.13). □

We next improve the denominator cutoff from $N(1+\log N)^{10}$ to N. The following elementary incomplete orthogonality will be used twice.

Lemma 2.5 (Incomplete Ramanujan orthogonality)**.** *If $p \neq p'$, then for every finite interval $I \subset \mathbb{Z}$,*

$$\left|\sum_{n\in I} c_p(n)c_{p'}(n)\right| \leq 2\sigma_1(p)\sigma_1(p'). \tag{2.14}$$

Proof. Insert

$$c_p(n) = \sum_{d|(p,n)} d\,\mu(p/d)$$

twice. The number of multiples of $\operatorname{lcm}(d,e)$ in I is $|I|/\operatorname{lcm}(d,e)+O(1)$. The main term vanishes when $p \neq p'$, by complete-period orthogonality of Ramanujan sums. Indeed, averaging their exponential definitions over one period $\operatorname{lcm}(p,p')$ leaves only pairs of reduced fractions satisfying $a/p + b/p' \in \mathbb{Z}$; reduction forces $p = p'$. The absolute sum of the endpoint errors is at most the right side of (2.14). □

Lemma 2.6 (Discrete Abel summation against a BV weight)**.** *Let (a_n) be indexed by $\mathbb{Z}$, or by a one-sided integer interval, and suppose that*

$$\sup_I \left|\sum_{n\in I} a_n\right| \leq M, \tag{2.15}$$

where the supremum is over all finite integer intervals in the index set. If W is a regulated function of bounded variation and tends to zero at every infinite end of the index set, then

$$\left|\sum_n a_n W(n)\right| \leq 2M(\|W\|_\infty + \mathrm{TV}_{\mathbb{R}}(W)). \tag{2.16}$$

The same conclusion holds with $\mathrm{TV}_{\mathbb{R}}(W)$ replaced by the discrete variation $\sum_n |W(n+1)-W(n)|$.

Proof. For integers $u \leq v$, put $A_u(k) = \sum_{u\leq n\leq k} a_n$. Ordinary summation by parts gives the exact identity

$$\sum_{n=u}^{v} a_n W(n) = A_u(v)W(v) + \sum_{n=u}^{v-1} A_u(n)(W(n) - W(n+1)). \tag{2.17}$$

By (2.15), its absolute value is at most

$$M\left(|W(v)| + \sum_{n=u}^{v-1} |W(n+1) - W(n)|\right).$$

On either infinite tail, both the endpoint value and the remaining variation tend to zero. Thus the finite sums satisfy the Cauchy criterion. Letting the endpoints tend to the ends of the index set in (2.17) proves (2.16); the displayed factor 2 harmlessly covers all one-sided and two-sided endpoint conventions. □

Proposition 2.7 (Natural denominator cutoff)**.** *As $N \to \infty$,*

$$\|Z_{N,\infty} - Z_{N,N}\|_2 = o(\log N), \tag{2.18}$$

and consequently

$$\boxed{\|S_N - Y_N\|_2 = o(\log N).} \tag{2.19}$$

Proof. Put $L = \log N$ and $P = N(1+L)^{10}$. For a dyadic block $Q < p \le 2Q$, $Q > N$, let

$$A_Q(n) = \sum_{Q<p\le 2Q} \frac{c_p(-n)}{p^2} H_*\left(\frac{n}{pN}\right), \qquad n > 0.$$

Write $h_p(n) = H_*(n/(pN))$. Expanding the square gives

$$\sum_{n\ge1} |A_Q(n)|^2 = \sum_{p,p'\asymp Q} \frac{1}{p^2p'^2} \sum_{n\ge1} c_p(-n)c_{p'}(-n)h_p(n)h_{p'}(n). \tag{2.20}$$

The midpoint convention in H_* preserves monotonicity. More precisely, at every integer $m \ge 1$,

$$H_*(m^-) - H_*(m) = H_*(m) - H_*(m^+) = \frac{1}{2m^2}.$$

Consequently each h_p is nonnegative and nonincreasing, tends to zero, and is bounded by $\zeta(2)$. Hence $h_p h_{p'}$ is also nonnegative and nonincreasing, and

$$\|h_p h_{p'}\|_\infty + \sum_{n\ge1} |h_p(n+1)h_{p'}(n+1) - h_p(n)h_{p'}(n)| \le 2\zeta(2)^2.$$

For $p \ne p'$, apply Lemma 2.6 to $a_n = c_p(-n)c_{p'}(-n)$, and use Lemma 2.5. Thus the inner sum in (2.20) is $O(\sigma_1(p)\sigma_1(p'))$, including every midpoint jump and every Abel boundary term. The off-diagonal contribution is therefore

$$\sum_{\substack{p,p'\asymp Q\\ p\ne p'}} \frac{\sigma_1(p)\sigma_1(p')}{p^2p'^2} \ll (\log\log(3Q))^2.$$

For the diagonal, use

$$\sum_{n \bmod p} c_p(n)^2 = p\varphi(p), \qquad \sum_{n\ge1} H_*(n/(pN))^2 \ll pN.$$

More explicitly, on the j-th period $jp < n \le (j+1)p$, monotonicity of H_* and complete-period orthogonality give a contribution at most

$$p\varphi(p)H_*(j/N)^2.$$

Since $H_*(u) \ll \min(1, u^{-1})$, $\sum_{j\ge0} H_*(j/N)^2 \ll N$. Thus subdivision into periods of length p gives

$$\sum_{p\asymp Q} \frac{1}{p^4} \sum_{n\ge1} c_p(n)^2 H_*(n/(pN))^2 \ll \sum_{p\asymp Q} \frac{N\varphi(p)}{p^3} \ll \frac{N}{Q}.$$

Thus

$$\|A_Q\|^2_{\ell^2(n\ge1)} \ll 1 + (\log\log(3Q))^2. \tag{2.21}$$

There are $O(\log(P/N)) = O(\log L)$ blocks. Minkowski, (2.21), and Lemma 2.4 give

$$\|Z_{N,\infty} - Z_{N,N}\|_2 \ll (\log L)^2 + o(1) = o(L),$$

which is (2.18).

Finally combine (2.6), Proposition 2.3 with $P = N$, and (2.18). The term $\frac{1}{2}\psi(N\alpha)$ has bounded L^2-norm, while

$$\|Y_{N,N} - Z_{N,N}\|_2 \ll \sqrt{L\log L} = o(L).$$

This proves (2.19). $\square$

3 The nonresonant shots

Throughout this section $L = \log N$, and $n \asymp K$ always means $K < n \leq 2K$, with K ranging over dyadic numbers. We first control the range

$$\frac{A}{L} < |p\delta_p| < \varepsilon$$

with a smooth cutoff, and then the fixed-away range $|p\delta_p| \geq \varepsilon$. The two arguments are different.

3.1 A mean-square Ramanujan tail

We use the following consequence of Chan and Kumchev's mean-square theorem for partial sums of Ramanujan sums [3, Theorem 1.2].

Lemma 3.1 (Ramanujan tail). *There is an absolute $B_{\text{tail}} > 30$ such that, uniformly for $K \geq 3$ and $2 \leq Q \leq \sqrt{K}$,*

$$\left\| \sum_{q>Q} \frac{c_q(n)}{q^2} \right\|_{\ell^2(1 \leq n \leq 2K)} \ll \frac{\sqrt{K}}{Q}. \tag{3.1}$$

The same bound holds for a finite tail $\sum_{Q<q\leq R}$, uniformly in R.

Proof. Put $C_x(n) = \sum_{q \leq x} c_q(n)$. Set

$$x_{\text{CK}} = x, \qquad y_{\text{CK}} = 2K, \qquad B_{\text{CK}} = 11.$$

When $x \leq \sqrt{y_{\text{CK}}}$, their elementary formula

$$\sum_{n \leq y_{\text{CK}}} |C_x(n)|^2 = \frac{y_{\text{CK}} x^2}{2\zeta(2)} + O(x^4 + x y_{\text{CK}} \log x)$$

is $O(Kx^2)$. When $\sqrt{y_{\text{CK}}} < x \leq T := K/(\log K)^{B_{\text{tail}}}$, apply part (ii) of their theorem with the displayed substitution. We verify both of its range hypotheses, including the lower logarithmic margin. Since $B_{\text{CK}} = 11$, $B_{\text{tail}} > 30$, and $x \leq T$,

$$x(\log x)^{2B_{\text{CK}}} \leq T(\log T)^{22} \leq K(\log K)^{22-B_{\text{tail}}} \leq 2K = y_{\text{CK}}$$

for all sufficiently large K. Since $x > \sqrt{y_{\text{CK}}}$, we also have

$$y_{\text{CK}} < x^2 \leq x^2(\log x)^{B_{\text{CK}}}.$$

The bounded remaining values of K are absorbed in the constant. Moreover, $\log(x y_{\text{CK}}) \asymp \log x \asymp \log K$ throughout this range. The function denoted by κ in that theorem satisfies $\sup_{u\in\mathbb{R}} |\kappa(u)| < \infty$, so its main term is $O(y_{\text{CK}} x^2) = O(Kx^2)$. Its error relative to $y_{\text{CK}} x^2$ is

$$O\Big((\log x)^{10} \left(x^{-1/2} + (y_{\text{CK}}/x)^{-1/2}\right)\Big) = O(1)$$

when $B_{\text{tail}} > 30$: indeed $x^{-1/2} \leq y_{\text{CK}}^{-1/4}$, and $y_{\text{CK}}/x \geq 2(\log K)^{B_{\text{tail}}}$ at the upper endpoint of the range. Thus

$$\sum_{n \leq 2K} |C_x(n)|^2 \ll Kx^2 \qquad (2 \leq x \leq K/(\log K)^{B_{\text{tail}}}). \tag{3.2}$$

Partial summation gives

$$\sum_{Q<q\leq R} \frac{c_q(n)}{q^2} = \frac{C_R(n)}{R^2} - \frac{C_Q(n)}{Q^2} + 2\int_Q^R \frac{C_x(n)}{x^3}\, \mathrm{d}x.$$

Minkowski and (3.2) bound the part below $T = K/(\log K)^{B_{\mathrm{tail}}}$ by $C\sqrt{K}/Q$. Above T, the divisor formula gives

$$|C_x(n)| \le x\tau(n), \qquad \|C_x\|_{\ell^2(n\le 2K)} \ll x\sqrt{K}(\log K)^{3/2}.$$

Its contribution is

$$\ll \frac{\sqrt{K}(\log K)^{3/2}}{T} = \frac{(\log K)^{B_{\mathrm{tail}}+3/2}}{\sqrt{K}},$$

and hence the argument has proved the more precise intermediate estimate

$$\left\| \sum_{q>Q} \frac{c_q(n)}{q^2} \right\|_{\ell^2(n\le 2K)} \ll \frac{\sqrt{K}}{Q} + E_K, \qquad E_K := \frac{(\log K)^{B_{\mathrm{tail}}+3/2}}{\sqrt{K}}, \quad 2 \le Q \le T. \tag{3.3}$$

The same estimate holds for finite tails, and the same divisor calculation also gives, uniformly in the upper endpoint,

$$\left\| \sum_{x<q\le R} \frac{c_q(n)}{q^2} \right\|_{\ell^2(n\le 2K)} \ll \frac{\sqrt{K}(\log K)^{3/2}}{x}, \qquad x \ge T. \tag{3.4}$$

For all sufficiently large K, $E_K = o(1)$, whereas $\sqrt{K}/Q \ge 1$ in the stated range; the bounded remaining values are absorbed into the constant. Letting $R \to \infty$ proves the infinite tail. Notice that the larger range $Q \le K/(\log K)^{B_{\mathrm{tail}}}$ would not follow from this argument: at its upper endpoint the displayed terminal error is too large by a factor $(\log K)^{3/2}$. The range stated here is exactly the one used below. $\square$

3.2 A smooth near-resonant estimate

Fix $0 < \varepsilon < 1/2$, $A \ge 1$, and put $a = A/L$. Fix Gevrey-order-two functions η, β on $[0,\infty)$, with

$$\eta = 0 \text{ on } [0,1], \quad \eta = 1 \text{ on } [2,\infty), \qquad \beta = 1 \text{ on } [0,1/2], \quad \beta = 0 \text{ on } [1,\infty),$$

and set

$$\rho(v) = \rho_{a,\varepsilon}(v) := \eta(|v|/a)\beta(|v|/\varepsilon).$$

Thus ρ is even, vanishes when $|v| \le a$ or $|v| \ge \varepsilon$, equals one when

$$2a \le |v| \le \varepsilon/2,$$

and satisfies

$$\left\|\rho^{(j)}\right\|_\infty \le C_\varepsilon^{j+1}(j!)^2 a^{-j}. \tag{3.5}$$

Define

$$Y^{\mathrm{near}}_{N,A,\varepsilon}(\alpha) = \sum_{\substack{p\le N\\ (p,q_p)=1}} \frac{V(N\delta_p)}{p\delta_p}\rho(p\delta_p). \tag{3.6}$$

Lemma 3.2 (Uniform inner-cutoff multipliers)**.** *Suppose that $0 < a \le \varepsilon/4$, and put*

$$W_a(v) = \begin{cases} \rho_{a,\varepsilon}(v)/v, & v \ne 0, \\ 0, & v = 0, \end{cases} \qquad J_a(t) = \int_{\mathbb{R}} W_a(v)e(-tv)\,\mathrm{d}v.$$

Then $W_a \in C_c^\infty(\mathbb{R})$. For every integer $m \ge 1$ and every $t \ne 0$,

$$|J_a(t)| + |tJ_a'(t)| \le C_\varepsilon^{m+1}(m!)^2 \min\left\{|t|, 1, (a|t|)^{-m}\right\}. \tag{3.7}$$

In particular, $\sup_{0<a\le\varepsilon/4, t\in\mathbb{R}} |J_a(t)| \ll_\varepsilon 1$. *Moreover,*

$$\int_0^\infty \sup_{1\le r\le 2} r|J_a'(rt)|\,\mathrm{d}t \ll_\varepsilon 1, \tag{3.8}$$

uniformly for $0 < a \le \varepsilon/4$*, and, for every integer* $j \ge 0$*,*

$$\left\|W_a^{(j)}\right\|_{L^2(\mathbb{R})} \le C_\varepsilon^{j+1}(j!)^2 a^{-j-1/2}. \tag{3.9}$$

All constants displayed here are independent of a*.*

Proof. Set

$$z_\varepsilon(v) = \beta(|v|/\varepsilon), \qquad h(u) = 1-\eta(|u|), \qquad h_a(v) = h(v/a).$$

Because $a \le \varepsilon/4$, the support of h_a is contained in $\{|v| \le 2a\}$, where $z_\varepsilon = 1$. On that support, $z_\varepsilon - h_a = \eta(|v|/a) = \rho_{a,\varepsilon}$; off that support, $h_a = 0$ and $\eta(|v|/a) = 1$. Hence

$$\rho_{a,\varepsilon} = z_\varepsilon - h_a. \tag{3.10}$$

Both terms on the right are fixed smooth compactly supported profiles, the second followed by dilation by a.

For every fixed integer $M \ge 2$, integration by parts M times and scaling give, for $t \ge 0$,

$$\sup_{1\le r\le 2} |\widehat{\rho}_{a,\varepsilon}(rt)| \le C_{\varepsilon,M}\left\{\varepsilon(1+\varepsilon t)^{-M} + a(1+at)^{-M}\right\}. \tag{3.11}$$

Since

$$J_a'(t) = -2\pi\mathrm{i}\,\widehat{\rho}_{a,\varepsilon}(t),$$

the integral of the right side of (3.11) over $t > 0$ is $O_{\varepsilon,M}(1)$. This proves (3.8) with the requested explicit envelope.

We next prove (3.7). The function $\rho_{a,\varepsilon}$ is even, so W_a is odd and

$$J_a(t) = -2\mathrm{i}\int_0^\infty \rho_{a,\varepsilon}(v)\frac{\sin(2\pi tv)}{v}\,\mathrm{d}v.$$

Since

$$\frac{\sin(2\pi tv)}{v}\,\mathrm{d}v = \mathrm{d}\,\mathrm{Si}(2\pi tv),$$

Stieltjes integration by parts yields

$$|J_a(t)| \le 2\sup_{x\in\mathbb{R}}|\,\mathrm{Si}(x)|\,\mathrm{TV}_{(0,\infty)}(\rho_{a,\varepsilon}) \ll_\varepsilon 1.$$

The total variation is uniform in a, either directly from the two transition regions or from (3.10). Similarly, Fourier integration by parts gives

$$|tJ_a'(t)| = 2\pi|t\widehat{\rho}_{a,\varepsilon}(t)| \le \left\|\rho_{a,\varepsilon}'\right\|_1 \ll_\varepsilon 1.$$

At the origin, $J_a(0) = 0$ by oddness and

$$|J_a'(t)| \le 2\pi\left\|\rho_{a,\varepsilon}\right\|_1 \ll_\varepsilon 1.$$

Thus

$$|J_a(t)| + |tJ_a'(t)| \ll_\varepsilon |t| \qquad (|t| \le 1).$$

It remains to record the high-frequency estimates with their dependence on the order. Leibniz's rule, the Gevrey bounds for the two transition profiles, and their supports imply, for $m \geq 1$,

$$\left\| W_a^{(m)} \right\|_1 \leq C_\varepsilon^{m+1} (m!)^2 a^{-m}, \qquad \left\| \rho_{a,\varepsilon}^{(m+1)} \right\|_1 \leq C_\varepsilon^{m+1} (m!)^2 a^{-m}. \tag{3.12}$$

For the first estimate, expand

$$W_a^{(m)} = \sum_{k=0}^{m} \binom{m}{k} \rho_{a,\varepsilon}^{(k)} \left(\frac{1}{v} \right)^{(m-k)}.$$

For $k = 0$, integration over $a < |v| < \varepsilon$ gives $O(m! a^{-m})$. For $k \geq 1$, the derivative of ρ is supported in

$$a < |v| < 2a, \qquad \varepsilon/2 < |v| < \varepsilon.$$

The inner region contributes at most

$$C_\varepsilon^{m+1} (m!)^2 a^{-k} a^{-(m-k+1)} a \ll C_\varepsilon^{m+1} (m!)^2 a^{-m},$$

and the outer region is smaller. The second estimate in (3.12) follows in the same way by scaling. Integrating by parts m times in J_a, and $m+1$ times in $\widehat{\rho}$, now gives

$$|J_a(t)| + |t J_a'(t)| \leq C_\varepsilon^{m+1} (m!)^2 (a|t|)^{-m}.$$

Together with the small- and intermediate-frequency bounds, this proves (3.7).

The same Leibniz expansion in L^2 gives

$$\left\| \rho_{a,\varepsilon}^{(k)} \left(\frac{1}{v} \right)^{(j-k)} \right\|_2 \leq C_\varepsilon^{j+1} (j!)^2 a^{-j-1/2}.$$

For $k = 0$, integrate $|v|^{-2(j+1)}$ from a to ε. For $k \geq 1$, the inner transition region contributes

$$a^{-k} a^{-(j-k+1)} a^{1/2} = a^{-j-1/2},$$

and the outer region is smaller. Summing over k proves (3.9). □

Proposition 3.3 (Near-resonant square function)**.** *For fixed $A \geq 1$ and $0 < \varepsilon < 1/2$,*

$$\limsup_{N \to \infty} \frac{\left\| Y_{N,A,\varepsilon}^{\text{near}} \right\|_2^2}{L^2} \ll_\varepsilon \frac{1}{A}. \tag{3.13}$$

Proof. For fixed A, ε, the relation $a = A/L \leq \varepsilon/4$ holds for all sufficiently large N, which is the range relevant to the limsup. Put $W = W_a$, $J = J_a$, and

$$b_p(\alpha) = \mathbf{1}_{\{(p,q_p)=1\}} \frac{\rho(p\delta_p)}{p\delta_p}.$$

For each p, there is the exact locally finite periodization

$$b_p(\alpha) = \sum_{\substack{q \in \mathbb{Z} \\ (p,q)=1}} W\left(p^2 \left(\alpha - \frac{q}{p} \right) \right). \tag{3.14}$$

Indeed, the support of the summand centered at q/p has radius ε/p^2. Distinct centers are separated by $1/p$, and $2\varepsilon/p^2 < 1/p$. Thus the supports are disjoint. On such a support, $|p\alpha - q| < \varepsilon/p < 1/2$, so $q = q_p$, and $p^2(\alpha - q/p) = p\delta_p$. Since $W \in C_c^\infty(\mathbb{R})$ and vanishes with all derivatives at its cutoff boundaries, (3.14) also proves that b_p is a periodic C^∞ function. Restricting to reduced q's removes entire disjoint smooth cells and creates no jump.

Changing variables in those cells gives, for every $n \in \mathbb{Z}$,

$$\begin{aligned}\widehat{b}_p(n) &= \frac{1}{p^2} \sum_{\substack{q \bmod p \\ (p,q)=1}} e(-nq/p) \int_{\mathbb{R}} W(v) e(-nv/p^2)\, \mathrm{d}v \\ &= \frac{c_p(-n)}{p^2} J(n/p^2).\end{aligned} \tag{3.15}$$

In particular $J(0) = 0$, since W is odd.

We begin with the constant coefficient $v_0 = 1/12$ in (2.1). Here and below in this proof, $n \asymp K$ means $K < n \le 2K$. Put

$$Q_K = \sqrt{aK}, \qquad R_K = \sqrt{\varepsilon K}.$$

Lemma 3.2 and

$$p \frac{\mathrm{d}}{\mathrm{d}p} J(n/p^2) = -2\frac{n}{p^2} J'(n/p^2)$$

give, for every fixed integer $m \ge 1$,

$$|J(n/p^2)| + p \left| \frac{\mathrm{d}}{\mathrm{d}p} J(n/p^2) \right| \ll_{m,\varepsilon} \begin{cases} (p/Q_K)^{2m}, & p < Q_K, \\ 1, & Q_K \le p \le R_K, \\ K/p^2, & p > R_K. \end{cases} \tag{3.16}$$

For the last line, use the small-t part of (3.7) when $n/p^2 \le 1$. If $1 < n/p^2 \le 2/\varepsilon$, the uniform bound is at most $C_\varepsilon K/p^2$: indeed $p^2 < n \le 2K$, so $K/p^2 \ge 1/2$, and the uniform $O_\varepsilon(1)$ bound is absorbed by the displayed quantity. We now write every boundary term in the multiplier summation. Let

$$\mathcal{H}_K = \ell^2\{n \in \mathbb{Z} : K < n \le 2K\}, \qquad u_p = \left(\frac{c_p(-n)}{p^2}\right)_{n \asymp K}, \qquad w_x = \left(J(n/x^2)\right)_{n \asymp K}.$$

For $2 \le U < \infty$, define the finite tail

$$\mathcal{R}_{x,U} = \sum_{x < p \le U} u_p.$$

Writing $\odot$ for coordinatewise multiplication, summation by parts in each coordinate gives the exact finite-dimensional identity

$$\sum_{2 < p \le U} u_p \odot w_p = \mathcal{R}_{2,U} \odot w_2 + \int_2^U \mathcal{R}_{x,U} \odot \frac{\mathrm{d}w_x}{\mathrm{d}x}\, \mathrm{d}x. \tag{3.17}$$

Indeed, $x \mapsto \mathcal{R}_{x,U}$ has jump $-u_p$ at p and $\mathcal{R}_{U,U} = 0$, so Stieltjes integration by parts is

$$-\int_{(2,U]} w_x \odot \mathrm{d}\mathcal{R}_{x,U} = \mathcal{R}_{2,U} \odot w_2 + \int_2^U \mathcal{R}_{x,U} \odot \mathrm{d}w_x.$$

This proves (3.17), including both endpoint terms.

Let

$$T = \frac{K}{(\log K)^{B_{\text{tail}}}}, \qquad E_K = \frac{(\log K)^{B_{\text{tail}}+3/2}}{\sqrt{K}}.$$

Since Ramanujan sums are real and even, the finite-tail form of (3.3) gives, uniformly for $2 \le x \le U \le T$,

$$\|\mathcal{R}_{x,U}\|_{\mathcal{H}_K} \ll \frac{\sqrt{K}}{x} + E_K. \tag{3.18}$$

For a fixed $m \geq 1$, put

$$\omega_{K,m}(x) = \begin{cases} (x/Q_K)^{2m}, & x < Q_K, \\ 1, & x \geq Q_K. \end{cases}$$

Equation (3.16) yields

$$\|w_x\|_{\ell^\infty(n \asymp K)} + x \left\| \frac{\mathrm{d}w_x}{\mathrm{d}x} \right\|_{\ell^\infty(n \asymp K)} \ll_{\varepsilon,m} \omega_{K,m}(x). \tag{3.19}$$

The uniform envelope in Lemma 3.2 also gives

$$\int_0^\infty \left\| \frac{\mathrm{d}w_x}{\mathrm{d}x} \right\|_{\ell^\infty(n \asymp K)} \mathrm{d}x \ll_\varepsilon 1. \tag{3.20}$$

To check the change of variables, write $n = rK$, $1 \leq r \leq 2$, and $t = K/x^2$; the left side is at most a constant times

$$\int_0^\infty \sup_{1 \leq r \leq 2} r|J'(rt)| \, \mathrm{d}t,$$

which is (3.8).

Assume first $K \geq 16/a$, so $Q_K \geq 4$, and set $U_0 = \min\{N, T\}$. Equations (3.17)–(3.20) give

$$\begin{aligned} \left\| \sum_{2 < p \leq U_0} u_p \odot w_p \right\|_{\mathcal{H}_K} &\leq \left(\frac{\sqrt{K}}{2} + E_K \right) \|w_2\|_\infty + \int_2^{U_0} \left(\frac{\sqrt{K}}{x} + E_K \right) \left\| \frac{\mathrm{d}w_x}{\mathrm{d}x} \right\|_\infty \mathrm{d}x \\ &\ll_{\varepsilon,m} \sqrt{K} \left\{ \frac{\omega_{K,m}(2)}{2} + \int_2^\infty \frac{\omega_{K,m}(x)}{x^2} \, \mathrm{d}x \right\} + E_K \\ &\ll_{\varepsilon,m} \frac{\sqrt{K}}{Q_K} + E_K. \end{aligned} \tag{3.21}$$

The last line follows from

$$Q_K^{-2m} \int_2^{Q_K} x^{2m-2} \, \mathrm{d}x + \int_{Q_K}^\infty x^{-2} \, \mathrm{d}x \ll_m Q_K^{-1}, \qquad \omega_{K,m}(2) \ll_m Q_K^{-1}.$$

If $T < N$, apply (3.17) on $(T, N]$. The terminal tail estimate (3.4) and (3.19) give

$$\begin{aligned} \left\| \sum_{T < p \leq N} u_p \odot w_p \right\|_{\mathcal{H}_K} &\ll_\varepsilon \sqrt{K} (\log K)^{3/2} \left(\frac{1}{T} + \int_T^\infty \frac{\mathrm{d}x}{x^2} \right) \\ &\ll_\varepsilon \frac{\sqrt{K} (\log K)^{3/2}}{T} = E_K. \end{aligned} \tag{3.22}$$

The finitely many terms $p \leq 2$ satisfy the same $O_{\varepsilon,m}(\sqrt{K}/Q_K)$ bound by the first line of (3.16). Consequently, whenever $K \geq 16/a$, uniformly over the dyadic block,

$$\left\| \sum_{p \leq N} \frac{c_p(n)}{p^2} J(n/p^2) \right\|_{\ell^2(n \asymp K)} \ll_\varepsilon \frac{\sqrt{K}}{Q_K} + E_K \ll_\varepsilon a^{-1/2}. \tag{3.23}$$

The assertion about E_K is uniform, not merely pointwise. Indeed, for all sufficiently large N, the function $x \mapsto (\log x)^{B_{\text{tail}}+3/2} x^{-1/2}$ is decreasing throughout $x \geq 16/a$, and therefore

$$\sup_{K \geq 16/a} E_K \leq C\sqrt{a} \left(\log(16/a) \right)^{B_{\text{tail}}+3/2} = o(1). \tag{3.24}$$

We spell out both ends of the frequency sum. For $1 \le n \le 16/a$, absolute convergence, the uniform bound for J, and the divisor formula give

$$\left| \sum_{p \le N} \frac{c_p(-n)}{p^2} J(n/p^2) \right| \ll \sum_{p \ge 1} \frac{|c_p(n)|}{p^2} \ll \sigma_{-1}(n). \tag{3.25}$$

Consequently, using $\sigma_{-1}(n) \ll 1 + \log\log(3n)$, these frequencies contribute at most

$$O_\varepsilon\left(a^{-1}(1 + \log\log(1/a))^2\right) = o(L/a) \tag{3.26}$$

to the squared norm.

At the other endpoint suppose that $Q_K > N$, equivalently $K > N^2/a$. For all sufficiently large N, this implies $N < T$, because $x/(\log x)^{B_{\rm tail}}$ is then increasing for $x \ge N^2/a$, and

$$T \ge \frac{N^2/a}{(\log(N^2/a))^{B_{\rm tail}}} > N.$$

Apply (3.17) with $U = N$. Every $x \le N$ then lies below Q_K, and hence

$$\|w_x\|_\infty + x\,\|w_x'\|_\infty \ll_{\varepsilon,m} (x/Q_K)^{2m}. \tag{3.27}$$

Using (3.18) in the exact Abel identity gives

$$\begin{aligned} \left\| \sum_{2<p\le N} u_p \odot w_p \right\|_{\mathcal{H}_K} &\ll_{\varepsilon,m} \sqrt{K} Q_K^{-2m}\left(1 + \int_2^N x^{2m-2}\,\mathrm{d}x\right) + E_K \left(\frac{N}{Q_K}\right)^{2m} \\ &\ll_{\varepsilon,m} \frac{\sqrt{K}}{Q_K}\left(\frac{N}{Q_K}\right)^{2m-1} + E_K\left(\frac{N}{Q_K}\right)^{2m}. \end{aligned} \tag{3.28}$$

The terms $p \le 2$ obey the same first bound directly. By (3.24), $E_K \le a^{-1/2} = \sqrt{K}/Q_K$ for all large N. Since $N/Q_K < 1$, we conclude that

$$\left\| \sum_{p \le N} \frac{c_p(-n)}{p^2} J(n/p^2) \right\|_{\ell^2(n \asymp K)} \ll_{\varepsilon,m} a^{-1/2}\left(\frac{N}{Q_K}\right)^{2m-1}. \tag{3.29}$$

On the dyadic blocks $K_r = 2^r N^2/a$, the square of the last bound is at most

$$\frac{1}{a}\left(\frac{N^2}{aK_r}\right)^{2m-1} = \frac{1}{a}\cdot 2^{-r(2m-1)}.$$

Thus the complete squared tail $K > N^2/a$ is $O_{\varepsilon,m}(a^{-1})$. Between $16/a$ and N^2/a there are $O(L)$ dyadic blocks, and each contributes $O_\varepsilon(a^{-1})$ by (3.23). Together with (3.26), this proves that the constant mode has squared norm

$$\ll_\varepsilon L/a = \frac{L^2}{A}. \tag{3.30}$$

Negative frequencies contribute the same amount by conjugation, and the zero frequency vanishes because $J(0) = 0$.

Now fix $\ell \ne 0$. Since N, ℓ, q_p are integers,

$$e(\ell N \delta_p) = e(\ell N p \alpha).$$

Differentiating the disjoint-cell representation (3.14) and changing variables in every cell gives, for every integer $j \geq 0$,

$$\begin{aligned} \left\| b_p^{(j)} \right\|_2^2 &= \varphi(p) p^{4j-2} \left\| W^{(j)} \right\|_2^2, \\ \left\| b_p^{(j)} \right\|_2 &\leq C_\varepsilon^{j+1} (j!)^2 p^{2j-1/2} a^{-j-1/2} \\ &= C_\varepsilon^{j+1} (j!)^2 \left(\frac{p^2}{a} \right)^j (ap)^{-1/2}. \end{aligned} \tag{3.31}$$

In particular,

$$\|b_p\|_2^2 \ll_\varepsilon \frac{1}{ap}. \tag{3.32}$$

Whenever $b_p(\alpha) \neq 0$,

$$\left| \alpha - \frac{q_p}{p} \right| = \frac{|p\delta_p|}{p^2} < \frac{1}{2p^2},$$

and $(p, q_p) = 1$. Legendre's criterion therefore makes p a continued-fraction denominator of α. Since such denominators satisfy $Q_{n+2} \geq 2Q_n$, only an absolute bounded number of active p's can lie in a dyadic interval. Let

$$P_0 = N \exp(-L^{1/3}).$$

The same recurrence shows that the total number of active denominators in $(P_0, N]$ is $O(L^{1/3})$. Pointwise Cauchy–Schwarz and (3.32) therefore give

$$\begin{aligned} \left\| \sum_{P_0 < p \leq N} b_p(\alpha) e(\ell N p \alpha) \right\|_2^2 &\ll L^{1/3} \sum_{P_0 < p \leq N} \|b_p\|_2^2 \\ &\ll \frac{L^{1/3}}{a} \sum_{P_0 < p \leq N} \frac{1}{p} \ll \frac{L^{2/3}}{a} = o(L/a). \end{aligned} \tag{3.33}$$

For $p \leq P_0$, divide into dyadic blocks $\mathcal{P}_s = (P_s/2, P_s]$, and set

$$F_{s,\ell}(\alpha) = \sum_{p \in \mathcal{P}_s} b_p(\alpha) e(\ell N p \alpha).$$

The same pointwise sparsity and (3.32) give

$$\|F_{s,\ell}\|_2^2 \ll \sum_{p \in \mathcal{P}_s} \|b_p\|_2^2 \ll_\varepsilon \frac{1}{a}. \tag{3.34}$$

Let $\Pi_{s,\ell}$ be Fourier projection onto the signed annulus

$$\mathcal{A}_{s,\ell} = \left\{ n \in \mathbb{Z} : \ \operatorname{sgn}(n) = \operatorname{sgn}(\ell), \quad \frac{|\ell| N P_s}{4} \leq |n| \leq 2|\ell| N P_s \right\}.$$

We use the following exact leakage inequality. If $f \in H^j(\mathbb{T})$, $m_0 \in \mathbb{Z}$, and every integer outside a frequency set $\mathcal{A}$ has distance at least D from m_0, then Plancherel gives

$$\begin{aligned} \|(I - \Pi_{\mathcal{A}})(e(m_0\alpha) f)\|_2^2 &= \sum_{\substack{k \in \mathbb{Z} \\ m_0 + k \notin \mathcal{A}}} |\widehat{f}(k)|^2 \\ &\leq D^{-2j} \sum_{k \in \mathbb{Z}} |k|^{2j} |\widehat{f}(k)|^2 = (2\pi D)^{-2j} \left\| f^{(j)} \right\|_2^2. \end{aligned} \tag{3.35}$$

For $p \in (P_s/2, P_s]$, the carrier $m_0 = \ell N p$ lies in $\mathcal{A}_{s,\ell}$, and its distance from the complement is at least $D_{s,\ell} = |\ell| N P_s/4$. Applying (3.35), the triangle inequality, and (3.31), we obtain

$$
\begin{aligned}
\|(I - \Pi_{s,\ell})F_{s,\ell}\|_2 &\le (2\pi D_{s,\ell})^{-j} \sum_{p \in \mathcal{P}_s} \left\| b_p^{(j)} \right\|_2 \\
&\le C_\varepsilon^{j+1} (j!)^2 |\ell|^{-j} \left(\frac{P_s}{Na} \right)^j \left(\frac{P_s}{a} \right)^{1/2}.
\end{aligned}
\tag{3.36}
$$

There is no unrecorded dependence on j in this estimate. Choose $j = \left\lceil 2L^{2/3} \right\rceil$. Since

$$
\log \frac{P_s}{Na} \le -L^{1/3} + \log(L/A),
$$

$|\ell|^{-j} \le 1$, and $\log(j!) \le j \log j$, the logarithm of the last line of (3.36) is at most

$$
-jL^{1/3} + j\log(L/A) + 2j\log j + O_\varepsilon(j) + \frac{1}{2}\log(P_s/a) = -\frac{3}{2}L + O_\varepsilon(L^{2/3}\log L).
$$

It follows that

$$
\sup_{s,\, \ell \neq 0} \|(I - \Pi_{s,\ell})F_{s,\ell}\|_2 = o(N^{-1}). \tag{3.37}
$$

The annuli $\mathcal{A}_{s,\ell}$, with P_s dyadic, have overlap multiplicity at most an absolute constant B. Consequently, frequency by frequency and then by Plancherel,

$$
\begin{aligned}
\left\| \sum_s \Pi_{s,\ell} F_{s,\ell} \right\|_2^2 &= \sum_{n \in \mathbb{Z}} \left| \sum_{s : n \in \mathcal{A}_{s,\ell}} \widehat{F}_{s,\ell}(n) \right|^2 \\
&\le B \sum_s \sum_{n \in \mathcal{A}_{s,\ell}} |\widehat{F}_{s,\ell}(n)|^2 \le B \sum_s \|F_{s,\ell}\|_2^2 \ll_\varepsilon L/a.
\end{aligned}
\tag{3.38}
$$

There are $O(L)$ blocks, so (3.37) also gives

$$
\sum_s \|(I - \Pi_{s,\ell})F_{s,\ell}\|_2 = o(1).
$$

The triangle inequality and (3.38) therefore give

$$
\left\| \sum_{p \le P_0} b_p(\alpha) e(\ell N p \alpha) \right\|_2^2 \ll_\varepsilon L/a, \tag{3.39}
$$

uniformly for $\ell \neq 0$. Together with (3.33), this is uniform in $\ell \neq 0$. Finally, the absolutely convergent Fourier series $V(x) = \sum_{\ell \in \mathbb{Z}} v_\ell e(\ell x)$ gives

$$
Y_{N,A,\varepsilon}^{\text{near}} = \sum_{\ell \in \mathbb{Z}} v_\ell \sum_{p \le N} b_p(\alpha) e(\ell N \delta_p).
$$

Use (3.30) for $\ell = 0$, the shifted estimate for $\ell \neq 0$, Minkowski's inequality, and $\sum_{\ell \neq 0} |v_\ell| < \infty$. We obtain

$$
\left\| Y_{N,A,\varepsilon}^{\text{near}} \right\|_2^2 \ll_\varepsilon \frac{L}{a} = \frac{L^2}{A},
$$

which proves (3.13). $\square$

3.3 The fixed-away estimate

Proposition 3.4 (Fixed-away shots)**.** *For every fixed* $0 < \varepsilon < 1/2$*,*

$$\Bigg\| \sum_{\substack{p \le N,\ (p,q_p)=1 \\ |p\delta_p| \ge \varepsilon}} \frac{V(N\delta_p)}{p\delta_p} \Bigg\|_2 = o(\log N). \tag{3.40}$$

More precisely, for every $0 < \varepsilon_0 < \varepsilon_1 < 1/2$*,*

$$\lim_{N\to\infty} \sup_{t\in[\varepsilon_0,\varepsilon_1]} \frac{1}{\log N} \Bigg\| \sum_{\substack{p \le N,\ (p,q_p)=1 \\ |p\delta_p| \ge t}} \frac{V(N\delta_p)}{p\delta_p} \Bigg\|_2 = 0. \tag{3.41}$$

Proof. Fix temporarily $0 < \delta < \varepsilon/2$, and take an even smooth function $\chi = \chi_{\varepsilon,\delta}$ such that

$$\chi(v) = 0 \quad (|v| \le \varepsilon - \delta), \qquad \chi(v) = 1 \quad (|v| \ge \varepsilon).$$

Let Y_N^χ be (1.1) with the extra factor $\chi(p\delta_p)$, and define its full periodization by

$$Z_N^\chi(\alpha) = \sum_{p\le N} \frac{1}{p} \operatorname{PV} \sum_{\substack{q\in\mathbb{Z} \\ (p,q)=1}} \frac{V(N(p\alpha - q))}{p\alpha - q} \chi(p(p\alpha - q)).$$

Because $1-\chi$ is supported in $|p(p\alpha - q)| < \varepsilon < 1/2$, there is at most one q in its support, necessarily q_p. Therefore

$$Y_N^\chi - Z_N^\chi = Y_N - Z_{N,N}, \tag{3.42}$$

and Proposition 2.3 gives

$$\|Y_N^\chi - Z_N^\chi\|_2^2 \ll L \log L = o(L^2). \tag{3.43}$$

Set

$$R_\chi(t) = \operatorname{PV} \int_{\mathbb{R}} \frac{\chi(v)}{v} e(-tv)\, \mathrm{d}v. \tag{3.44}$$

Put $\kappa_\chi = 1-\chi$. This function is even and compactly supported. For $t \ne 0$, the Fourier transform of $\operatorname{PV}(1/v)$ gives

$$R_\chi(t) = -\mathrm{i}\pi \operatorname{sgn}(t) - \operatorname{PV} \int_{\mathbb{R}} \frac{\kappa_\chi(v)}{v} e(-tv)\, \mathrm{d}v. \tag{3.45}$$

The second term in (3.45) tends to zero as $t \to 0$. Indeed, after pairing v and $-v$, its absolute value is at most

$$2\int_0^C \kappa_\chi(v) \frac{|\sin(2\pi t v)|}{v}\, \mathrm{d}v = O_\chi(|t|).$$

Thus the symmetric value and the two one-sided limits are

$$R_\chi(0) = 0, \qquad R_\chi(0^+) = -\mathrm{i}\pi, qquad R_\chi(0^-) = \mathrm{i}\pi. \tag{3.46}$$

For $t \ne 0$, differentiation in (3.45) gives

$$R'_\chi(t) = 2\pi\mathrm{i}\widehat{\kappa_\chi}(t).$$

Dirichlet's integral applied to (3.44) shows that $R_\chi(t) \to 0$ at both infinities. Since $\widehat{\kappa_\chi}$ is Schwartz, integration from the appropriate infinity shows that R_χ is smooth and rapidly decreasing on each open half-line. More precisely, for all $r, J \ge 0$,

$$|R_\chi^{(r)}(t)| \le C_{r,J,\chi}(1+|t|)^{-J} \qquad (t \ne 0), \tag{3.47}$$

where for $r=0$ the two half-lines are treated separately. In particular, R_χ, with the value in (3.46), is a regulated BV function with one jump; it is not being regarded as a Schwartz function on all of $\mathbb{R}$.

Auxiliary scaled-multiplier estimate. Let $R=R_\chi$. If $p,p'\in(Q,2Q]$ and $a,a'\in\mathbb{R}$, then, for every $J\ge 0$,

$$\begin{gathered} W(t)=R\left(\frac{t-a}{p^2}\right)\overline{R\left(\frac{t-a'}{p'^2}\right)}, \\ \|W\|_\infty+\mathrm{TV}_{\mathbb{R}}(W)\ll_{J,\chi}\left(1+\frac{|a-a'|}{Q^2}\right)^{-J}. \end{gathered}\tag{3.48}$$

Here the values at a and a' are induced by $R(0)=0$. Moreover, suppose that $\ell\neq 0$, $a=\ell Np$, $a'=\ell Np'$, and put

$$H=\frac{|\ell|N}{Q},\qquad I_{\ell,Q}=\left\{t:\ \operatorname{sgn}(t)=\operatorname{sgn}(\ell),\quad \frac{|\ell|NQ}{2}\le|t|\le 3|\ell|NQ\right\},\qquad E_{\ell,Q}=\mathbb{R}\setminus I_{\ell,Q}.$$

Then

$$\left\|\mathbf{1}_{E_{\ell,Q}}W\right\|_\infty+\mathrm{TV}_{\mathbb{R}}(\mathbf{1}_{E_{\ell,Q}}W)\ll_{J,\chi}H^{-J}.\tag{3.49}$$

The convention at the two endpoints of $I_{\ell,Q}$ is immaterial.

Proof of the auxiliary estimate. Write $E_M(x)=(1+|x|)^{-M}$. Uniformly for $1/4\le\lambda\le 4$, $d\in\mathbb{R}$, and $M>J+1$,

$$\sup_x E_M(x)E_M(\lambda x-d)+\int_{\mathbb{R}}E_M(x)E_M(\lambda x-d)\,\mathrm{d}x\ll_{J,M}(1+|d|)^{-J}.\tag{3.50}$$

To verify this, split the line into $|x|\ge|d|/8$ and its complement. On the first set one extracts $(1+|d|)^{-J}$ from the first factor and integrates the remaining envelope. On the complement, $|\lambda x-d|\ge|d|/2$, and one instead extracts the same power from the second factor. The supremum estimate is identical.

Away from a,a', ordinary differentiation gives

$$W'(t)=\frac{1}{p^2}R'\left(\frac{t-a}{p^2}\right)\overline{R\left(\frac{t-a'}{p'^2}\right)}+\frac{1}{p'^2}R\left(\frac{t-a}{p^2}\right)\overline{R'\left(\frac{t-a'}{p'^2}\right)}.$$

In the first integral use $t=a+p^2x$, and in the second use $t=a'+p'^2x$. Since $p^2/p'^2\in[1/4,4]$, (3.47) and (3.50) bound both $\int_{\mathbb{R}\setminus\{a,a'\}}|W'(t)|\,\mathrm{d}t$ and $\|W\|_\infty$ by the right side of (3.48).

It remains to count the point jumps. If $a\neq a'$, then the total variation contributed at a, including the assigned value $W(a)=0$, is

$$2\pi\left|R\left(\frac{a-a'}{p'^2}\right)\right|;$$

the analogous contribution at a' is $2\pi|R((a'-a)/p^2)|$. Both have the required separation decay. If $a=a'$, the two one-sided products are bounded and the assigned value is zero, so the total point contribution is at most $2\pi^2$. This proves (3.48), including unequal scales and all possible coincidences of the jump points.

For (3.49), both centers lie strictly inside $I_{\ell,Q}$, while for every $t\in E_{\ell,Q}$,

$$1+\frac{|t-a|}{p^2}\gg 1+H,\qquad 1+\frac{|t-a'|}{p'^2}\gg 1+H.$$

Thus multiplication by $\mathbf{1}_{E_{\ell,Q}}$ kills the original jump points. On $E_{\ell,Q}$, use the displayed derivative formula, take the supremum of one far-away factor, and integrate the other. This contributes

$O_{J,\chi}(H^{-J})$. The only new jumps are at the two endpoints of $I_{\ell,Q}$; both factors there have normalized distance $\gg H$, so these point contributions obey the same bound. This proves (3.49). □

We also record the interchange that leads to the Fourier coefficients below. The original symmetric principal value agrees with its Abel regularization. Indeed, once $|q|$ is large, the cutoff factor is one, and pairing q with $-q$ gives

$$V(Np\alpha)\left(\frac{1}{p\alpha-q}+\frac{1}{p\alpha+q}\right)=O_p(q^{-2}).$$

Thus the paired tails are absolutely summable, and Abel's theorem applies to the resulting convergent symmetric series. To compare the literal Abel weights with a symmetrically weighted pairing, note that, for $q>|p\alpha|$, the two weights are $e^{-\lambda_{\mathrm{A}}q}e^{\pm\lambda_{\mathrm{A}}p\alpha}$. Replacing both by $e^{-\lambda_{\mathrm{A}}q}$ changes the paired tail by

$$O_p\left(\lambda_{\mathrm{A}}\sum_{q\geq 1}\frac{e^{-\lambda_{\mathrm{A}}q}}{q}\right)=O_p(\lambda_{\mathrm{A}}\log(2/\lambda_{\mathrm{A}}))=o(1).$$

The symmetrically weighted paired series is dominated by its $O_p(q^{-2})$ tail and therefore converges to the original symmetric principal value. We may therefore first truncate each q-sum symmetrically and insert the Abel factor $\exp(-\lambda_{\mathrm{A}}|p\alpha-q|)$, $\lambda_{\mathrm{A}}>0$. The sum is then finite and, after the truncation is removed, absolutely integrable. The Fourier series (2.1) is absolutely and uniformly convergent, so it may be inserted term by term before unfolding the residue classes modulo p. For the ℓ-th term, the changes of variables $x=p\alpha-q$ and then $v=px$ give the damped transform

$$\int_{\mathbb{R}}\frac{\chi(v)}{v}e\left(-\frac{n-\ell Np}{p^2}v\right)e^{-\lambda_{\mathrm{A}}|v|/p}\,\mathrm{d}v.$$

Letting $\lambda_{\mathrm{A}}\downarrow 0$ symmetrically gives $R_\chi((n-\ell Np)/p^2)$. This passage is pointwise in the real frequency and uniformly bounded; uniform convergence across frequency zero is neither true nor needed. More precisely, for the principal kernel,

$$\mathrm{PV}\int_{\mathbb{R}}\frac{e(-sv)e^{-\lambda_{\mathrm{A}}|v|/p}}{v}\,\mathrm{d}v=-2\mathrm{i}\arctan\left(\frac{2\pi ps}{\lambda_{\mathrm{A}}}\right),$$

which is uniformly bounded and converges pointwise to $-\mathrm{i}\pi\,\mathrm{sgn}(s)$. For the correction, pairing v and $-v$ gives

$$-2\mathrm{i}\int_0^C\kappa_\chi(v)\frac{\sin(2\pi sv)}{v}e^{-\lambda_{\mathrm{A}}v/p}\,\mathrm{d}v,$$

which is uniformly bounded by Dirichlet's estimate and converges for each s as $\lambda_{\mathrm{A}}\downarrow 0$. Since $p\leq N$ is a finite sum and $\sum_\ell|v_\ell|<\infty$, dominated convergence now removes the Abel factor and exchanges the principal value, the Fourier series, and the cell unfolding. This is the same symmetric truncation convention as in Lemma 2.1.

Using (2.1), the n-th Fourier coefficient of the ℓ-th mark mode of Z_N^χ, without the factor v_ℓ, is

$$B_\ell(n)=\sum_{p\leq N}\frac{c_p(-n)}{p^2}R_\chi\left(\frac{n-\ell Np}{p^2}\right). \tag{3.51}$$

The unshifted mode. For $n\asymp K$, the transition in (3.51) with $\ell=0$ is $Q_K=\sqrt{K}$. From (3.47), for a suitable fixed $C_\chi\geq 1$ and every j,

$$|R_\chi(n/p^2)|+p\left|\frac{\mathrm{d}}{\mathrm{d}p}R_\chi(n/p^2)\right|\ll_{j,\chi}\begin{cases}(p/Q_K)^j, & p<Q_K/C_\chi,\\ 1, & p\geq Q_K/C_\chi.\end{cases}$$

We now spell out the vector-valued Abel summation, including its boundaries. Put

$$\mathcal{H}_K = \ell^2\{n \in \mathbb{Z} : K < n \le 2K\}, \qquad u_p = \left(\frac{c_p(-n)}{p^2}\right)_{n \asymp K}, \qquad w_x = (R_\chi(n/x^2))_{n \asymp K},$$

and, for a finite upper endpoint U, set

$$\mathcal{R}_{x,U} = \sum_{x<p\le U} u_p.$$

The exact identity (3.17), applied with this multiplier, reads

$$\sum_{2<p\le U} u_p \odot w_p = \mathcal{R}_{2,U} \odot w_2 + \int_2^U \mathcal{R}_{x,U} \odot \frac{\mathrm{d}w_x}{\mathrm{d}x}\,\mathrm{d}x. \tag{3.52}$$

There is no omitted upper boundary, because $\mathcal{R}_{U,U} = 0$.

For all sufficiently large K, put $T = K/(\log K)^{B_{\mathrm{tail}}} \ge 2$ and $U_0 = \min\{N, T\}$; the bounded set of smaller blocks will be absorbed into the final constant. The finite-tail form of (3.3) gives, uniformly for $2 \le x \le U_0$,

$$\|\mathcal{R}_{x,U_0}\|_{\mathcal{H}_K} \ll \frac{\sqrt{K}}{x} + E_K. \tag{3.53}$$

For a fixed integer $j > 2$, define

$$\omega_{K,j}(x) = \min\left\{1, \left(\frac{C_\chi x}{Q_K}\right)^j\right\}.$$

The preceding multiplier bound implies

$$\|w_x\|_{\ell^\infty(n\asymp K)} + x\left\|\frac{\mathrm{d}w_x}{\mathrm{d}x}\right\|_{\ell^\infty(n\asymp K)} \ll_{j,\chi} \omega_{K,j}(x). \tag{3.54}$$

The half-line Schwartz bounds also imply the vector-valued total variation estimate

$$\int_0^\infty \sup_{K\le n\le 2K}\left|\frac{\mathrm{d}}{\mathrm{d}x}R_\chi(n/x^2)\right|\mathrm{d}x \ll_\chi 1. \tag{3.55}$$

To see this, put $n = rK$, $1 \le r \le 2$, and $t = K/x^2$; the left side is bounded by a constant times $\int_0^\infty \sup_{1\le r\le 2} r|R'_\chi(rt)|\,\mathrm{d}t$. This is finite by (3.47). Applying (3.52), (3.53), and (3.54), and using

$$\frac{\omega_{K,j}(2)}{2} + \int_2^\infty \frac{\omega_{K,j}(x)}{x^2}\,\mathrm{d}x \ll_{j,\chi} \frac{1}{Q_K},$$

we obtain

$$\left\|\sum_{2<p\le U_0} u_p \odot w_p\right\|_{\mathcal{H}_K} \ll_{j,\chi} \frac{\sqrt{K}}{Q_K} + E_K = O_\chi(1) + o(1). \tag{3.56}$$

The E_K-part includes the boundary at 2 and is bounded using (3.55).

If $T < N$, apply the same exact identity to $(T, N]$. The finite form of (3.4) gives

$$\|\mathcal{R}_{x,N}\|_{\mathcal{H}_K} \ll \frac{\sqrt{K}(\log K)^{3/2}}{x}, \qquad T \le x \le N,$$

where now $\mathcal{R}_{x,N} = \sum_{x<p\le N} u_p$. Moreover,

$$\left|\frac{\mathrm{d}}{\mathrm{d}x}R_\chi(n/x^2)\right| \ll_\chi \min(x^{-1}, Kx^{-3}).$$

The lower boundary at T is $O(\sqrt{K}(\log K)^{3/2}/T) = O(E_K)$, since R_χ is bounded; the upper boundary again vanishes. The remaining Abel integral is

$$\ll \sqrt{K}(\log K)^{3/2}K \int_T^\infty x^{-4}\,\mathrm{d}x = o(E_K).$$

The finitely many terms $p \le 2$ satisfy an $O_\chi(1)$ bound by the first multiplier estimate. We have therefore proved

$$\|B_0\|_{\ell^2(n\asymp K)} \ll_\chi 1. \tag{3.57}$$

We also record the high-frequency tail rather than subsume it in (3.57). Suppose that $Q_K = \sqrt{K} > N$, equivalently $K > N^2$. For all sufficiently large N, this implies $T > N$: the function $x/(\log x)^{B_{\mathrm{tail}}}$ is increasing in this range, and at $x = N^2$ it is larger than N. Furthermore, once N is sufficiently large, the function $x \mapsto (\log x)^{B_{\mathrm{tail}}+3/2}x^{-1/2}$ is decreasing on $[N^2, \infty)$. Hence

$$\sup_{K\ge N^2} E_K \le \frac{(\log N^2)^{B_{\mathrm{tail}}+3/2}}{N} = o(1),$$

so every use of $E_K = o(1)$ below is uniform over the high-frequency blocks. Apply (3.52) with $U = N$. Every $x \le N$ lies below Q_K, so (3.54) and (3.53) give, for each fixed $j > 2$,

$$\begin{aligned}\Bigg\|\sum_{2<p\le N} u_p \odot w_p\Bigg\|_{\mathcal{H}_K} &\ll_{j,\chi} \sqrt{K}Q_K^{-j}\left(1 + \int_2^N x^{j-2}\,\mathrm{d}x\right) + E_K\left(\frac{N}{Q_K}\right)^j \\ &\ll_{j,\chi} \left(\frac{N}{Q_K}\right)^{j-1}.\end{aligned} \tag{3.58}$$

In the last step we used $\sqrt{K}/Q_K = 1$, $E_K = o(1)$, and $N/Q_K < 1$. The terms $p \le 2$ obey the same bound directly from (3.47), after increasing the fixed decay exponent if necessary. On the dyadic blocks $K_r = 2^r N^2$, the square of (3.58) is $O_{j,\chi}(2^{-r(j-1)})$, and is therefore summable. The range $K \le N^2$ contains $O(\log N) = O(L)$ dyadic blocks, each bounded by (3.57).

Finally, the reality of $\chi(v)/v$ gives $R_\chi(-t) = \overline{R_\chi(t)}$, and Ramanujan sums are real and even. Hence $B_0(-n) = \overline{B_0(n)}$, so the negative-frequency blocks have the same bounds. At frequency zero, $B_0(0) = 0$ by (3.46). Combining the positive, negative, and zero frequencies proves

$$\sum_n |B_0(n)|^2 \ll_\chi L. \tag{3.59}$$

Shifted modes. Fix $\ell \neq 0$. For $p, p' \asymp Q$, apply the auxiliary scaled-multiplier estimate above with $a = \ell N p$ and $a' = \ell N p'$. It gives, for

$$W_{p,p'}(t) = R_\chi\left(\frac{t-\ell Np}{p^2}\right)\overline{R_\chi\left(\frac{t-\ell Np'}{p'^2}\right)},$$

$$\|W_{p,p'}\|_\infty + \mathrm{TV}_{\mathbb{R}}(W_{p,p'}) \ll_{J,\chi} \left(1 + \frac{|\ell|N|p-p'|}{Q^2}\right)^{-J}. \tag{3.60}$$

Before making the dyadic decomposition, remove the singleton $p = 1$. Since $c_1(n) = 1$, its contribution is

$$B_\ell^{[1]}(n) = R_\chi(n - \ell N).$$

Translation by the integer ℓN, together with (3.47) and $R_\chi(0) = 0$, gives the uniform bound

$$\sum_{n\in\mathbb{Z}} |B_\ell^{[1]}(n)|^2 = \sum_{m\in\mathbb{Z}} |R_\chi(m)|^2 \ll_\chi 1. \tag{3.61}$$

Thus it remains to decompose the terms $2 \le p \le N$ into the blocks $Q < p \le 2Q$ below; the singleton can be restored at the end by the triangle inequality. The function $W_{p,p'}$ tends to zero at both infinities. Therefore Lemmas 2.6 and 2.5, together with (3.60), imply, when $p \neq p'$,

$$\left| \sum_{n \in \mathbb{Z}} c_p(n) c_{p'}(n) W_{p,p'}(n) \right| \ll_{J,\chi} \sigma_1(p)\sigma_1(p') \left(1 + \frac{|\ell| N |p - p'|}{Q^2}\right)^{-J}. \tag{3.62}$$

Let $B_{\ell,Q}$ denote the dyadic $Q < p \le 2Q$ part of (3.51). We first write out its diagonal contribution. Put $a = \ell N p$. Since $a/p = \ell N \in \mathbb{Z}$, every interval $(a + kp, a + (k+1)p] \cap \mathbb{Z}$ is a complete residue system modulo p, and hence

$$\sum_{a+kp<n\le a+(k+1)p} c_p(n)^2 = p\varphi(p).$$

The half-line bounds (3.47) now give

$$\begin{aligned} \sum_{n \in \mathbb{Z}} c_p(n)^2 \left| R_\chi\left(\frac{n - \ell N p}{p^2}\right)\right|^2 &\le p\varphi(p) \sum_{k \in \mathbb{Z}} \sup_{k/p < x \le (k+1)/p} |R_\chi(x)|^2 \\ &\ll_\chi p^2 \varphi(p). \end{aligned} \tag{3.63}$$

Consequently the diagonal part of $\|B_{\ell,Q}\|_2^2$ is

$$\ll_\chi \sum_{Q<p\le 2Q} \frac{\varphi(p)}{p^2} \ll_\chi 1.$$

For the off-diagonal, put $\lambda = |\ell| N/Q^2$, use $\sigma_1(p)/p \ll \log\log(3Q)$, and choose $J > 2$. Equation (3.62) gives

$$\begin{aligned} &\sum_{\substack{p,p' \asymp Q \\ p \neq p'}} \frac{\sigma_1(p)\sigma_1(p')}{p^2 p'^2} (1 + \lambda|p - p'|)^{-J} \\ &\quad \ll \frac{(\log\log(3Q))^2}{Q^2} Q \sum_{1 \le h \le Q} (1 + \lambda h)^{-J} \\ &\quad \ll (\log\log(3Q))^2 \min\left(1, \frac{Q}{|\ell| N}\right). \end{aligned} \tag{3.64}$$

Combining this with (3.63) proves, uniformly for $\ell \neq 0$, the slightly stronger form

$$\|B_{\ell,Q}\|_2^2 \ll_\chi 1 + (\log\log(3Q))^2 \min\left(1, \frac{Q}{|\ell| N}\right). \tag{3.65}$$

It remains to square-sum in Q. Let

$$Q_0 = \frac{N}{L^{C_{\mathrm{far}}}}, \qquad C_{\mathrm{far}} > 10,$$

and let $\Pi_{\ell,Q}$ project onto

$$\operatorname{sgn}(n) = \operatorname{sgn}(\ell), \qquad \frac{|\ell| N Q}{2} \le |n| \le 3|\ell| N Q.$$

Put $H = |\ell| N/Q$ and let $E_{\ell,Q}$ be the complement in $\mathbb{R}$ of the displayed carrier annulus. Expanding by Plancherel gives the exact identity

$$\begin{aligned} \|(1 - \Pi_{\ell,Q}) B_{\ell,Q}\|_2^2 = \sum_{p,p' \asymp Q} \frac{1}{p^2 p'^2} \sum_{n \in \mathbb{Z}} c_p(n) c_{p'}(n) \\ \cdot \mathbf{1}_{E_{\ell,Q}}(n) W_{p,p'}(n). \end{aligned} \tag{3.66}$$

For $p \neq p'$, the projected estimate (3.49) and Lemmas 2.6 and 2.5 bound the inner sum by

$$O_{J,\chi}(\sigma_1(p)\sigma_1(p')H^{-J}). \tag{3.67}$$

For $p = p'$, every retained n satisfies $|n - \ell N p|/p^2 \gg H$. Subdivide into the periods used in (3.63) and take a Schwartz exponent larger than $J+1$. The remaining supremum sum is bounded by

$$p\int_{|x|\gg H}(1+|x|)^{-J-2}\,\mathrm{d}x + O(H^{-J-2}) \ll_J pH^{-J}.$$

It follows that

$$\sum_{n\in E_{\ell,Q}\cap\mathbb{Z}} c_p(n)^2\left|R_\chi\left(\frac{n-\ell Np}{p^2}\right)\right|^2 \ll_{J,\chi} p^2\varphi(p)H^{-J}. \tag{3.68}$$

After applying the outside weights in (3.66), the diagonal contribution is $O_{J,\chi}(H^{-J})$, because $\sum_{p\asymp Q}\varphi(p)/p^2 \ll 1$. The off-diagonal contribution from (3.67) is $O_{J,\chi}((\log\log(3Q))^2H^{-J})$. We have therefore proved, including both projection-boundary jumps, that for every J,

$$\|(1-\Pi_{\ell,Q})B_{\ell,Q}\|_2^2 \ll_{J,\chi} (1+\log\log(3Q))^2\left(\frac{Q}{|\ell|N}\right)^J. \tag{3.69}$$

The carrier annuli have bounded overlap as Q ranges dyadically: for each integer frequency n, at most a fixed number of dyadic values of Q can satisfy $n \in I_{\ell,Q}$, since those values have ratio at most 6. Therefore

$$\begin{aligned}\left\|\sum_Q \Pi_{\ell,Q}B_{\ell,Q}\right\|_2^2 &= \sum_n\left|\sum_Q \mathbf{1}_{I_{\ell,Q}}(n)B_{\ell,Q}(n)\right|^2\\ &\ll \sum_Q\|B_{\ell,Q}\|_2^2.\end{aligned} \tag{3.70}$$

For $Q \le Q_0$, the sum of (3.65) over dyadic Q is $O_\chi(L)$: the constant terms contribute $O(L)$, while

$$\sum_{Q\le Q_0}(\log\log(3Q))^2\frac{Q}{|\ell|N} = o(1).$$

Moreover, after taking square roots in (3.69),

$$\sum_{Q\le Q_0}\|(1-\Pi_{\ell,Q})B_{\ell,Q}\|_2 = o(1)$$

on choosing J large. Bounded overlap of the projected annuli in (3.70) therefore gives

$$\left\|\sum_{Q\le Q_0}B_{\ell,Q}\right\|_2^2 \ll_\chi L. \tag{3.71}$$

There are $O(\log L)$ blocks above Q_0. Each has squared norm $O_\chi((\log L)^2)$, so the triangle inequality gives

$$\sup_{\ell\neq 0}\sum_n|B_\ell(n)|^2 \ll_\chi L + (\log L)^4 = o(L^2).$$

Minkowski, $\sum_{\ell\neq 0}|v_\ell| < \infty$, and (3.59) prove

$$\|Z_N^\chi\|_2^2 \ll_\chi L + (\log L)^4 = o(L^2).$$

Together with (3.43), the same holds for Y_N^χ.

Finally remove the smoothing. The difference between Y_N^χ and the sharp sum in (3.40) is supported on

$$\varepsilon-\delta<|p\delta_p|<\varepsilon.$$

Every active p is a convergent denominator, so at most $O(L)$ are active at one α, since $Q_n \ge F_{n+1}$ for the Fibonacci numbers F_n. Each summand is $O_\varepsilon(1)$, and

$$\int_0^1 \sum_{p\le N} \mathbf{1}_{\{\varepsilon-\delta<|p\delta_p|<\varepsilon,\,(p,q_p)=1\}}\,d\alpha \le 2\delta \sum_{p\le N}\frac{\varphi(p)}{p^2} \ll \delta L.$$

Pointwise Cauchy–Schwarz thus bounds the squared L^2-norm of the difference by $O_\varepsilon(\delta L^2)$. First let $N\to\infty$ and then $\delta\downarrow 0$. This proves (3.40).

It remains to prove the quantified uniformity (3.41). Fix $0<\delta<\varepsilon_0/2$. For every $t\in[\varepsilon_0,\varepsilon_1]$, construct $\chi_{t,\delta}$ from one fixed smooth transition profile so that it vanishes on $|v|\le t-\delta$ and equals one on $|v|\ge t$. All these profiles vanish on a fixed neighborhood of zero, and their transition points remain in a fixed compact set. Hence, for each fixed r, J,

$$\sup_{t\in[\varepsilon_0,\varepsilon_1]} C_{r,J,\chi_{t,\delta}}<\infty, \tag{3.72}$$

where the bound may depend on the presently fixed δ. Every periodization and multiplier estimate above is consequently uniform in t. In particular, (3.43) and the bound for Z_N^χ give the explicit common estimate

$$\sup_{t\in[\varepsilon_0,\varepsilon_1]}\left\|Y_N^{\chi_{t,\delta}}\right\|_2^2 \ll L\log L + C_\delta(L+(\log L)^4) = o_\delta(L^2).$$

Equivalently,

$$\sup_{t\in[\varepsilon_0,\varepsilon_1]}\left\|Y_N^{\chi_{t,\delta}}\right\|_2 = o_\delta(L). \tag{3.73}$$

Let $F_N(t)$ denote the sharp sum in (3.41), and put $D_{N,t}=F_N(t)-Y_N^{\chi_{t,\delta}}$. This difference is supported on

$$t-\delta<|p\delta_p|<t.$$

Each active summand is $O_{\varepsilon_0}(1)$. Moreover, its reduced fraction satisfies

$$\left|\alpha-\frac{q_p}{p}\right|<\frac{\varepsilon_1}{p^2}<\frac{1}{2p^2}.$$

Legendre's criterion therefore makes q_p/p a continued-fraction convergent. Since consecutive convergent denominators satisfy $Q_{n+2}\ge 2Q_n$, the number of active $p\le N$ is $O(L)$, uniformly in α and t. On the other hand, direct cell integration gives, uniformly in t,

$$\int_0^1 \sum_{p\le N} \mathbf{1}_{\{t-\delta<|p\delta_p|<t,\ (p,q_p)=1\}}\,d\alpha \le 2\delta\sum_{p\le N}\frac{\varphi(p)}{p^2}+O(\delta)\ll \delta L. \tag{3.74}$$

The $O(\delta)$ term harmlessly covers the endpoint cells for $p=1$. Pointwise Cauchy–Schwarz, followed by (3.74), yields

$$\sup_{t\in[\varepsilon_0,\varepsilon_1]}\|D_{N,t}\|_2^2 \ll_{\varepsilon_0} \delta L^2. \tag{3.75}$$

Combining (3.73) and (3.75),

$$\limsup_{N\to\infty}\sup_{t\in[\varepsilon_0,\varepsilon_1]} L^{-1}\|F_N(t)\|_2 \ll_{\varepsilon_0} \sqrt{\delta}.$$

Letting $\delta\downarrow 0$ proves (3.41). Thus the order of limits is: first $N\to\infty$ with δ fixed, and then $\delta\downarrow 0$. □

3.4 The complete small-jump estimate

Proposition 3.5 (Minor-arc truncation)**.** *For every $\gamma > 0$,*

$$\lim_{A\to\infty} \limsup_{N\to\infty} \mathrm{Leb}\left\{\alpha : \frac{1}{L}\left|\sum_{\substack{p\le N,\ (p,q_p)=1\\ L|p\delta_p|>A}} \frac{V(N\delta_p)}{p\delta_p}\right| > \gamma\right\} = 0. \tag{3.76}$$

Proof. Fix $0 < \varepsilon < 1/4$, and abbreviate

$$v_p = p\delta_p, \qquad \xi_p = Lv_p, \qquad u_p = N\delta_p \bmod 1.$$

Let

$$F_N(t) = \sum_{\substack{p\le N,\ (p,q_p)=1\\ |v_p|\ge t}} \frac{V(u_p)}{v_p}.$$

Proposition 3.4 gives $\|F_N(\varepsilon/2)\|_2 = o(L)$.

For all sufficiently large N, the fixed-profile cutoff used in Proposition 3.3 yields the exact decomposition, outside null boundary sets,

$$\sum_{\substack{p\le N,\ (p,q_p)=1\\ |\xi_p|>A}} \frac{V(u_p)}{v_p} = F_N(\varepsilon/2) + Y^{\mathrm{near}}_{N,A,\varepsilon} + E^{\mathrm{lo}}_{N,A} - E^{\mathrm{up}}_{N,\varepsilon}, \tag{3.77}$$

where

$$E^{\mathrm{lo}}_{N,A} = \sum_{\substack{p\le N,\ (p,q_p)=1\\ A<|\xi_p|<2A}} (1-\eta(|\xi_p|/A))\,\frac{V(u_p)}{v_p},$$

$$E^{\mathrm{up}}_{N,\varepsilon} = \sum_{\substack{p\le N,\ (p,q_p)=1\\ \varepsilon/2<|v_p|<\varepsilon}} \beta(|v_p|/\varepsilon)\frac{V(u_p)}{v_p}.$$

This identity, rather than a containment argument, is needed because cancellation does not compare a weighted transition sum with its containing sharp band.

The near term satisfies

$$\limsup_{N\to\infty} L^{-2}\|Y^{\mathrm{near}}_{N,A,\varepsilon}\|_2^2 \ll_\varepsilon A^{-1}$$

by Proposition 3.3. For the lower transition, define $\mathcal{P}$ to be the limiting Poisson process in Lemma 4.1, and put

$$h_A(\xi,u) = (1-\eta(|\xi|/A))\,\mathbf{1}_{\{A<|\xi|<2A\}}\frac{V(u)}{\xi}.$$

Lemma 4.1 and the continuous mapping theorem (the boundary has zero limiting intensity) give

$$\frac{E^{\mathrm{lo}}_{N,A}}{L} \Longrightarrow \int h_A\,\mathrm{d}\mathcal{P}.$$

The process in Lemma 4.1 uses $p < N$; the omitted $p = N$ term is nonzero on a set of measure $O_A((LN)^{-1})$, so it does not affect this convergence. The limiting shot has mean zero by symmetry in ξ, and

$$\mathrm{Var}\left(\int h_A\,\mathrm{d}\mathcal{P}\right) = \frac{1}{\zeta(2)}\int h_A(\xi,u)^2\,\mathrm{d}\xi\,\mathrm{d}u \ll \frac{1}{A}.$$

Thus Portmanteau and Chebyshev bound its contribution to (3.76) by $O(A^{-1}\gamma^{-2})$.

It remains to justify the upper transition. Put $r_0 = \varepsilon/2$, $r_1 = \varepsilon$, and $h(t) = \beta(t/\varepsilon)$. Stieltjes layer cake gives exactly

$$E^{\mathrm{up}}_{N,\varepsilon} = h(r_0)F_N(r_0) + \int_{r_0}^{r_1} h'(t)F_N(t)\,\mathrm{d}t.$$

This identity follows first pointwise in α, by applying the one-variable Stieltjes identity to each of the finitely many shots and then summing. As a map from $[r_0, r_1]$ to the separable space $L^2(\mathbb{T})$, $t \mapsto F_N(t)$ is strongly measurable: its joint representative is measurable in (t, α). Moreover, for fixed N, throughout this interval

$$|F_N(t, \alpha)| \le \frac{N\|V\|_\infty}{r_0},$$

so it is Bochner integrable. The pointwise identity is therefore also an identity in $L^2(\mathbb{T})$. Minkowski's inequality and (3.41) therefore give

$$\frac{\left\|E^{\mathrm{up}}_{N,\varepsilon}\right\|_2}{L} \le \left(|h(r_0)| + \int_{r_0}^{r_1} |h'(t)|\,\mathrm{d}t\right) \sup_{r_0 \le t \le r_1} \frac{\|F_N(t)\|_2}{L} \longrightarrow 0.$$

Apply Chebyshev to the near term, use the preceding bounds in (3.77), and finally let $A \to \infty$. This proves (3.76). $\square$

4 The marked Poisson process and the limit law

4.1 A continued-fraction marked Poisson theorem

Let m be normalized Haar measure on $\mathbb{T}$. We prove the one-dimensional marked theorem directly; no rank-two specialization of a Rogers formula is used.

Lemma 4.1 (Signed marked resonances)**.** *For α uniform on $\mathbb{T}$, the point process*

$$\mathcal{Q}_N = \sum_{\substack{1 \le p < N \\ (p, q_p) = 1}} \delta_{(\log p/L,\; Lp\delta_p,\; N\delta_p \bmod 1)} \tag{4.1}$$

converges vaguely on $[0,1] \times (\mathbb{R} \setminus \{0\}) \times \mathbb{T}$ to a Poisson process with intensity

$$\lambda\, dt\, d\xi\, dm(u), \qquad \lambda = \frac{6}{\pi^2} = \frac{1}{\zeta(2)}. \tag{4.2}$$

Consequently,

$$\mathcal{P}_N = \sum_{\substack{1 \le p < N \\ (p, q_p) = 1}} \delta_{(Lp\delta_p,\; N\delta_p \bmod 1)} \tag{4.3}$$

converges vaguely on $(\mathbb{R} \setminus \{0\}) \times \mathbb{T}$ to the Poisson process with intensity

$$\lambda\, d\xi\, dm(u). \tag{4.4}$$

Proof. Roadmap. We first establish the unmarked factorial measures in deterministic continued-fraction index time and transfer them from Gauss to Lebesgue measure. We then prove asymptotic Haar marking by an oscillatory cylinder estimate, treating early and late configurations separately. Finally, factorial measures are upgraded to point-process convergence, index time is replaced by denominator time, and the endpoint time layers are restored by a direct rational-cell estimate.

Convergents and rare digits. Write

$$\alpha = [0; a_1, a_2, \ldots]$$

and denote its regular convergents by P_n/Q_n, with

$$P_{-1} = 1, \quad P_0 = 0, \qquad Q_{-1} = 0, \quad Q_0 = 1.$$

Let $T(x) = \{1/x\}$, $x_n = T^n\alpha$, and $y_n = Q_{n-1}/Q_n$. The determinant identities for convergents give

$$Q_n\alpha - P_n = (-1)^n \frac{\theta_n}{Q_n}, \qquad \theta_n = \frac{x_n}{1 + x_n y_n} = \frac{1}{a_{n+1} + x_{n+1} + y_n}. \tag{4.5}$$

These identities include $n = 0$, with $y_0 = 0$.

Fix $B < \infty$. If a point in (4.3) satisfies $|Lp\delta_p| \le B$, then, for all sufficiently large N,

$$\left|\alpha - \frac{q_p}{p}\right| < \frac{1}{2p^2}.$$

Since q_p/p is reduced, Legendre's criterion [9, Chapter II, §7, Theorem 19, p. 30] says that it is a convergent. Thus, on every fixed ξ-window, (4.1) is exactly

$$\sum_{Q_n < N} \delta_{(\log Q_n/L,\ (-1)^n L\theta_n,\ (-1)^n N\theta_n/Q_n \bmod 1)}. \tag{4.6}$$

Let ν_G be Gauss measure,

$$d\nu_G(x) = \frac{dx}{\log 2\,(1+x)}.$$

Set

$$\kappa = \frac{12\log 2}{\pi^2}.$$

The continued-fraction digits are exponentially ψ-mixing: there are $C > 0$ and $\vartheta > 1$ such that, if

$$F \in \sigma(a_1, \ldots, a_j), \qquad H \in \sigma(a_{j+d}, a_{j+d+1}, \ldots),$$

then

$$|\nu_G(F \cap H) - \nu_G(F)\nu_G(H)| \le C\vartheta^{-d}\nu_G(F)\nu_G(H). \tag{4.7}$$

We use the form and notation recorded in [5, (1.7)]. We record two consequences of (4.7) that will be used repeatedly. Suppose that $\mathscr{A}$ and $\mathscr{H}$ are σ-fields for which

$$|\nu_G(A \cap B) - \nu_G(A)\nu_G(B)| \le \eta\, \nu_G(A)\nu_G(B) \qquad (A \in \mathscr{A},\ B \in \mathscr{H}).$$

Then, for arbitrary complex-valued $U_1 \in L^1(\mathscr{A})$ and $U_2 \in L^1(\mathscr{H})$,

$$|\mathbb{E}_{\nu_G}(U_1U_2) - (\mathbb{E}_{\nu_G}U_1)(\mathbb{E}_{\nu_G}U_2)| \le 2\eta\, \mathbb{E}_{\nu_G}|U_1|\, \mathbb{E}_{\nu_G}|U_2|. \tag{4.8}$$

Indeed, for nonnegative simple functions this follows by expanding both functions into indicator functions and applying the event inequality term by term. Monotone convergence gives the result for arbitrary nonnegative integrable functions. For a complex function U_1, write

$$U_1 = (\Re U_1)^+ - (\Re U_1)^- + \mathrm{i}\big((\Im U_1)^+ - (\Im U_1)^-\big),$$

and similarly for U_2. The sums of the L^1-norms of the four nonnegative parts are at most $\sqrt{2}\,\mathbb{E}|U_1|$ and $\sqrt{2}\,\mathbb{E}|U_2|$, respectively, which proves (4.8).

A second consequence is the following quasi-Bernoulli bound. If $m_1 < \cdots < m_s$ and $E_j \in \sigma(a_{m_j})$, then

$$\nu_G\left(\bigcap_{j=1}^{s} E_j\right) \le C_s \prod_{j=1}^{s} \nu_G(E_j). \tag{4.9}$$

To see this, expose the events from left to right and apply (4.7) at each step. Every gap is at least one, so each application costs at most $1 + C\vartheta^{-1}$. This explains precisely what is meant below by the quasi-Bernoulli bound.

For a compact interval $J \Subset (0, \infty)$, the last formula in (4.5) gives, on the event $L\theta_n \in J$,

$$\left| L\theta_n - \frac{L}{a_{n+1}} \right| \ll_J \frac{1}{L}. \tag{4.10}$$

The same conclusion, with a constant depending on a fixed enlargement of J, holds when $L/a_{n+1} \in J$. Indeed, either event forces $a_{n+1} \asymp_J L$, and

$$\left| \frac{L}{a_{n+1}} - \frac{L}{a_{n+1} + x_{n+1} + y_n} \right| \leq \frac{2L}{a_{n+1}^2}.$$

Thus, whenever either of the two indicators differs from the other, both value coordinates lie in one fixed compact subset of $(0, \infty)$, and one of them lies within $O_J(L^{-1})$ of an endpoint of J. Moreover, direct summation of

$$\nu_G(a_1 = q) = \frac{1}{\log 2} \log \frac{(q+1)^2}{q(q+2)} = \frac{1}{\log 2} \left(\frac{1}{q^2} + O(q^{-3}) \right)$$

shows that

$$\nu_G \left\{ \frac{L}{a_1} \in J \right\} = \frac{|J|}{L \log 2} + O_J(L^{-2}). \tag{4.11}$$

Unmarked factorial measures. We give the factorial-moment argument explicitly. Fix r, compact index-time intervals I_j, compact value intervals $J_j \Subset (0, \infty)$, and prescribed parities. Put $b_L = \lceil C_{\mathrm{mix}} \log L \rceil$, with C_{mix} large, and sum over distinct positions $n_j \in \kappa L I_j$ having the prescribed parities. For the digit events

$$A_{n,j} = \{L/a_{n+1} \in J_j\},$$

(4.11) gives

$$\nu_G(A_{n,j}) = \frac{|J_j|}{L \log 2} + O_{J_j}(L^{-2}).$$

If the ordered positions have successive gaps at least b_L, repeated use of (4.7) gives

$$\nu_G \left(\bigcap_{j=1}^{r} A_{n_j,j} \right) = \prod_{j=1}^{r} \nu_G(A_{n_j,j}) \left(1 + O_r(\vartheta^{-b_L}) \right). \tag{4.12}$$

If some gap is smaller than b_L, there are only $O_r(b_L L^{r-1})$ possible tuples. Applying (4.7) successively with gap one bounds every such intersection by $O_r(L^{-r})$. Hence all short-gap tuples contribute

$$O_r(b_L/L) = o(1). \tag{4.13}$$

For a value-window endpoint $z > 0$, the replacement error in (4.10) is contained in

$$\{\mathrm{dist}(L/a_{n+1}, z) \leq C/L\},$$

whose probability is $O_z(L^{-2})$, by the same direct digit summation as in (4.11). Combining this boundary event with $r-1$ ordinary rare events and using the gap-one bound shows that the sum of all replacement errors is $O_r(L^r L^{-(r+1)}) = O_r(L^{-1})$.

Thus (4.12), (4.13), and the Riemann sums over the index windows prove convergence of every unmarked factorial measure in deterministic index time. The number of integers of either parity in $\kappa L I_j$ is $\frac{1}{2}\kappa L|I_j| + O(1)$, and $\kappa/(2\log 2) = 6/\pi^2 = \lambda$. Explicitly, for prescribed parities and labeled intervals as above,

$$\sum_{\substack{n_j \in \kappa L I_j \\ n_1,\dots,n_r \text{ distinct} \\ n_j \text{ of the prescribed parity}}} \nu_G\left(\bigcap_{j=1}^{r}\{L\theta_{n_j} \in J_j\}\right) \longrightarrow \lambda^r \prod_{j=1}^{r} |I_j|\,|J_j|, \qquad \lambda = \frac{6}{\pi^2}. \tag{4.14}$$

We now prove explicitly the uniform moment bounds that will be used below. Let $\mathscr{I}_L$ be any union of a fixed number of deterministic index-time windows, so that $\#\mathscr{I}_L = O(L)$, let $K \Subset (0,\infty)$, and put

$$R_L = \sum_{n \in \mathscr{I}_L} \mathbf{1}_{\{L\theta_n \in K\}}.$$

Since

$$\theta_n = \frac{1}{a_{n+1} + x_{n+1} + y_n}, \qquad 0 < x_{n+1} + y_n < 2,$$

the event $L\theta_n \in K$ implies $c_K L \le a_{n+1} \le C_K L$ for suitable positive constants c_K, C_K. The latter one-digit event has Gauss probability $O_K(L^{-1})$. Therefore, for every fixed $s \ge 1$, (4.9) gives

$$\begin{aligned}
\mathbb{E}_{\nu_G}(R_L)_s &= \sum_{\substack{n_1,\dots,n_s \in \mathscr{I}_L \\ \text{all distinct}}} \nu_G\left(\bigcap_{j=1}^{s}\{L\theta_{n_j} \in K\}\right) \\
&\le \sum_{\substack{n_1,\dots,n_s \in \mathscr{I}_L \\ \text{all distinct}}} \nu_G\left(\bigcap_{j=1}^{s}\{c_K L \le a_{n_j+1} \le C_K L\}\right) \ll_{K,s} L^s L^{-s} \ll_{K,s} 1.
\end{aligned} \tag{4.15}$$

Here $(x)_s = x(x-1)\cdots(x-s+1)$. Since $d\,\mathrm{Leb}/d\nu_G = \log 2\,(1+x)$ is bounded, the same estimate holds under Lebesgue measure. Finally,

$$R_L^s = \sum_{j=0}^{s} \left\{ \begin{matrix} s \\ j \end{matrix} \right\} (R_L)_j, \tag{4.16}$$

where the braces are Stirling numbers of the second kind. We have therefore proved

$$\sup_L \mathbb{E}_{\nu_G} R_L^s + \sup_L \mathbb{E}_{\mathrm{Leb}} R_L^s < \infty \qquad (s \ge 1). \tag{4.17}$$

Denominator time and the unmarked intensity. The exact identity

$$\log(Q_n + Q_{n-1}x_n) = \sum_{j=0}^{n-1} -\log x_j \tag{4.18}$$

holds. The Gauss map preserves ν_G [9, Chapter III], and its ergodicity also follows directly from (4.7). Since $-\log x \in L^1(\nu_G)$, Birkhoff's ergodic theorem [1] implies, locally uniformly for $s \ge 0$,

$$\frac{\log Q_{\lfloor \kappa L s \rfloor}}{L} \longrightarrow s \quad \text{in probability.} \tag{4.19}$$

Indeed,

$$\int_0^1 -\log x \, d\nu_G(x) = \frac{\pi^2}{12\log 2} = \kappa^{-1},$$

and the two sides of (4.18) differ from $\log Q_n$ by at most $\log 2$. To obtain the claimed uniformity, write $\mathscr{S}_n = \sum_{j<n} -\log x_j$ and $\bar{r} = \kappa^{-1}$. On every Birkhoff generic orbit, for any $C < \infty$ and $\delta > 0$, choose n_0 so that $|\mathscr{S}_n - n\bar{r}| \le \delta n$ for $n \ge n_0$. Then

$$\frac{1}{L}\max_{n_0 \le n \le CL} |\mathscr{S}_n - n\bar{r}| \le C\delta, \qquad \frac{1}{L}\max_{n<n_0} |\mathscr{S}_n - n\bar{r}| \longrightarrow 0.$$

First let $L \to \infty$ and then $\delta \downarrow 0$. Together with the bounded difference above and the harmless floor in $\lfloor \kappa L s \rfloor$, this proves (4.19) locally uniformly. Since Lebesgue and Gauss measure are equivalent, the same almost-sure statement holds for both.

Equations (4.11) and (4.19), together with the parity restriction in (4.5), show that each signed half has intensity

$$\frac{\kappa}{2\log 2} = \frac{6}{\pi^2}. \tag{4.20}$$

Thus the deterministic index-time unmarked process has intensity λ, and (4.19) identifies the candidate denominator-time intensity. The rigorous process-level time change is carried out after marked convergence below. The uniform factorial bounds needed in that argument have already been proved above.

Transfer from Gauss to Lebesgue measure. We next make the transfer from Gauss measure to Lebesgue measure quantitative. Put

$$w = \frac{d\,\mathrm{Leb}}{d\nu_G} = \log 2\,(1+x), \qquad \mathscr{A}_b = \sigma(a_1, \ldots, a_b),$$

and set

$$w_L = \mathbb{E}_{\nu_G}(w \mid \mathscr{A}_{b_L}).$$

We choose the version of this conditional expectation which is constant on each half-open depth-b_L cylinder and, at every rational endpoint, assign one of the two adjacent cylinder values. The endpoints form a null set for both measures. A depth-b continued-fraction cylinder has length

$$\frac{1}{Q_b(Q_b + Q_{b-1})} \le F_{b+1}^{-2}.$$

Since w is Lipschitz and this cylinderwise conditional average lies between the minimum and maximum of w on the cylinder,

$$\|w - w_L\|_\infty \ll F_{b_L+1}^{-2} = o(1). \tag{4.21}$$

Also w_L is uniformly bounded and $\mathbb{E}_{\nu_G} w_L = \mathbb{E}_{\nu_G} w = 1$.

We record an aggregate version of this transfer. Suppose that $X_{\mathbf{n}}$ is a collection of integrable functions, each measurable with respect to digits beginning at a position d_L, where $d_L - b_L \gg L$, and suppose that

$$\sum_{\mathbf{n}} \mathbb{E}_{\nu_G} |X_{\mathbf{n}}| = O(1).$$

Using $w = w_L + (w - w_L)$, followed by (4.8), gives

$$\begin{aligned}\sum_{\mathbf{n}} |\mathbb{E}_{\mathrm{Leb}} X_{\mathbf{n}} - \mathbb{E}_{\nu_G} X_{\mathbf{n}}| &\le \|w - w_L\|_\infty \sum_{\mathbf{n}} \mathbb{E}_{\nu_G} |X_{\mathbf{n}}| \\ &\quad + 2C\vartheta^{-(d_L - b_L)} \mathbb{E}_{\nu_G} |w_L| \sum_{\mathbf{n}} \mathbb{E}_{\nu_G} |X_{\mathbf{n}}| = o(1). \end{aligned} \tag{4.22}$$

For deterministic index-time intervals bounded away from zero, the earliest relevant digit position is $\gg_\varepsilon L$. Apply (4.22) specifically to the digit-replaced tuple functions

$$X_{\mathbf{n}} = \prod_{j=1}^{r} \mathbf{1}_{A_{n_j,j}}, \qquad A_{n,j} = \{L/a_{n+1} \in J_j\}.$$

These functions are measurable with respect to the digit future beginning at the earliest a_{n_j+1}, and their summed $L^1(\nu_G)$-norm is $O(1)$ by (4.15). Thus their factorial limits are the same under Gauss and Lebesgue measure. The earlier aggregate $O(L^{-1})$ replacement of the exact events $\{L\theta_{n_j} \in J_j\}$ by $A_{n_j,j}$ also holds under Lebesgue measure, because

$$\mathrm{Leb}(E) \leq \|w\|_\infty \, \nu_G(E)$$

for every error event E. This transfers the exact-θ unmarked factorial limits without treating an exact-θ event as a future digit event. Equivalently, one may first discard the initial $2b_L$ positions, whose expected rare count is $O(b_L/L) = o(1)$, and retain a gap of b_L. The lower endpoint is restored by the direct cell estimate (4.52) below. This proves the same unmarked factorial limits, moment bounds, and time change for Lebesgue measure.

The following exact one-point calculation is a consistency check and is not used in the multi-point argument below. Conditional on a continued-fraction prefix through n, Lebesgue measure gives the tail x_n density

$$\frac{1+y_n}{(1+x_n y_n)^2}\, dx_n.$$

Under $z = Lx_n/(1+x_n y_n) = L\theta_n$, this becomes exactly $(1+y_n)L^{-1}\, dz$. Since

$$N(Q_n\alpha - P_n) = (-1)^n \frac{N}{LQ_n} z,$$

for every nonzero $h \in \mathbb{Z}$ and compact interval J,

$$\left| \int_J e\left(h\frac{Nz}{LQ_n}\right) dz \right| \leq \frac{LQ_n}{2|h|N}.$$

This tends to zero uniformly on $Q_n \leq N^{1-\varepsilon}$. The joint independence of several such marks is the content of the oscillatory argument below.

The oscillatory marking estimate. It remains to show that the torus coordinates in (4.6) become independent Haar marks. We prove this through the Fourier coefficients of the factorial measures. Fix r, compact positive value intervals, prescribed parities, and compact deterministic index-time intervals contained in $[\varepsilon, 1-\varepsilon]$. A general labeled r-th factorial rectangle is a sum over distinct indices, without an a priori order. Partition that sum according to the at most $r!$ possible relative orders of the indices, and, in each part, relabel the time intervals, value intervals, parities, and Fourier frequencies accordingly. It is therefore enough to treat one such part. For ordered indices $n_1 < \cdots < n_r$, attach the Fourier factor

$$\exp\left(2\pi \mathrm{i} N \sum_{j=1}^{r} h_j (Q_{n_j}\alpha - P_{n_j})\right), \qquad \mathbf{h} = (h_1, \ldots, h_r) \in \mathbb{Z}^r. \tag{4.23}$$

We claim that the sum of the corresponding mixed factorial moment converges to zero whenever $\mathbf{h} \neq 0$. When $\mathbf{h} = 0$, the unmarked calculation above gives the required product measure.

Choose

$$0 < \Delta < \rho/100, \qquad 0 < \rho < \varepsilon/100. \tag{4.24}$$

Partition the deterministic time intervals into boxes of diameter at most Δ. For each resulting ordered box configuration, write $t_1 < \cdots < t_r$ for representative times. Discard the whole configuration if $|t_i - t_j| \leq 11\rho$ for some $i \neq j$. There are finitely many configurations for fixed Δ, ρ; the unmarked factorial convergence already proved above shows that their aggregate mass converges to the corresponding product-Lebesgue mass. The latter is $O_r(\rho)$, since it lies in the union of $\binom{r}{2}$ diagonal strips of width $22\rho + 2\Delta$. Thus this deletion costs $O_r(\rho) + o(1)$. In every retained configuration, arbitrary actual times in the chosen boxes differ by more than

$$11\rho - 2\Delta > 10\rho.$$

In particular, all ordering and separation restrictions are imposed by the boxes themselves, rather than by a later tuplewise deletion.

The unmarked factorial bounds also allow us to work on the following event. Fix once and for all $C_* > 3$, and set

$$\mathcal{G}_{L,\Delta} = \left\{ \sup_{0 \le n \le C_* L} \left| \frac{\log Q_n}{L} - \frac{n}{\kappa L} \right| \le \Delta \right\}. \tag{4.25}$$

The probability of its complement tends to zero by (4.19). Since the rare counts have bounded moments of every fixed order, their factorial powers are uniformly integrable. Indeed, for every fixed r, (4.17) and Cauchy–Schwarz give

$$\mathbb{E}\Big[R_L^r \mathbf{1}_{\mathcal{G}_{L,\Delta}^c}\Big] \le (\mathbb{E}R_L^{2r})^{1/2} \mathbb{P}(\mathcal{G}_{L,\Delta}^c)^{1/2} = o(1). \tag{4.26}$$

Here $\mathbb{E}, \mathbb{P}$ may refer to either Gauss or Lebesgue measure, because (4.17) transfers as above and $d\,\mathrm{Leb}\,/d\nu_G$ is bounded. Therefore inserting or deleting $\mathcal{G}_{L,\Delta}$ changes the factorial moments by $o(1)$. On $\mathcal{G}_{L,\Delta}$,

$$N^{t_j - 2\Delta} \le Q_{n_j} \le N^{t_j + 2\Delta}. \tag{4.27}$$

Let k be the largest index for which $h_k \ne 0$. The time separation and (4.27) give

$$D := \sum_{j \le k} h_j Q_{n_j}, \qquad |D| \asymp_{\mathbf{h}} Q_{n_k}. \tag{4.28}$$

Indeed, for $j < k$,

$$\frac{Q_{n_j}}{Q_{n_k}} \le N^{t_j - t_k + 4\Delta} \le N^{-10\rho + 6\Delta} = o(1),$$

so $D = h_k Q_{n_k}(1 + o(1))$, uniformly over the retained tuples. The integer numerator terms in (4.23) disappear, so its phase is $e(ND\alpha)$.

For a block ending at n_ℓ, let

$$\mathcal{B}_{\mathbf{n},\ell} = \bigcap_{j \le \ell} \left\{ N^{t_j - 2\Delta} \le Q_{n_j} \le N^{t_j + 2\Delta} \right\}. \tag{4.29}$$

This event is measurable on the cylinder through n_ℓ, and it contains $\mathcal{G}_{L,\Delta}$ for every retained tuple.

Oscillatory cylinder-sum estimate. We isolate the estimate used in both the early and late cases. Fix $1 \le \ell \le r$, let $\mathscr{I}_1, \dots, \mathscr{I}_\ell$ be the integer index boxes corresponding to representative times $t_1 < \cdots < t_\ell$, and assume that $k \le \ell$ is the largest index for which $h_k \ne 0$. Thus $\#\mathscr{I}_j = O(L)$. For $\mathbf{n} = (n_1, \dots, n_\ell)$, put, cylinderwise,

$$D_{\mathbf{n}} = \sum_{j \le k} h_j Q_{n_j}, \qquad E_{\mathbf{n},\ell} = \mathcal{B}_{\mathbf{n},\ell} \cap \bigcap_{j=1}^{\ell} \{L\theta_{n_j} \in J_j\}.$$

Then

$$\sum_{\substack{n_j \in \mathscr{I}_j \\ n_1 < \cdots < n_\ell}} \left| \int_0^1 e(ND_{\mathbf{n}}(\alpha)\alpha) \mathbf{1}_{E_{\mathbf{n},\ell}}(\alpha)\, d\alpha \right| \ll_{\mathbf{h},r,\mathbf{J}} L^{\ell-1} N^{2t_\ell - 1 - t_k + 6\Delta}. \tag{4.30}$$

Here is the complete count. Decompose the integral for a fixed tuple into the half-open continued-fraction cylinders C of depth n_ℓ. On C, all P_{n_j}, Q_{n_j}, $j \le \ell$, and hence $D_{\mathbf{n}}$, are constant. Moreover,

$$L\theta_{n_j} = (-1)^{n_j} L Q_{n_j} (Q_{n_j}\alpha - P_{n_j})$$

is affine in α. The intersection of C with all value constraints is therefore either empty or a single interval $I_{\mathbf{n},C}$, and $\mathcal{B}_{\mathbf{n},\ell}$ is constant on C. By the separation of the representative times and $\mathcal{B}_{\mathbf{n},\ell}$,

$$|D_{\mathbf{n},C}| \geq |h_k| Q_{n_k} - \sum_{j<k} |h_j| Q_{n_j} \gg_{\mathbf{h}} N^{t_k - 2\Delta}$$

for all sufficiently large N, uniformly over the retained tuples. Consequently,

$$\left| \int_{I_{\mathbf{n},C}} e(N D_{\mathbf{n},C} \alpha)\, d\alpha \right| \leq \frac{1}{\pi N |D_{\mathbf{n},C}|} \ll_{\mathbf{h}} N^{-1-t_k+2\Delta}. \tag{4.31}$$

A finite regular continued-fraction word is determined by its reduced terminal convergent P/Q, apart from the two usual terminal expansions. Indeed, the unique canonical expansion has last digit at least two, and its only alternative replaces the terminal digit a by $a-1, 1$; the Euclidean algorithm gives no other finite expansion. Hence the number of all finite cylinders whose terminal denominator is at most R is at most

$$2 \sum_{Q \leq R} \varphi(Q) + O(1) \ll R^2.$$

On $\mathcal{B}_{\mathbf{n},\ell}$ we may take $R = N^{t_\ell + 2\Delta}$. Once the deepest word, and hence n_ℓ, is fixed, there are at most $O(L^{\ell-1})$ choices for $n_1, \ldots, n_{\ell-1}$. Notice that the cylinder count already includes the summation over n_ℓ; no additional factor L is present. Multiplying this count by (4.31) gives

$$L^{\ell-1} N^{2t_\ell + 4\Delta} N^{-1-t_k+2\Delta},$$

which proves (4.30).

Replacing $\mathbf{1}_{\mathcal{G}_{L,\Delta}}$ by $\mathbf{1}_{\mathcal{B}_{\mathbf{n},\ell}}$ changes the aggregate factorial expression only on $\mathcal{G}^c_{L,\Delta}$, hence by $o(1)$ in view of (4.26). We use $\ell = r$ in the early case and $\ell = j_0$ only for the prefix marginal in the late case; no future denominator condition is passed through a mixing estimate.

Early and late oscillatory cases. Put $\sigma = (1+t_k)/2$. Discard every remaining box configuration for which

$$|t_j - \sigma| < 4\rho \quad \text{for some } j > k.$$

The unmarked product factorial measure bounds the prelimit discarded mass by $O_r(\rho) + o(1)$. If every later event has time below $\sigma - 4\rho$, apply (4.30) with $\ell = r$; its exponent is at most $-8\rho + 6\Delta$, so the contribution is $o(1)$.

It remains to treat the case in which a later event has time above $\sigma + 4\rho$. Let $j_0 \geq k$ be the last event below $\sigma - 4\rho$; include all constraints through j_0 in a prefix block. Write

$$\mathbf{n}_- = (n_1, \ldots, n_{j_0}), \qquad \mathbf{n}_+ = (n_{j_0+1}, \ldots, n_r).$$

Put $s = \sigma + \rho$ and use the deterministic depth

$$m_s = \lceil \kappa L(s + 2\Delta) \rceil. \tag{4.32}$$

Define the genuinely prefix-measurable event

$$\mathcal{G}^-_{L,\Delta}(m_s) = \left\{ \sup_{0 \leq n \leq m_s} \left| \frac{\log Q_n}{L} - \frac{n}{\kappa L} \right| \leq \Delta \right\}. \tag{4.33}$$

We have $\mathcal{G}_{L,\Delta} \subset \mathcal{G}^-_{L,\Delta}(m_s) \subset \mathcal{B}_{\mathbf{n},j_0}$, and the complement of $\mathcal{G}^-_{L,\Delta}(m_s)$ has probability $o(1)$. For the second inclusion, each retained index satisfies

$$\left| \frac{n_j}{\kappa L} - t_j \right| \leq \Delta;$$

combining this with the defining bound for $\mathcal{G}^-_{L,\Delta}(m_s)$ gives $|L^{-1}\log Q_{n_j} - t_j| \le 2\Delta$, exactly the condition in $\mathcal{B}_{\mathbf{n},j_0}$. Here $m_s < C_* L$ for all large N, by $t_k \le 1-\varepsilon$ and (4.24), so the first inclusion uses only the range present in (4.25). On $\mathcal{G}^-_{L,\Delta}(m_s)$,

$$Q_{m_s} \ge N^{s+\Delta}, \qquad m_s > n_{j_0}, \qquad n_{j_0+1} - m_s \ge c_\rho L.$$

Indeed,

$$\frac{n_{j_0}}{\kappa L} \le \sigma - 4\rho + \Delta, \qquad \frac{m_s}{\kappa L} \ge \sigma + \rho + 2\Delta,$$

while

$$\frac{n_{j_0+1}}{\kappa L} \ge \sigma + 4\rho - \Delta, \qquad \frac{m_s}{\kappa L} \le \sigma + \rho + 2\Delta + O(L^{-1}).$$

The two index gaps are therefore bounded below by a positive constant times ρL. The lower bound for Q_{m_s} follows from $m_s/(\kappa L) \ge s + 2\Delta$ and the defining error Δ in $\mathcal{G}^-_{L,\Delta}(m_s)$.

Late case: freezing the prefix. Let $\mathscr{N}_-$ denote the retained prefix tuples. Define first the shallow-cylinder prefix function

$$G^{\mathcal{B}}_{\mathbf{n}_-}(\alpha) = e(ND_{\mathbf{n}_-}(\alpha)\alpha)\mathbf{1}_{\mathcal{B}_{\mathbf{n},j_0}}(\alpha) \prod_{j\le j_0} \mathbf{1}_{\{L\theta_{n_j}(\alpha)\in J_j\}}.$$

The explicit cylinder estimate (4.30), now with $\ell = j_0$, gives

$$\begin{aligned}\sum_{\mathbf{n}_-\in\mathscr{N}_-} \left|\mathbb{E}_{\mathrm{Leb}} G^{\mathcal{B}}_{\mathbf{n}_-}\right| &\ll L^{j_0-1} N^{2t_{j_0}-1-t_k+6\Delta} \\ &\le L^{j_0-1}N^{-8\rho+6\Delta} = o(1). \end{aligned} \tag{4.34}$$

Next replace the local denominator event by the genuinely prefix-measurable good event:

$$G^-_{\mathbf{n}_-}(\alpha) = e(ND_{\mathbf{n}_-}(\alpha)\alpha)\mathbf{1}_{\mathcal{G}^-_{L,\Delta}(m_s)}(\alpha) \prod_{j\le j_0} \mathbf{1}_{\{L\theta_{n_j}(\alpha)\in J_j\}}.$$

Since $\mathcal{G}^-_{L,\Delta}(m_s) \subset \mathcal{B}_{\mathbf{n},j_0}$, the two functions agree on $\mathcal{G}^-_{L,\Delta}(m_s)$. Thus

$$\begin{aligned}\sum_{\mathbf{n}_-\in\mathscr{N}_-} \mathbb{E}_{\mathrm{Leb}} \left|G^-_{\mathbf{n}_-} - G^{\mathcal{B}}_{\mathbf{n}_-}\right| &\le \mathbb{E}_{\mathrm{Leb}}\left[R_L^{j_0}\mathbf{1}_{(\mathcal{G}^-_{L,\Delta}(m_s))^c}\right] \\ &\le \left(\mathbb{E}_{\mathrm{Leb}} R_L^{2j_0}\right)^{1/2} \mathrm{Leb}\big((\mathcal{G}^-_{L,\Delta}(m_s))^c\big)^{1/2} = o(1). \end{aligned} \tag{4.35}$$

We now freeze this function on depth-m_s cylinders. Let $\mathscr{C}_{m_s}$ be the half-open continued-fraction cylinders of depth m_s, which partition the irrational points up to a null set. We keep one fixed half-open endpoint convention and assign all rational endpoints arbitrarily; every identity in this paragraph is consequently an almost-everywhere identity, which is all that the integrals require. For every $C \in \mathscr{C}_{m_s}$, choose an irrational representative $\alpha_C \in C$, and write

$$D_{\mathbf{n}_-,C} = \sum_{j\le k} h_j Q_{n_j}(C).$$

The event $\mathcal{G}^-_{L,\Delta}(m_s)$ is constant on every such cylinder. Define

$$F_{\mathbf{n}_-}(\alpha) = \sum_{C\in\mathscr{C}_{m_s}} \mathbf{1}_C(\alpha)\mathbf{1}_{\{C\subset\mathcal{G}^-_{L,\Delta}(m_s)\}} e(ND_{\mathbf{n}_-,C}\alpha_C) \prod_{j\le j_0} \mathbf{1}_{\{L\theta_{n_j}(\alpha_C)\in J_j\}}. \tag{4.36}$$

This function is bounded and $\sigma(a_1,\dots,a_{m_s})$-measurable.

A depth-m_s cylinder has length at most $Q_{m_s}^{-2}$. On a cylinder retained by $\mathcal{G}_{L,\Delta}^{-}(m_s)$, the denominator bounds above imply

$$\sup_{\alpha\in C}\left|e(ND_{\mathbf{n}_-,C}\alpha)-e(ND_{\mathbf{n}_-,C}\alpha_C)\right| \ll_{\mathbf{h}} \frac{N^{1+t_k+2\Delta}}{N^{2s+2\Delta}} = O_{\mathbf{h}}(N^{-2\rho}) =: \delta_N. \tag{4.37}$$

Similarly, since $t_j \le t_{j_0} \le \sigma - 4\rho$,

$$\begin{aligned}\sup_{\alpha\in C}\left|L\theta_{n_j}(\alpha)-L\theta_{n_j}(\alpha_C)\right| &\le \frac{LQ_{n_j}^2}{Q_{m_s}^2}\\ &\ll LN^{2(\sigma-4\rho+2\Delta)-2(\sigma+\rho+\Delta)} \le LN^{-10\rho+2\Delta} =: \eta_N \qquad (j\le j_0).\end{aligned} \tag{4.38}$$

For an interval J, write $J^{(\eta)}$ for its closed η-neighborhood. The two freezing estimates give the pointwise bound

$$\begin{aligned}|G_{\mathbf{n}_-}^- - F_{\mathbf{n}_-}| \ll \delta_N &\prod_{j\le j_0}\mathbf{1}_{\{L\theta_{n_j}\in J_j^{(\eta_N)}\}}\\ &+\sum_{i\le j_0}\mathbf{1}_{\{\mathrm{dist}(L\theta_{n_i},\partial J_i)\le\eta_N\}}\prod_{\substack{j\le j_0\\ j\ne i}}\mathbf{1}_{\{L\theta_{n_j}\in J_j^{(\eta_N)}\}}.\end{aligned} \tag{4.39}$$

For every fixed $z>0$, uniformly for $0<\eta\le\min(1,z/2)$,

$$\nu_G\{|L\theta_n - z|\le\eta\} \ll_z \frac{\eta}{L}+\frac{1}{L^2}. \tag{4.40}$$

Indeed, on this event $L\theta_n$ remains in a fixed compact subset of $(0,\infty)$, and (4.10) places it inside

$$\left\{\left|\frac{L}{a_{n+1}}-z\right|\le\eta+\frac{C_z}{L}\right\}.$$

Direct summation of the one-digit probabilities proves (4.40).

For intersections with the other rare constraints, enlarge those constraints to their corresponding one-digit events and use (4.9). There are $O(L^{j_0})$ prefix tuples, so the total contribution of any one boundary coordinate is

$$O\left(L^{j_0}\left(\frac{\eta_N}{L}+\frac{1}{L^2}\right)L^{-(j_0-1)}\right) = O(\eta_N + L^{-1}).$$

The first term in (4.39) contributes $O(\delta_N)$ by (4.15). The densities of Gauss and Lebesgue measure relative to one another are bounded, so the estimates hold under either measure. Consequently,

$$\sum_{\mathbf{n}_-\in\mathcal{N}_-}\mathbb{E}_{\mathrm{Leb}}|G_{\mathbf{n}_-}^- - F_{\mathbf{n}_-}| \ll \delta_N+\eta_N+L^{-1} = o(1). \tag{4.41}$$

The same calculation with all r rare factors present gives

$$\sum_{\mathbf{n}_-,\mathbf{n}_+}\mathbb{E}_{\mathrm{Leb}}\left(|G_{\mathbf{n}_-}^- - F_{\mathbf{n}_-}|\prod_{j>j_0}\mathbf{1}_{\{L\theta_{n_j}\in J_j\}}\right) = o(1). \tag{4.42}$$

More precisely, after enlarging every exact rare event to its one-digit event, the phase term in this sum is

$$O(\delta_N L^r L^{-r}) = O(\delta_N),$$

whereas a term with one prefix boundary strip is

$$O\left(L^{r}\left(\frac{\eta_N}{L}+\frac{1}{L^2}\right)L^{-(r-1)}\right) = O(\eta_N + L^{-1}).$$

There are only $j_0 \le r$ possible boundary coordinates. This proves (4.42) without multiplying a merely prefix-level error by the number of future tuples. Combining (4.34), (4.35), and (4.41) proves the required absolute summability, because

$$\sum_{\mathbf{n}_-} |\mathbb{E}_{\mathrm{Leb}} F_{\mathbf{n}_-}| \le \sum_{\mathbf{n}_-} |\mathbb{E}_{\mathrm{Leb}} G^{\mathcal{B}}_{\mathbf{n}_-}| + \sum_{\mathbf{n}_-} \mathbb{E}_{\mathrm{Leb}} |G^{-}_{\mathbf{n}_-} - G^{\mathcal{B}}_{\mathbf{n}_-}| + \sum_{\mathbf{n}_-} \mathbb{E}_{\mathrm{Leb}} |F_{\mathbf{n}_-} - G^{-}_{\mathbf{n}_-}|.$$

Thus

$$\sum_{\mathbf{n}_- \in \mathscr{N}_-} |\mathbb{E}_{\mathrm{Leb}} F_{\mathbf{n}_-}| = o(1). \tag{4.43}$$

The frozen factor is itself dominated by enlarged one-digit rare events. In fact, if $F_{\mathbf{n}_-}(\alpha) \neq 0$ and $\alpha \in C$, then (4.38) gives

$$L\theta_{n_j}(\alpha) \in J_j^{(\eta_N)} \qquad (j \le j_0).$$

Applying (4.10) on this fixed enlarged compact window, with a constant uniform in $j \le j_0$, yields

$$|F_{\mathbf{n}_-}(\alpha)| \le \prod_{j \le j_0} \mathbf{1}_{\{L/a_{n_j+1} \in J_j^{(\eta_N + C_{\mathbf{J}}/L)}\}}. \tag{4.44}$$

Late case: prefix–future mixing. All later mark frequencies are zero. By (4.10), their exact value constraints can be replaced, with aggregate error $o(1)$, by the one-digit constraints $L/a_{n+1} \in J$. To see that this remains true after multiplication by $F_{\mathbf{n}_-}$, use (4.44). If a single future indicator changes, its digit lies in an endpoint strip of probability $O(L^{-2})$; the frozen prefix and the remaining future coordinates are bounded by $r-1$ one-digit rare events of probability $O(L^{-1})$. Hence (4.9) gives the full aggregate bound

$$O\big(L^r L^{-2} L^{-(r-1)}\big) = O(L^{-1}).$$

Let $H_{\mathbf{n}_+}$ denote the product of these future one-digit indicators, and let $\mathscr{N}_+$ be the retained future tuples. Then $H_{\mathbf{n}_+}$ is measurable with respect to digits beginning at $a_{n_{j_0+1}+1}$, at distance at least $c_\rho L$ from the prefix σ-field. For the fixed separated collection of index boxes now under consideration, every retained prefix tuple precedes every retained future tuple, and the boxes themselves enforce all ordering and separation conditions. Hence the retained full tuples are exactly the Cartesian product $\mathscr{N}_- \times \mathscr{N}_+$, which justifies the product double sums below.

We also replace the global event $\mathcal{G}_{L,\Delta}$ by $\mathcal{G}^-_{L,\Delta}(m_s)$. Since the former is contained in the latter, note explicitly that on any fixed digit window $L/a_{n+1} \in K_0 \Subset (0,\infty)$,

$$0 \le \frac{L}{a_{n+1}} - L\theta_n = \frac{L(x_{n+1}+y_n)}{a_{n+1}(a_{n+1}+x_{n+1}+y_n)} \le \frac{2L}{a_{n+1}^2} = O_{K_0}(L^{-1}).$$

Thus all digit-replaced events occurring below are contained, for large L, in a fixed enlargement of the corresponding exact θ-windows. Take R_L here on a fixed slightly enlarged compact value window containing every exact prefix window, every enlargement in (4.44), and every digit-replaced future window. It therefore dominates every summand that occurs in this replacement, and the aggregate difference is at most

$$\mathbb{E}\left[R_L^r \mathbf{1}_{\mathcal{G}^c_{L,\Delta}}\right] = o(1)$$

by (4.26). Together with (4.42), this shows that the late part of the original factorial Fourier coefficient is

$$\sum_{\mathbf{n}_- \in \mathscr{N}_-} \sum_{\mathbf{n}_+ \in \mathscr{N}_+} \mathbb{E}_{\mathrm{Leb}}(F_{\mathbf{n}_-} H_{\mathbf{n}_+}) + o(1). \tag{4.45}$$

Recall $w = d\,\mathrm{Leb}/d\nu_G$ and $w_L = \mathbb{E}_{\nu_G}(w \mid a_1, \ldots, a_{b_L})$. Since $b_L < m_s$, $w_L F_{\mathbf{n}_-}$ is prefix measurable. The estimates just proved, together with (4.15), give

$$\sum_{\mathbf{n}_-} \mathbb{E}_{\nu_G}|w_L F_{\mathbf{n}_-}| = O(1), \qquad \sum_{\mathbf{n}_+} \mathbb{E}_{\nu_G} H_{\mathbf{n}_+} = O(1).$$

The functional mixing inequality (4.8) now yields the aggregate covariance bound

$$\begin{aligned}\sum_{\mathbf{n}_-,\mathbf{n}_+} &|\mathbb{E}_{\nu_G}(w_L F_{\mathbf{n}_-} H_{\mathbf{n}_+}) - \mathbb{E}_{\nu_G}(w_L F_{\mathbf{n}_-})\mathbb{E}_{\nu_G} H_{\mathbf{n}_+}| \\ &\le 2C\vartheta^{-c_\rho L} \left(\sum_{\mathbf{n}_-} \mathbb{E}_{\nu_G}|w_L F_{\mathbf{n}_-}|\right)\left(\sum_{\mathbf{n}_+} \mathbb{E}_{\nu_G} H_{\mathbf{n}_+}\right) = o(1). \end{aligned} \tag{4.46}$$

Applying the same inequality to $|F_{\mathbf{n}_-}|$ and $H_{\mathbf{n}_+}$ shows that

$$\sum_{\mathbf{n}_-,\mathbf{n}_+} \mathbb{E}_{\nu_G}|F_{\mathbf{n}_-}|H_{\mathbf{n}_+} = O(1).$$

Thus replacing w by w_L in (4.45) costs at most

$$\|w - w_L\|_\infty \sum_{\mathbf{n}_-,\mathbf{n}_+} \mathbb{E}_{\nu_G}|F_{\mathbf{n}_-}|H_{\mathbf{n}_+} = o(1).$$

Moreover,

$$\sum_{\mathbf{n}_-} |\mathbb{E}_{\nu_G}(w_L F_{\mathbf{n}_-}) - \mathbb{E}_{\mathrm{Leb}} F_{\mathbf{n}_-}| \le \|w - w_L\|_\infty \sum_{\mathbf{n}_-} \mathbb{E}_{\nu_G}|F_{\mathbf{n}_-}| = o(1). \tag{4.47}$$

For completeness, (4.22), applied to the future tuple indicators, also gives

$$\sum_{\mathbf{n}_+} |\mathbb{E}_{\mathrm{Leb}} H_{\mathbf{n}_+} - \mathbb{E}_{\nu_G} H_{\mathbf{n}_+}| = o(1). \tag{4.48}$$

Combining (4.45), (4.46), (4.47), and (4.43), the absolute value of the late contribution is at most

$$\left(\sum_{\mathbf{n}_-} |\mathbb{E}_{\mathrm{Leb}} F_{\mathbf{n}_-}|\right)\left(\sum_{\mathbf{n}_+} \mathbb{E}_{\nu_G} H_{\mathbf{n}_+}\right) + o(1) = o(1).$$

Finally let $N \to \infty$, then $\Delta \downarrow 0$, and then $\rho \downarrow 0$. We have proved that every nonzero joint Fourier coefficient of the marked factorial measures vanishes.

Signed extension and the point-process criterion. We have so far written the argument for positive value intervals and prescribed parity. This covers the whole signed state space. Indeed, by (4.5),

$$(-1)^n L\theta_n > 0 \quad \Longleftrightarrow \quad n \text{ is even.}$$

Thus an interval $J \Subset (0, \infty)$ selects even indices, whereas an interval $J \Subset (-\infty, 0)$ selects odd indices after reflection to the positive interval $-J$. Reflection preserves Lebesgue length, so both signed halves have the same intensity. Every compact subset of $\mathbb{R} \setminus \{0\}$ has positive distance from zero and is the union of its positive and negative parts. Finite disjoint decompositions therefore reduce every relatively compact signed rectangle to the two cases just proved. No rectangle accumulating at zero is needed, because such a set is not relatively compact in $\mathbb{R} \setminus \{0\}$.

Trigonometric polynomials are uniformly dense in the continuous functions on $\mathbb{T}^r$. The zero Fourier coefficient is the unmarked product factorial measure, and all other coefficients vanish. Hence, the factorial measures converge weakly in the torus variables. For a Haar-continuity rectangle, sandwich its torus indicator between continuous functions whose Haar integrals differ by an arbitrarily small amount; the uniform unmarked factorial bound controls the corresponding aggregate error. Therefore, on $[\varepsilon, 1-\varepsilon]$, all limiting factorial measures are

$$\lambda^r \prod_{j=1}^{r} dt_j \, d\xi_j \, dm(u_j).$$

We now spell out the passage from factorial measures to point processes. Let

$$\widetilde{\mathcal{Q}}_N = \sum_{n \geq 0} \delta_{(n/(\kappa L),\, (-1)^n L\theta_n,\, (-1)^n N\theta_n/Q_n \bmod 1)}$$

be the process in deterministic index time. On

$$E_\varepsilon = [\varepsilon, 1-\varepsilon] \times (\mathbb{R} \setminus \{0\}) \times \mathbb{T},$$

let $\mathscr{R}_\varepsilon$ be the countable semiring of relatively compact half-open rectangles whose Euclidean endpoints are rational, whose value-coordinate closures avoid zero, and whose torus endpoints belong to a fixed countable dense subset of $\mathbb{T}$. A torus arc that wraps around zero is split into two arcs. This is a dissecting, convergence-determining semiring, and every member is a continuity set for the diffuse measure

$$\Lambda = \lambda \, dt \, d\xi \, dm.$$

Take pairwise disjoint $B_1, \ldots, B_q \in \mathscr{R}_\varepsilon$, and write $X_{N,i} = \widetilde{\mathcal{Q}}_N(B_i)$. The factorial-measure calculation above gives, for every $s_1, \ldots, s_q \geq 0$,

$$\mathbb{E} \prod_{i=1}^{q} (X_{N,i})_{s_i} \longrightarrow \prod_{i=1}^{q} \Lambda(B_i)^{s_i}. \tag{4.49}$$

The right-hand side is the joint factorial moment of independent Poisson variables Z_i with means $\Lambda(B_i)$.

Let $T_N = \sum_i X_{N,i}$. After reflecting every negative value window to its positive magnitude, T_N is bounded by a finite union of rare counts of the form used in (4.15). Consequently, for every fixed m,

$$\sup_N \mathbb{E}(T_N)_m < \infty, \qquad \sup_N \mathbb{E}T_N^m < \infty,$$

where the second bound follows from (4.16). In particular, the bound with one order larger makes every polynomial of a fixed degree in $(X_{N,1}, \ldots, X_{N,q})$ uniformly integrable. Every ordinary joint moment is a finite linear combination of the joint factorial moments in (4.49); hence all ordinary joint moments converge to those of $(Z_1, \ldots, Z_q)$. That law is moment determinate because its moment-generating function

$$\exp\left(\sum_{i=1}^{q} \Lambda(B_i)(e^{t_i} - 1) \right)$$

is finite in a neighborhood of the origin. The multivariate method of moments therefore yields

$$(X_{N,1}, \ldots, X_{N,q}) \Longrightarrow (Z_1, \ldots, Z_q). \tag{4.50}$$

For every finite union U of members of $\mathscr{R}_\varepsilon$, we have $\Lambda(\partial U) = 0$, and the first factorial moment and (4.50) give

$$\mathbb{E}\widetilde{\mathcal{Q}}_N(U) \longrightarrow \Lambda(U), \qquad \mathbb{P}\{\widetilde{\mathcal{Q}}_N(U) = 0\} \longrightarrow e^{-\Lambda(U)},$$

respectively. Kallenberg's mean-and-void criterion states that these two limits on the finite unions of a countable dissecting semiring of continuity sets imply convergence to the Poisson random measure [6, Theorem 4.7]. Applied to $\mathscr{R}_\varepsilon$ on this locally compact second-countable space, it therefore gives vague convergence to the Poisson process with intensity Λ. Notice also that the mean bound supplies local tightness directly: every compact set is covered by such a finite union U, and $\sup_N \mathbb{E}\widetilde{\mathcal{Q}}_N(U) < \infty$, so Markov's inequality controls the probability of more than M points uniformly as $M \to \infty$. Hence the marked point process converges in deterministic index time.

Time change and endpoint restoration. We give the time-change estimate at the level of test functions. Let f be continuous with compact support whose time projection lies in $[\varepsilon, 1-\varepsilon]$, and let

$$\omega_f(\delta) = \sup_{\substack{|t-t'|\le\delta \\ \xi\in\mathbb{R}\setminus\{0\},\, u\in\mathbb{T}}} |f(t,\xi,u) - f(t',\xi,u)|.$$

Extend f by zero outside a slightly larger compact time interval; then $\omega_f(\delta) \to 0$ as $\delta \downarrow 0$. Because the value projection of $\operatorname{supp} f$ is compact in $\mathbb{R}\setminus\{0\}$, choose a compact interval $K_f \Subset (0,\infty)$ containing

$$\{|\xi| : (t,\xi,u) \in \operatorname{supp} f\}.$$

Let $R_L^{(f)}$ count all rare convergents with $0 \le n \le C_* L$ and magnitude $L\theta_n \in K_f$. This contains every index-time point relevant to f. It also contains every denominator-time point relevant to f, because $Q_n \ge F_{n+1}$ and $Q_n < N$ imply $n < C_* L$ for all sufficiently large N. By (4.17),

$$\sup_L \mathbb{E} R_L^{(f)} + \sup_L \mathbb{E}(R_L^{(f)})^2 < \infty.$$

On $\mathcal{G}_{L,\Delta}$, with $\Delta < \varepsilon/2$, every relevant convergent satisfies

$$\left| \frac{n}{\kappa L} - \frac{\log Q_n}{L} \right| \le \Delta.$$

Moreover, a point with either time coordinate in $\operatorname{supp} f$ has the other coordinate in $[\varepsilon/2, 1-\varepsilon/2]$. In particular $Q_n < N$, so the cutoff in (4.6) is present on both sides. Using the continued-fraction identification (4.6), we obtain

$$\left| \langle \widetilde{\mathcal{Q}}_N, f\rangle - \langle \mathcal{Q}_N, f\rangle \right| \le \omega_f(\Delta) R_L^{(f)} \qquad \text{on } \mathcal{G}_{L,\Delta}. \tag{4.51}$$

On the complement, the same difference is at most $2\|f\|_\infty R_L^{(f)}$. Hence

$$\begin{aligned}\mathbb{E}\left| \langle \widetilde{\mathcal{Q}}_N, f\rangle - \langle \mathcal{Q}_N, f\rangle \right| &\le \omega_f(\Delta)\mathbb{E}R_L^{(f)} + 2\|f\|_\infty \mathbb{E}\left[R_L^{(f)} \mathbf{1}_{\mathcal{G}^c_{L,\Delta}}\right] \\ &\le O_f(\omega_f(\Delta)) + o_{N\to\infty}(1),\end{aligned}$$

where Cauchy–Schwarz, the second-moment bound, and $\mathbb{P}(\mathcal{G}^c_{L,\Delta}) = o(1)$ were used in the last line. First let $N \to \infty$, and then let $\Delta \downarrow 0$. The converging-together argument for a countable convergence-determining family of such f's transfers the marked process from index time to denominator time on every fixed interior layer. The endpoint layers are restored next by the direct cell bound.

It remains only to restore the two endpoint layers. Here no time-change argument is needed. Directly subdividing α into the reduced rational cells about q/p gives, for every set of denominators $I \subset [1,N)$,

$$\mathbb{E}\#\{p \in I : (p,q_p) = 1, |Lp\delta_p| \le B\} \le \frac{2B}{L}\sum_{p\in I}\frac{\varphi(p)}{p^2} + O_B(L^{-1}). \tag{4.52}$$

Indeed, for each of the $\varphi(p)$ reduced numerators the relevant α-interval has length at most $2B/(Lp^2)$. For $p > 1$ these reduced cells are interior; the $O_B(L^{-1})$ error covers the two endpoint cells for $p = 1$. We also have

$$\sum_{p\le x} \frac{\varphi(p)}{p^2} = \frac{6}{\pi^2}\log x + O(1),$$

as follows by inserting

$$\frac{\varphi(p)}{p} = \sum_{d|p} \frac{\mu(d)}{d}$$

and interchanging the sums. The resulting inner sum uses $\sum_{m\le y} m^{-1} = \log y + O(1)$, while the leading coefficient is $\sum_{d\ge 1} \mu(d)/d^2 = 1/\zeta(2)$. Therefore the expected number of points in either discarded time layer $[0,\varepsilon]$ or $[1-\varepsilon,1]$, on a fixed ξ-window, is $O_B(\varepsilon) + o(1)$. More explicitly, if $f \in C_c([0,1]\times(\mathbb{R}\setminus\{0\})\times\mathbb{T})$, its value projection is contained in $\{|\xi| \le B_f\}$ for some B_f. Choose a continuous function $\chi_\varepsilon : [0,1] \to [0,1]$ which is zero on $[0,\varepsilon]\cup[1-\varepsilon,1]$ and one on $[2\varepsilon, 1-2\varepsilon]$. Then $f(t,\xi,u)\chi_\varepsilon(t)$ is a compactly supported continuous test function on an interior layer, and

$$\mathbb{E}\,|\langle \mathcal{Q}_N, f\rangle - \langle \mathcal{Q}_N, f\chi_\varepsilon\rangle| \le \|f\|_\infty (O_{B_f}(\varepsilon) + o(1)).$$

The limiting Poisson process satisfies the analogous bound $O_f(\varepsilon)$ by its intensity measure. Markov's inequality and the converging-together argument therefore restore the two layers. Letting $\varepsilon \downarrow 0$ proves (4.2). Projection in the t-coordinate is proper because $[0,1]$ is compact, so the point-process mapping theorem proves (4.4). $\square$

4.2 Finite Poisson shots

For $A > 0$, define

$$X_{N,A} = \sum_{\substack{p\le N,\ (p,q_p)=1\\ |Lp\delta_p|\le A}} \frac{V(N\delta_p)}{Lp\delta_p}. \tag{4.53}$$

The $p = N$ term may be deleted: the measure of its support is $O_A((LN)^{-1})$.

Continue to write $\mathcal{P} = \{(\xi_j, u_j)\}$ for the Poisson process with intensity (4.4), and put

$$X_A = \sum_{|\xi_j|\le A} \frac{V(u_j)}{\xi_j}. \tag{4.54}$$

The sum is almost surely finite and no point has $\xi_j = 0$.

Lemma 4.2 (Finite-shot convergence)**.** *For every fixed $A > 0$,*

$$X_{N,A} \Longrightarrow X_A.$$

Proof. For $0 < r < A$, restrict both point processes to $r \le |\xi| \le A$. The shot functional is continuous at every simple configuration with no point on the boundary, an event of limiting probability one. Lemma 4.1 and the continuous mapping theorem give convergence of the restricted shots.

It remains to let $r \downarrow 0$. For the prelimit, a union bound and subdivision into reduced cells give

$$\mathrm{Leb}\{\mathcal{P}_N((-r,r)\times\mathbb{T}) > 0\} \le \frac{2r}{L}\sum_{p\le N}\frac{\varphi(p)}{p^2} \ll r.$$

The limiting probability is at most $2r/\zeta(2)$. With probability $1 - O(r)$, deleting $|\xi| < r$ changes neither shot. Let $r \downarrow 0$. $\square$

4.3 The Cauchy exponent

Lemma 4.3 (Poisson integral)**.** *As $A \to \infty$,*

$$X_A \Longrightarrow X, \qquad \mathbb{E}e^{\mathrm{i}tX} = \exp\left(-\frac{|t|}{2\pi}\right).$$

Proof. The exponential formula for a Poisson process and symmetry in ξ give

$$\begin{aligned}\log \mathbb{E}e^{\mathrm{i}tX_A} &= \frac{1}{\zeta(2)}\int_{-A}^{A}\int_0^1 \left(e^{\mathrm{i}tV(u)/\xi} - 1\right) \mathrm{d}u\,\mathrm{d}\xi \\ &= \frac{2}{\zeta(2)}\int_0^A\int_0^1 (\cos(tV(u)/\xi) - 1)\, \mathrm{d}u\,\mathrm{d}\xi.\end{aligned}$$

For $a \geq 0$,

$$\int_0^\infty (\cos(a/\xi) - 1)\, \mathrm{d}\xi = a\int_0^\infty \frac{\cos y - 1}{y^2}\, \mathrm{d}y = -\frac{\pi a}{2}.$$

Moreover,

$$0 \leq 1 - \cos(tV(u)/\xi) \leq \min\left(2, \frac{t^2V(u)^2}{2\xi^2}\right) \leq \min(2, C_t\xi^{-2}),$$

which is integrable for $\xi > 0$. Dominated convergence in the nonpositive symmetric integrand therefore yields

$$\lim_{A\to\infty} \log \mathbb{E}e^{\mathrm{i}tX_A} = -\frac{\pi|t|}{\zeta(2)}\int_0^1 V(u)\,\mathrm{d}u.$$

Since

$$\int_0^1 V(u)\,\mathrm{d}u = \frac{1}{12}, \qquad \zeta(2) = \frac{\pi^2}{6},$$

the exponent is $-|t|/(2\pi)$. Lévy's continuity theorem completes the proof. □

4.4 Completion of the proof

Proof of Theorem 1.1. By Proposition 2.7,

$$\frac{S_N - Y_N}{L} \longrightarrow 0 \quad \text{in } L^2, \text{ hence in probability.}$$

For fixed A, the part of Y_N/L with $|Lp\delta_p| \leq A$ is $X_{N,A}$. Proposition 3.5 states precisely that the complementary part goes to zero in probability when first $N \to \infty$ and then $A \to \infty$. Lemma 4.2 gives $X_{N,A} \Rightarrow X_A$ for fixed A, and Lemma 4.3 gives $X_A \Rightarrow X$. Moreover, for every $\eta > 0$, the union bound gives

$$\begin{aligned}\mathbb{P}\left\{\left|\frac{S_N}{L} - X_{N,A}\right| > \eta\right\} \leq{}& \mathbb{P}\left\{\frac{|S_N - Y_N|}{L} > \frac{\eta}{2}\right\} \\ &+ \mathbb{P}\left\{\left|\frac{Y_N}{L} - X_{N,A}\right| > \frac{\eta}{2}\right\}.\end{aligned}$$

Consequently,

$$\lim_{A\to\infty}\limsup_{N\to\infty} \mathbb{P}\left\{\left|\frac{S_N}{L} - X_{N,A}\right| > \eta\right\} = 0.$$

The converging-together theorem [2, Theorem 3.2] now yields

$$\frac{S_N}{\log N} \Longrightarrow X.$$

The Cauchy law with characteristic function $\exp(-|t|/(2\pi))$ has continuous distribution function

$$g(c) = \frac{1}{2} + \frac{1}{\pi}\arctan(2\pi c).$$

Weak convergence therefore gives convergence of the distribution functions at every $c \in \mathbb{R}$. The function g is nondecreasing, with limits 0 and 1 at $-\infty$ and $+\infty$, respectively. □

5 Two elementary consistency checks

For completeness, the finite-N law has no atoms. Away from the finitely many points a/k, $k \le N$,

$$S_N(\alpha) = \frac{N}{2} - \frac{N(N+1)}{2}\alpha + \sum_{k=1}^{N} \lfloor k\alpha \rfloor .$$

It is therefore affine with nonzero slope on every complementary interval. Also, away from those points,

$$S_N(1-\alpha) = -S_N(\alpha).$$

Thus every finite-N law is exactly symmetric and $\mathrm{Leb}\{S_N \le 0\} = 1/2$, in agreement with the limiting Cauchy law.

The second moment is not uniformly integrable on the logarithmic scale. Fourier orthogonality gives

$$\int_0^1 \psi(k\alpha)\psi(\ell\alpha)\,\mathrm{d}\alpha = \frac{(k,\ell)^2}{12k\ell},$$

so positivity of all covariance terms and the diagonal terms give

$$\mathbb{E}\left(\frac{S_N}{\log N}\right)^2 = \frac{\mathrm{Var}\, S_N}{(\log N)^2} \ge \frac{N}{12(\log N)^2} \longrightarrow \infty.$$

Thus $\{(S_N/\log N)^2\}_N$ is not uniformly integrable. This divergence of second moments does not contradict the weak convergence in Theorem 1.1.

References

[1] G. D. Birkhoff, Proof of the ergodic theorem, *Proc. Natl. Acad. Sci. USA* **17** (1931), 656–660, doi:10.1073/pnas.17.12.656.

[2] P. Billingsley, *Convergence of Probability Measures*, second edition, Wiley, New York, 1999.

[3] T. H. Chan and A. V. Kumchev, On sums of Ramanujan sums, *Acta Arith.* **152** (2012), 1–10, doi:10.4064/aa152-1-1.

[4] P. Erdős, Problems and results on diophantine approximations, *Compositio Math.* **16** (1964), 52–65, https://www.numdam.org/item/CM_1964__16__52_0/.

[5] A. Ghosh, M. S. Kirsebom, and P. Roy, Continued fractions, the Chen–Stein method and extreme value theory, *Ergodic Theory Dynam. Systems* **41** (2021), 461–470, doi:10.1017/etds.2019.64.

[6] O. Kallenberg, *Random Measures*, third edition, Akademie-Verlag, Berlin, 1983.

[7] H. Kesten, Uniform distribution mod 1, *Ann. of Math. (2)* **71** (1960), 445–471.

[8] H. Kesten, Uniform distribution mod 1. II, *Acta Arith.* **7** (1961/1962), 355–380.

[9] A. Ya. Khinchin, *Continued Fractions*, University of Chicago Press, Chicago, 1964.