

THE PÓLYA–VINOGRADOV INEQUALITY WITH $\log \log q$ FOR EVERY CHARACTER

DEEP BHATTACHARJEE  AND UHASHI BHATTACHARYA 

ABSTRACT. Let χ be a primitive Dirichlet character modulo $q \geq 3$ and let $M(\chi)$ be the largest modulus of a partial sum $\sum_{n \leq x} \chi(n)$. The Pólya–Vinogradov inequality bounds $M(\chi)$ by a constant times $\sqrt{q} \log q$. Paley showed that $\sqrt{q} \log \log q$ is the most that can be hoped for, and Montgomery and Vaughan proved that this is the truth if the generalised Riemann hypothesis holds. Without any hypothesis we prove that $M(\chi) \leq C\sqrt{q} \log \log q$ for every primitive character χ to every modulus $q \geq 3$, with an absolute constant C , and that $|L'/L(1, \chi)| \leq C \log \log q$ for every nonprincipal character. We also prove, from the elementary bounds of Chebyshev alone, that for every $\varepsilon > 0$ there is a constant C_ε such that at most q^ε of the primitive characters modulo q have $M(\chi) > C_\varepsilon \sqrt{q} \log \log q$, and at most q^ε of the nonprincipal characters have $|L'/L(1, \chi)| > C_\varepsilon \log \log q$. The proof has two parts. The first is a transfer theorem for a single character: if $\theta < 1$ is fixed and the L -functions of the twists $\chi\psi$ of χ by the characters ψ of the moduli up to $(\log q)^{16}$ have no zero in the rectangle $\operatorname{Re} s > \theta$, $|\Im s| \leq 5q^2$, then $M(\chi) \leq C_\theta \sqrt{q} \log \log q$. It is proved by a contour shift for Riesz means of prime sums, a dissection of the circle into major and minor arcs, and a weak form of the Montgomery–Vaughan bound for exponential sums with multiplicative coefficients, which we prove with Selberg’s sieve as the only arithmetic input. The second part counts the characters that fail the hypothesis of the transfer theorem by a zero-density theorem in the manner of Ingham and Montgomery, which bounds the number of zeros of all L -functions to a modulus in such a rectangle by a power of the modulus that tends to zero as θ tends to one. The same transfer theorem shows that a zero-free half-plane $\operatorname{Re} s > \theta$ for all Dirichlet L -functions gives the bound for every character, and such a half-plane exists: OpenAI has proved that every Dirichlet L -function, and every finite-order Hecke L -function over $\mathbb{Q}(\sqrt{-3})$, has no zero with $\operatorname{Re} s > 11/12$, by a mean square of twisted Möbius sums over the Eisenstein integers, two Poisson summations, the transformation of Kubota’s cubic theta function and the quadratic large sieve. Their proof is reproduced in full in a second appendix, with credit to them, so that the paper is complete in itself: the zero-density theorem is proved in the first appendix with explicit constants, and every statement of the paper is proved in the text.

Note. The results of this paper were first made public on 15 October 2026, as release v1.0.0 of the repository <https://github.com/creelie/Breakthrough-problem-s-3> (concept DOI 10.5281/zenodo.23383715); later versions are archived on Zenodo as 10.5281/zenodo.23563651, 10.5281/zenodo.23629139, 10.5281/zenodo.23629649, 10.5281/zenodo.23637263 and 10.5281/zenodo.23638793. An earlier and incomplete version of the paper appeared on Hexagon on 15 October 2026 as the preprint hexagon:2610.00256, available at <https://hexagonmath.org/2610.00256>, and was revised there on 21 October 2026. That version claimed the bound $\sqrt{q} \log \log q$ for every primitive character; it deduced it from a zero-free half-plane for all Dirichlet L -functions announced in an unrefereed preprint of another group, which was cited as an external input, and it had a short abstract. The second version, release v2.0.0 of the same repository, archived as 10.5281/zenodo.23671158, proved every statement except that one in the paper, the zero-density theorem of Montgomery included, from

Date: 26 October 2026.

2020 Mathematics Subject Classification. 11L40; 11M06; 11M26; 11N36.

Key words and phrases. Pólya–Vinogradov inequality; character sums; zero-density theorems; Selberg’s sieve.

Corresponding author: Deep Bhattacharjee. Email: d.bhattacharjee@erl-forschung.de, itsdeep@live.com, eggspl@proton.me; ORCID 0000-0003-0466-750X..

Ushashi Bhattacharya: email rai.bhattacharya.in@gmail.com; ORCID 0009-0002-2254-3914.

Chebyshev's bounds alone, and stated the bound for every character as a theorem resting on the zero-free half-plane $\operatorname{Re} s > 7/8$ of that preprint, whose machine-checked proof the authors compiled and checked. In this version the paper is complete in itself: Appendix B reproduces in full, with credit to OpenAI, their proof of the zero-free half-plane $\operatorname{Re} s > 11/12$ for all finite-order Hecke L -functions over $\mathbb{Q}(\sqrt{-3})$, from which Theorem 1.5 follows by the transfer theorem, and the title states the result for every character. The source of this revision is in release v3.0.0 of the same repository, archived as 10.5281/zenodo.23672710, with programs that check the explicit constants of the proofs, compute the sums of this paper for small moduli, and contain the compiled Lean 4 proof of the half-plane $\operatorname{Re} s > 7/8$ as an independent check.

1. INTRODUCTION

Let χ be a primitive Dirichlet character modulo $q \geq 3$ and put

$$M(\chi) = \max_{x \geq 1} \left| \sum_{n \leq x} \chi(n) \right|.$$

The Pólya–Vinogradov inequality states that $M(\chi) \ll \sqrt{q} \log q$ [6, Chapter 23].¹ Paley proved that the factor $\log q$ cannot be replaced by anything smaller than $\log \log q$: for infinitely many primitive quadratic characters $M(\chi) > c\sqrt{q} \log \log q$ [25]. Montgomery and Vaughan proved the matching upper bound $M(\chi) \ll \sqrt{q} \log \log q$ under the generalised Riemann hypothesis [22]. Without hypothesis, Granville and Soundararajan lowered the exponent of $\log q$ for characters of fixed odd order [10],² and Bober, Goldmakher, Granville and Koukoulopoulos determined the distribution of $M(\chi)/(\sqrt{q} \log \log q)$ over the characters to a prime modulus in the range where that ratio is at most $\log \log q$ minus a large constant [4].

Montgomery and Vaughan use the Riemann hypothesis in one place: on the major arcs of a circle method, for the prime sums of the characters $\chi\psi$ twisted by additive characters, where ψ runs over the characters of a small modulus r . What that step needs is a power saving in these prime sums up to the height q , and a power saving follows from the absence of zeros of $L(s, \chi\psi)$ in a rectangle of fixed width to the left of the line $\operatorname{Re} s = 1$ and of height a power of q . Our first theorem makes this precise.

Theorem 1.1 (transfer theorem). *Let $1/2 \leq \theta < 1$. There is a constant C_θ with the following property. Let χ be a primitive character modulo $q \geq 3$, put $M = \lceil \log q \rceil$ and $R = M^{16}$, and suppose that for every $r \leq R$ and every character ψ modulo r ,*

$$L(s, \chi\psi) \neq 0 \quad \text{for } \operatorname{Re} s > \theta, |\Im s| \leq 5q^2. \quad (1.1)$$

Then $M(\chi) \leq C_\theta \sqrt{q} \log \log q$.

The number of zeros of all the L -functions to a modulus in such a rectangle is bounded by a zero-density theorem of the kind first proved by Ingham and Montgomery, Theorem 2.3 below, which we prove in Appendix A: the bound is a power of the modulus whose exponent tends to zero as θ tends to one, times a power of the logarithm. Counting the characters χ for which (1.1) fails gives our main result.

Theorem 1.2. *For every $\varepsilon > 0$ there is a constant C_ε such that for every $q \geq 3$ the number of primitive characters χ modulo q with*

$$M(\chi) > C_\varepsilon \sqrt{q} \log \log q$$

is at most q^ε .

¹The proof expands χ in additive characters through the Gauss sum $\tau(\bar{\chi})$, whose modulus is $\sqrt{q}$ for primitive χ , and bounds the resulting geometric series trivially; Lemma 5.1 below is the same argument with the harmonic sums kept.

²For a primitive character of odd order g they obtain $M(\chi) \ll \sqrt{q}(\log q)^{1-\delta_g/2+o(1)}$ with $\delta_g = 1 - \frac{g}{\pi} \sin \frac{\pi}{g}$, and under the generalised Riemann hypothesis the same exponent of $\log \log q$.

For a prime q there are $q - 2$ primitive characters, so Theorem 1.2 gives the bound of Montgomery and Vaughan for all but a vanishing proportion $q^{\varepsilon-1}$ of them. We do not know a reference for this statement. The results of Granville and Soundararajan and of Bober, Goldmakher, Granville and Koukoulopoulos on the distribution of large character sums [10, 4] are proved by a different method, through the structure of the characters with large sums, and they bound the number of characters with $M(\chi)$ above a given multiple of $\sqrt{q} \log \log q$ only in the range of multiples below $\log \log q$ minus a large constant; the complementary range, which is the one of Theorem 1.2, is where the zero-density theorem enters.

The transfer theorem also shows what a zero-free half-plane gives.

Corollary 1.3. *Let $1/2 \leq \theta < 1$ and suppose that $L(s, \xi) \neq 0$ for $\operatorname{Re} s > \theta$ for every nonprincipal Dirichlet character ξ . Then $M(\chi) \leq C_\theta \sqrt{q} \log \log q$ for every primitive character χ modulo $q \geq 3$, with the constant of Theorem 1.1.*

Such a half-plane exists, with $\theta = 11/12$: it is Theorem B.1 of Appendix B, due to OpenAI, and Theorem 1.5 below records what it gives.

The same zero-free rectangle controls the logarithmic derivative of $L(s, \chi)$ at $s = 1$, and here the rectangle needed is small: its height is a power of $\log q$.

Theorem 1.4. *For every $\varepsilon > 0$ there is a constant C_ε such that for every $q \geq 3$ the number of nonprincipal characters χ modulo q with*

$$\left| \frac{L'}{L}(1, \chi) \right| > C_\varepsilon \log \log q$$

is at most q^ε .

Theorems 1.1 to 1.4 are proved in this paper from Chebyshev's bounds alone; the zero-density theorem they use is proved in Appendix A. The bound for every character needs a zero-free half-plane for all Dirichlet L -functions, and one is now known. OpenAI has proved that every finite-order Hecke L -function over $\mathbb{Q}(\sqrt{-3})$ has no zero with $\operatorname{Re} s > 11/12$ [24]; by base change, $L_K(s, \chi \circ N) = L(s, \chi)L(s, \chi\chi_{-3})$ for a Dirichlet character χ , so no Dirichlet L -function has a zero there either. The proof works with the Möbius function of the Eisenstein integers: a zero to the right of the line $\operatorname{Re} s = 11/12$ would prevent a power saving in the smoothed Möbius sums of the field, and that power saving is obtained from a mean square over a family of such sums twisted by sextic residue symbols, which is estimated by two applications of Poisson summation, the transformation of Kubota's cubic theta function at the cusps, Patterson's formula for its Fourier coefficients and the quadratic large sieve over number fields. Appendix B reproduces that proof in full, with credit to its authors, so that the paper is complete in itself. The same group had earlier proved the weaker half-plane $\operatorname{Re} s > 7/8$ with a proof written in Lean 4 and checked by the Lean kernel against the Mathlib library [23]; we compiled that proof from its source and ran the kernel check ourselves, and the build, its axiom report and a Lean proof of the two bounds below from that input are in the archive cited in the data availability section, as an independent check of the conclusion.

Theorem 1.5. *Let $\theta = 11/12$, and let C_θ and C_5 be the constants of Theorem 1.1 and of Proposition 7.1. Then $M(\chi) \leq C_\theta \sqrt{q} \log \log q$ for every primitive character χ modulo $q \geq 3$, and $|L'/L(1, \chi)| \leq C_5 \log \log q$ for every nonprincipal character χ modulo $q \geq 3$.*

Proof. By Theorem B.1, $L(s, \xi) \neq 0$ for $\operatorname{Re} s > 11/12$ for every nonprincipal Dirichlet character ξ . The first bound is then Corollary 1.3 with $\theta = 11/12$. For the second, $L(s, \chi) \neq 0$ in the rectangle of Proposition 7.1, which gives the bound. $\square$

The constants of Theorem 1.5 are absolute. Since the half-plane $\operatorname{Re} s > 7/8$ of [23] is wider, the same proof with $\theta = 7/8$ gives the two bounds with the smaller constants $C_{7/8}$ and C_5 ; the statement with $\theta = 11/12$ is the one whose proof is in this paper. Under the generalised Riemann hypothesis the bound $|L'/L(1, \chi)| \ll \log \log q$ holds for every nonprincipal character [18]. The constants C_ε of Theorems 1.2 and 1.4 tend to infinity as ε

tends to zero, since $\theta = 1 - \varepsilon/48$ enters the proof through the width $1 - \theta$ of the zero-free rectangle; in the proof they grow like a power of $1/\varepsilon$. The exceptional characters are those whose L -function, or one of whose twists by a character of small modulus, has a zero near the line $\operatorname{Re} s = 1$ at a height at most $5q^2$; the theorems say nothing about those characters, and the Euler–Kronecker constant of $\mathbb{Q}(\zeta_q)$, which is a sum over all the characters modulo q , is not controlled by them.

The proof. Figure 1 shows how the parts fit together. Lemma 5.1 reduces $M(\chi)$ to the sums $\sum_{m \leq H} \bar{\chi}(m)e(m\beta)/m$ with $H \leq q/2$. On the major arcs, where β is close to a fraction with denominator at most R , the integers m with a large prime factor are written as pm' , the additive character $e(pmb/r)$ is expanded in the multiplicative characters ψ modulo r , and the prime sums of $\chi\psi$ are bounded with a power saving by a contour shift for the Riesz mean $\sum_{n \leq x} \chi\psi(n)\Lambda(n)(1 - n/x)$, Proposition 3.7. The contour is a rectangle of height x^2 and not a whole vertical strip, which is why the hypothesis (1.1) is a rectangle and not a half-plane: on the line $\operatorname{Re} s = 2$ the kernel $x^s/(s(s+1))$ decays like $|\Im s|^{-2}$, so the part of the integral above the height x^2 is bounded by a constant. On the minor arcs the sum is bounded by Theorem 4.1, a weak form of the theorem of Montgomery and Vaughan on exponential sums with multiplicative coefficients, whose proof needs only the spacing of the multiples of β and Selberg’s upper bound for prime pairs. The integers with only small prime factors are few enough to be counted trivially. Section 6 counts the characters for which (1.1) fails: a zero of $L(s, \chi\psi)$ is a zero of the L -function of a character modulo qr , the map $\chi \mapsto \chi\psi$ is injective for fixed ψ , and Theorem 2.3 bounds the zeros of all the characters modulo qr with $r \leq R$ together. Section 7 proves Theorem 1.4 by the same contour shift, now crossing the pole of $L'/L(s+1, \chi)/(s(s+1))$ at $s = 0$, whose residue is $L'/L(1, \chi)$. Appendix B proves the zero-free half-plane that Theorem 1.5 feeds into the transfer theorem; Figure 4 there shows its parts. The counting arguments are of the same kind as the counts attached to finite complexes and group actions in the authors’ earlier work [1, 2, 3].

The paper is organised as follows. Section 2 collects Chebyshev’s bounds, Selberg’s sieve and the zero-density theorem. Section 3 bounds Riesz means of prime sums by a contour shift in a rectangle. Section 4 proves the bound for exponential sums with multiplicative coefficients. Section 5 proves Theorem 1.1, Section 6 proves Theorem 1.2 and Corollary 1.3, and Section 7 proves Theorem 1.4.

2. PRELIMINARIES

Throughout, p and ℓ denote primes, Λ is the von Mangoldt function, $\theta(x) = \sum_{p \leq x} \log p$, $\psi(x) = \sum_{n \leq x} \Lambda(n)$, $e(x) = e^{2\pi i x}$, and $\|x\|$ is the distance from x to the nearest integer. We write $\omega(n)$ for the number of prime factors of n , $\Omega(n)$ for the number of prime factors counted with multiplicity, and $P(n)$ for the largest prime factor of $n \geq 2$, with $P(1) = 1$. The letter θ is also used for the abscissa of a zero-free rectangle; the two uses are never close to each other.

Lemma 2.1. *For $x \geq 1$ and integers $N \geq 2$,*

- (i) $\theta(x) \leq x \log 4$;
- (ii) $\psi(x) - \theta(x) \leq 2\sqrt{x} \log x$ and $\psi(x) \leq 4x$;
- (iii) $\sum_{p \leq N} \frac{\log p}{p} \leq \log N + \log 4$;
- (iv) $\sum_{p \leq N} \frac{1}{p} \leq \log \log N + 10$;
- (v) $\sum_{n \leq x} \frac{\Lambda(n)}{n} \leq \log x + 3$.

Proof. (i) It suffices to treat integers $m \geq 1$, and we use induction on m . The cases $m = 1, 2$ are clear, and for even $m > 2$ we have $\theta(m) = \theta(m-1)$. Let $m = 2n+1$. The primes in $(n+1, 2n+1]$ divide $\binom{2n+1}{n}$, and $2\binom{2n+1}{n} \leq 2^{2n+1}$ because $\binom{2n+1}{n}$ and $\binom{2n+1}{n+1}$ are equal terms of the binomial expansion of $(1+1)^{2n+1}$. Hence $\theta(2n+1) \leq \theta(n+1) + n \log 4 \leq (2n+1) \log 4$.

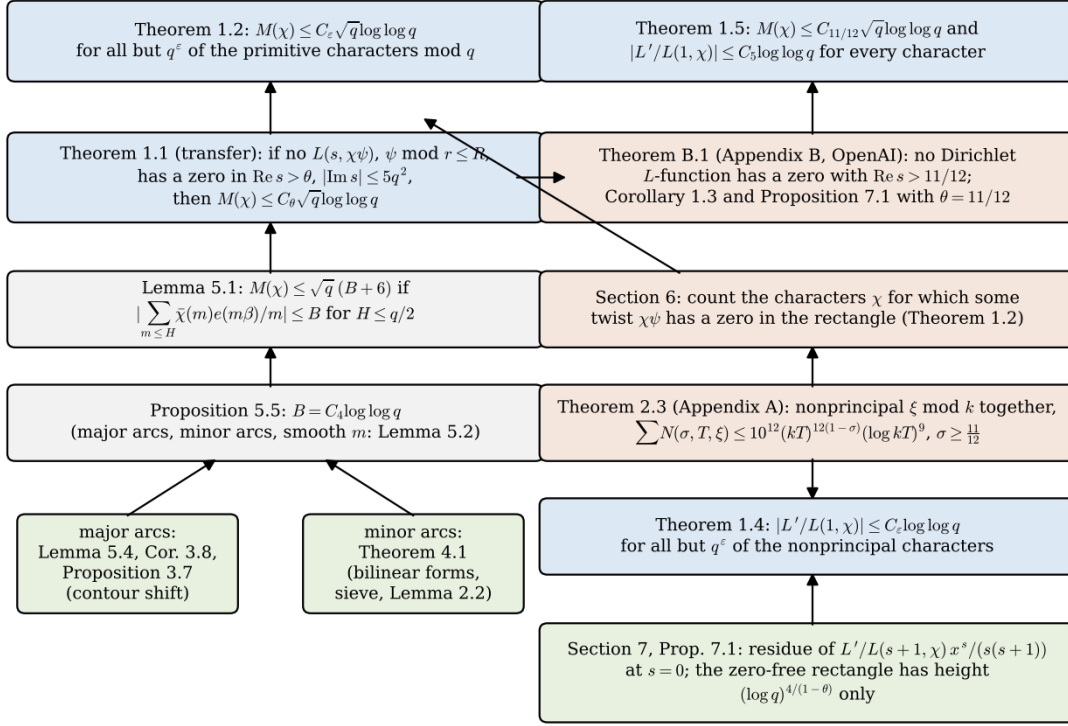


FIGURE 1. The structure of the proofs. The character sum is reduced to the sums $\sum_{m \leq H} \bar{\chi}(m)e(m\beta)/m$, which are bounded on the major arcs by the prime sums of the twists $\chi\psi$ (Proposition 3.7, Corollary 3.8, Lemma 5.4) and on the minor arcs by Theorem 4.1. The zero-density theorem bounds the number of characters for which the major-arc input fails.

(ii) We have $\psi(x) - \theta(x) = \sum_{k \geq 2} \theta(x^{1/k})$, where only $k \leq \log x / \log 2$ contribute. Each term is at most $\sqrt{x} \log 4$ by (i), so the sum is at most $(\log x / \log 2) \sqrt{x} \log 4 = 2\sqrt{x} \log x$. Since $\log x \leq \sqrt{x}$, $\psi(x) \leq (\log 4 + 2)x \leq 4x$.

(iii) By Legendre's formula and (i),

$$N \log N \geq \log N! \geq \sum_{p \leq N} \left\lfloor \frac{N}{p} \right\rfloor \log p \geq N \sum_{p \leq N} \frac{\log p}{p} - N \log 4.$$

(iv) Let $A(u) = \sum_{p \leq u} \log p / p$. Partial summation and (iii) give

$$\sum_{p \leq N} \frac{1}{p} = \frac{A(N)}{\log N} + \int_2^N \frac{A(u)}{u \log^2 u} du \leq 1 + \frac{\log 4}{\log 2} + \int_2^N \left(\frac{1}{u \log u} + \frac{\log 4}{u \log^2 u} \right) du,$$

and the integral is at most $\log \log N - \log \log 2 + 2$.

(v) The primes contribute at most $\log x + \log 4$ by (iii) applied to $N = \lfloor x \rfloor$ (for $x < 2$ the sum is empty). The prime powers p^k with $k \geq 2$ contribute at most $\sum_p \log p \sum_{k \geq 2} p^{-k} = \sum_p \log p / (p(p-1))$. The primes $p \leq 7$ give $\log 2/2 + \log 3/6 + \log 5/20 + \log 7/42 < 0.66$, and the primes $p \geq 11$ give at most $\sum_{p \geq 11} 2\sqrt{p}/(p(p-1)) \leq 2 \sum_{n \geq 11} n^{-3/2} \leq 2 \int_{10}^\infty u^{-3/2} du < 0.64$, since $\log p \leq 2\sqrt{p}$. The total is at most $\log x + \log 4 + 1.3 < \log x + 3$. $\square$

We use Selberg's upper bound sieve in the following form; for the method see [11, Chapter 3], and we give the proof of the form we need.³

³The weights λ_d are those of Selberg's quadratic form, and the only feature used later is the bound $|\lambda_d| \leq 1$, which makes the remainder term at most a sum of $3^{\omega(d)} |R_d|$ over $d \leq y$.

Lemma 2.2. *Let $a_1, \dots, a_X$ be integers, P a squarefree integer, ν a multiplicative function with $0 < \nu(p) < 1$ for $p \mid P$, and $y \geq 1$. For $d \mid P$ put $R_d = \#\{n \leq X : d \mid a_n\} - \nu(d)X$, and for $l \mid P$ put $g(l) = \prod_{p \mid l} \nu(p)/(1 - \nu(p))$. Then*

$$\#\{n \leq X : (a_n, P) = 1\} \leq \frac{X}{S} + \sum_{\substack{d \mid P \\ d \leq y}} 3^{\omega(d)} |R_d|, \quad S = \sum_{\substack{l \mid P \\ l \leq \sqrt{y}}} g(l).$$

Proof. Let D be the set of divisors d of P with $d \leq \sqrt{y}$, and let $(\lambda_d)_{d \in D}$ be real numbers with $\lambda_1 = 1$. If $(a_n, P) = 1$ then $\sum_{d \in D, d \mid a_n} \lambda_d = 1$, and otherwise this sum is real, so

$$\begin{aligned} \#\{n : (a_n, P) = 1\} &\leq \sum_{n \leq X} \left(\sum_{\substack{d \in D \\ d \mid a_n}} \lambda_d \right)^2 \\ &= X \sum_{d_1, d_2 \in D} \lambda_{d_1} \lambda_{d_2} \nu([d_1, d_2]) + \sum_{d_1, d_2 \in D} \lambda_{d_1} \lambda_{d_2} R_{[d_1, d_2]}. \end{aligned}$$

Write Q for the double sum in the first term. Since ν is multiplicative and d_1, d_2 are squarefree, $\nu([d_1, d_2]) = \nu(d_1)\nu(d_2)/\nu((d_1, d_2))$. Let h be the multiplicative function on the divisors of P with $h(p) = 1/\nu(p) - 1 = 1/g(p)$; then $1/\nu(m) = \sum_{l \mid m} h(l)$ for $m \mid P$, and

$$Q = \sum_{l \mid P} h(l) \left(\sum_{\substack{d \in D \\ l \mid d}} \lambda_d \nu(d) \right)^2 = \sum_{l \in D} \frac{y_l^2}{g(l)}, \quad y_l = \sum_{\substack{d \in D \\ l \mid d}} \lambda_d \nu(d).$$

Möbius inversion gives $\lambda_d \nu(d) = \sum_{l \in D, d \mid l} \mu(l/d) y_l$, so the condition $\lambda_1 = 1$ reads $\sum_{l \in D} \mu(l) y_l = 1$. By Cauchy's inequality

$$1 \leq \left(\sum_{l \in D} \frac{y_l^2}{g(l)} \right) \left(\sum_{l \in D} g(l) \right) = QS,$$

with equality for $y_l = \mu(l)g(l)/S$. We make this choice, so that $Q = 1/S$. Then, writing $l = de$,

$$\lambda_d = \frac{1}{\nu(d)S} \sum_{\substack{l \in D \\ d \mid l}} \mu(l/d) \mu(l) g(l) = \frac{\mu(d)g(d)}{\nu(d)S} \sum_{\substack{e \in D, (e,d)=1 \\ de \leq \sqrt{y}}} g(e).$$

Since $g(d)/\nu(d) = \prod_{p \mid d} (1 - \nu(p))^{-1} = \prod_{p \mid d} (1 + g(p)) = \sum_{f \mid d} g(f)$, we get $|\lambda_d| \leq S^{-1} \sum_{f \mid d} \sum_e g(fe)$, and the products fe are distinct elements of D ; hence $|\lambda_d| \leq 1$. Finally $[d_1, d_2] \leq d_1 d_2 \leq y$, and every $d \mid P$ is the least common multiple of at most $3^{\omega(d)}$ pairs (d_1, d_2) , so the second term is at most $\sum_{d \mid P, d \leq y} 3^{\omega(d)} |R_d|$. $\square$

For a Dirichlet character ξ modulo k , real σ and $T \geq 2$ let

$$N(\sigma, T, \xi) = \#\{\rho = \beta + i\gamma : L(\rho, \xi) = 0, \beta \geq \sigma, |\gamma| \leq T\}.$$

For $\sigma > 0$ the zeros counted are those of the L -function of the primitive character inducing ξ , since the Euler factors removed in passing from that character to ξ vanish only on the line $\operatorname{Re} s = 0$. The zero-density theorem we need is the following; it is proved in Appendix A, with the constant written out.

Theorem 2.3. *Let $c_0 = 10^{12}$. For every $k \geq 3$, every $T \geq 2$ and every σ with $11/12 \leq \sigma \leq 1$,*

$$\sum_{\substack{\xi \bmod k \\ \xi \neq \xi_0}} N(\sigma, T, \xi) \leq c_0 (kT)^{12(1-\sigma)} (\log kT)^9, \quad (2.1)$$

the sum running over the nonprincipal characters modulo k .

Theorems 12.1 and 12.2 of [21] give the sum over all characters modulo k , principal character included, with the bounds $\ll (kT)^{3(1-\sigma)/(2-\sigma)}(\log kT)^9$ for $1/2 \leq \sigma \leq 4/5$ and $\ll (kT)^{2(1-\sigma)/\sigma}(\log kT)^{14}$ for $4/5 \leq \sigma \leq 1$, with unspecified absolute constants; these are stronger than (2.1) in the exponent of kT . We prove the weaker form (2.1) because its constant is explicit and because the proof, which follows Ingham [16] and Montgomery, is short enough to be given in full.⁴ In the proof of Theorem 1.2 only the shape of (2.1) matters, namely that the exponent of kT tends to zero with $1 - \sigma$.

3. RIESZ MEANS OF PRIME SUMS

For an arithmetic function a and $x \geq 1$ put

$$S(a; x) = \sum_{n \leq x} a(n) \left(1 - \frac{n}{x}\right), \quad S(\xi; x) = S(\xi \Lambda; x) = \sum_{n \leq x} \xi(n) \Lambda(n) \left(1 - \frac{n}{x}\right).$$

The weight $1 - n/x$ makes the Mellin kernel absolutely integrable, which keeps the contour shift below free of truncation errors. The aim of this section is Proposition 3.7, which bounds $S(\xi; x)$ with a power saving when $L(s, \xi)$ has no zero in a rectangle of height $x^2 + 2$ to the left of $\operatorname{Re} s = 1$.

Lemma 3.1. *Let $a: \mathbb{N} \rightarrow \mathbb{C}$ satisfy $\sum_n |a(n)| n^{-\sigma} < \infty$ for $\sigma > 1$, and let $F(s) = \sum_n a(n) n^{-s}$ for $\operatorname{Re} s > 1$.*

- (i) *For $\operatorname{Re} s > 1$, $\int_1^\infty S(a; x) x^{-s-1} dx = \frac{F(s)}{s(s+1)}$.*
- (ii) *For $x > 0$, $S(a; x) = \frac{1}{2\pi} \int_{-\infty}^\infty F(2+it) \frac{x^{2+it}}{(2+it)(3+it)} dt$.*

Proof. (i) For $n \geq 1$, $\int_n^\infty (1 - n/x) x^{-s-1} dx = n^{-s}/s - n^{-s}/(s+1) = n^{-s}/(s(s+1))$. Since $\sum_n |a(n)| \int_n^\infty x^{-\sigma-1} dx = \sigma^{-1} \sum_n |a(n)| n^{-\sigma} < \infty$, the sum and the integral may be exchanged.

(ii) For $y > 0$ let $k(y) = \frac{1}{2\pi i} \int_{(2)} y^s \frac{ds}{s(s+1)}$. If $y \geq 1$, moving the line of integration to $\operatorname{Re} s = -U$, with $U > 2$, picks up the residues 1 at $s = 0$ and $-1/y$ at $s = -1$. For each fixed U , the horizontal integrals tend to zero as their height tends to infinity: their lengths are fixed and the denominator is of order the square of that height. On the new vertical line,

$$\int_{-\infty}^\infty \frac{|y^{-U+it}|}{|(-U+it)(1-U+it)|} dt \leq y^{-U} \int_{-\infty}^\infty \frac{dt}{(U-1)^2 + t^2} = \frac{\pi y^{-U}}{U-1}.$$

This tends to zero for $y \geq 1$, including $y = 1$, so $k(y) = 1 - 1/y$ there. For $y < 1$, shifting instead to $\operatorname{Re} s = U$ crosses no pole and leaves an integral bounded by $\pi y^U/U$, which tends to zero. As $|s(s+1)|^{-1}$ is integrable on $\operatorname{Re} s = 2$ and $\sum_n |a(n)| n^{-2} < \infty$, summation and integration may be exchanged, and the right side equals $\sum_n a(n) k(x/n) = S(a; x)$. $\square$

Lemma 3.2. *Let χ be nonprincipal modulo q . For $\operatorname{Re} s = \sigma > 0$, $|L(s, \chi)| \leq q|s|/\sigma$.*

Proof. Put $A(t) = \sum_{n \leq t} \chi(n)$; then $|A(t)| \leq q$, since the sum of χ over a period vanishes, and partial summation gives $L(s, \chi) = s \int_1^\infty A(t) t^{-s-1} dt$ for $\sigma > 0$. Hence $|L(s, \chi)| \leq q|s| \int_1^\infty t^{-\sigma-1} dt = q|s|/\sigma$. $\square$

Lemma 3.3. *Let $A_0 = \frac{3}{2} \sum_\ell \ell^{-3/2}$. For every Dirichlet character χ the series $E(s) = -\sum_\ell \log(1 - \chi(\ell) \ell^{-s})$ is holomorphic on $\operatorname{Re} s > 1$, $e^{E(s)} = L(s, \chi)$ there, and $|E(s)| \leq A_0$ for $\operatorname{Re} s \geq 3/2$. Moreover $|L'/L(s, \chi)| \leq 2$ for $\operatorname{Re} s \geq 2$.*

Proof. For $\operatorname{Re} s \geq 1$ we have $|\chi(\ell) \ell^{-s}| \leq 1/2$, and $|\log(1 - w)| \leq \sum_k |w|^k/k \leq \frac{3}{2}|w|$ for $|w| \leq 1/2$: indeed, the terms with $k \geq 2$ have total at most $|w|^2/(2(1 - |w|)) \leq |w|/2$. The series therefore converges locally uniformly on $\operatorname{Re} s > 1$ and is bounded by A_0 on $\operatorname{Re} s \geq 3/2$,

⁴Whether the zeros are counted in $0 < \gamma \leq T$ or in $|\gamma| \leq T$ affects only the constant, since the zeros of $L(s, \bar{\xi})$ are the conjugates of the zeros of $L(s, \xi)$ and the sum is over all characters modulo k .

and the Euler product gives $e^E = L$. For $\operatorname{Re} s \geq 2$ the derivative $E'(s) = -\sum_n \chi(n)\Lambda(n)n^{-s}$ has modulus at most $\sum_{n \geq 2} \log n/n^2$, and since $\log u/u^2$ decreases for $u \geq 2$ this is at most $\log 2/4 + \int_2^\infty \log u u^{-2} du = \log 2/4 + (1 + \log 2)/2 < 2$. $\square$

Lemma 3.4 (Borel–Carathéodory). *Let h be holomorphic on a neighbourhood of the disc $|z - c| \leq R$ and let $0 < r < R$. Then*

$$\max_{|z-c| \leq r} |h(z) - h(c)| \leq \frac{2r}{R-r} \max_{|w-c|=R} \operatorname{Re}(h(w) - h(c)).$$

Proof. Put $g(z) = h(c+z) - h(c)$ and $A = \max_{|w|=R} \operatorname{Re} g(w)$. The mean of $\operatorname{Re} g$ over the circle $|w| = R$ is $\operatorname{Re} g(0) = 0$, so $A \geq 0$. If $A = 0$, then $\operatorname{Re} g$ vanishes on the circle, hence on the disc, and g is an imaginary constant equal to $g(0) = 0$. Let $A > 0$. By the maximum principle $\operatorname{Re} g \leq A$ on $|z| \leq R$, so $|2A - g| \geq A$ there and $\phi = g/(2A - g)$ is holomorphic. Since $|g|^2 - |2A - g|^2 = 4A(\operatorname{Re} g - A) \leq 0$, we have $|\phi| \leq 1$, and $\phi(0) = 0$. Schwarz's lemma gives $|\phi(z)| \leq |z|/R$, and from $g = 2A\phi/(1 + \phi)$ we get $|g(z)| \leq 2A|z|/(R - |z|)$. $\square$

Lemma 3.5. *Let $0 < \delta < 1/1000$, $1/2 \leq a \leq 1$, $t \in \mathbb{R}$, and let χ be nonprincipal modulo q with $L(s, \chi) \neq 0$ for $\operatorname{Re} s > a + 2\delta$, $|\Im s - t| \leq 2$. Put $K_1 = (1 + 2A_0)/(2\delta^2)$. For $a + 8\delta \leq \sigma \leq 2$,*

$$\left| \frac{L'}{L}(\sigma + it, \chi) \right| \leq K_1 \log(2q(3 + |t|)^2).$$

Proof. Let $c = 2 + it$, $R = 2 - a - 3\delta$, $r = 2 - a - 6\delta$ and $\Lambda_0 = \log(2q(3 + |t|)^2) \geq \log 18 > 1$. The disc $|z - c| < 2 - a - 2\delta$ has radius at most $3/2$, so it lies in the strip $|\Im z - t| \leq 2$, and on it $\operatorname{Re} z > a + 2\delta$; hence $L(z, \chi)$ has no zero there and has a holomorphic logarithm h . On $|z - c| < 1/2$ both h and the function E of Lemma 3.3 are logarithms of L , so they differ by a constant in $2\pi i\mathbb{Z}$; we choose $h = E$ there. Then $|h(c)| \leq A_0$.

If $|z - c| \leq R$, then $\operatorname{Re} z \geq a + 3\delta \geq 1/2$ and $|z| \leq |c| + 3/2 \leq 4 + |t|$, so Lemma 3.2 gives $\operatorname{Re} h(z) = \log |L(z, \chi)| \leq \log(2q(4 + |t|)) \leq \Lambda_0$. Lemma 3.4 gives, for $|z - c| \leq r$,

$$|h(z) - h(c)| \leq \frac{2r}{R-r} (\Lambda_0 + A_0) \leq \frac{\Lambda_0 + A_0}{\delta},$$

as $r \leq 3/2$ and $R - r = 3\delta$. Hence $|h(z)| \leq A_0 + (\Lambda_0 + A_0)/\delta \leq (1 + 2A_0)\Lambda_0/\delta$ on $|z - c| \leq r$. For $w = \sigma + it$ with $a + 8\delta \leq \sigma \leq 2$ we have $|w - c| = 2 - \sigma \leq r - 2\delta$, so the circle $|z - w| = 2\delta$ lies in $|z - c| \leq r$, and Cauchy's estimate gives $|L'/L(w, \chi)| = |h'(w)| \leq (1 + 2A_0)\Lambda_0/(2\delta^2)$. $\square$

Lemma 3.6. *For $q \geq 1$, $t \in \mathbb{R}$ and $s = \sigma + it$ with $\sigma \geq 1/2$,*

$$\log(2q(3 + |t|)^2) \leq 10 \log(q + 2)(1 + |t|)^{1/2}, \quad |s(s + 1)| \geq \frac{1}{8}(1 + |t|)^2.$$

Proof. We have $\log(2q) \leq 2 \log(q + 2)$, and $\log y \leq 2y^{1/2}$ with $3 + |t| \leq 4(1 + |t|)$ gives $\log(3 + |t|) \leq 4(1 + |t|)^{1/2}$; the first bound follows from $\log(q + 2) \geq 1$. For the second, $|s| \geq \max(\sigma, |t|) \geq (1 + |t|)/4$ and $|s + 1| \geq \max(\sigma + 1, |t|) \geq (1 + |t|)/2$. $\square$

Proposition 3.7. *Let $0 < \delta < 1/1000$, $1/2 \leq a \leq 1$, and put $K = 90K_1 + 6$ with K_1 as in Lemma 3.5. Let $x \geq 1$ and let ξ be nonprincipal modulo k with $L(s, \xi) \neq 0$ for $\operatorname{Re} s > a + 2\delta$, $|\Im s| \leq x^2 + 2$. Then*

$$|S(\xi; x)| \leq K x^{a+8\delta} \log(k + 2).$$

Proof. Put $\sigma_0 = a + 8\delta$, $T = x^2$ and

$$\Phi(s) = -\frac{L'}{L}(s, \xi) \frac{x^s}{s(s + 1)}.$$

For $\operatorname{Re} s > 1$, $-L'/L(s, \xi) = \sum_n \xi(n)\Lambda(n)n^{-s}$, so Lemma 3.1(ii) gives $2\pi S(\xi; x) = \int_{-\infty}^\infty \Phi(2 + it) dt$. We split the integral at $|t| = T$.

For $|t| > T$ we use the line $\operatorname{Re} s = 2$ as it stands: $|L'/L(2 + it, \xi)| \leq 2$ by Lemma 3.3, $|x^{2+it}| = x^2$ and $|s(s + 1)| \geq (1 + |t|)^2/8$ by Lemma 3.6, so

$$\int_{|t| > T} |\Phi(2 + it)| dt \leq 2 \int_T^\infty \frac{16x^2}{(1 + t)^2} dt = \frac{32x^2}{1 + T} \leq 32.$$

For $|t| \leq T$ we move the line of integration to $\operatorname{Re} s = \sigma_0$ inside the rectangle $\sigma_0 \leq \operatorname{Re} s \leq 2$, $|\Im s| \leq T$ (Figure 2). The function Φ is holomorphic on a neighbourhood of this rectangle, since $L(s, \xi)$ has no zero with $\operatorname{Re} s \geq \sigma_0 > a + 2\delta$ and $|\Im s| \leq T + 2$, and $s(s+1) \neq 0$ there. By Cauchy's theorem the integral of Φ over the right side equals the integral over the left side plus the integrals over the two horizontal sides, with the appropriate orientations. On the rectangle, Lemma 3.5 applies at every point $\sigma + it$, because the hypothesis gives $L(s, \xi) \neq 0$ for $\operatorname{Re} s > a + 2\delta$ and $|\Im s - t| \leq 2$ when $|t| \leq T$, and with Lemma 3.6 it gives

$$|\Phi(\sigma + it)| \leq 80K_1 \log(k+2) x^\sigma (1+|t|)^{-3/2} \quad (\sigma_0 \leq \sigma \leq 2, |t| \leq T).$$

Since $\int_{\mathbb{R}} (1+|t|)^{-3/2} dt = 4$, the left side contributes at most $320K_1 \log(k+2) x^{\sigma_0}$. Each horizontal side has length $2 - \sigma_0 \leq 3/2$ and lies at height T , so it contributes at most

$$\frac{3}{2} \cdot 80K_1 \log(k+2) x^2 (1+T)^{-3/2} \leq 120K_1 \log(k+2) x^{-1} \leq 120K_1 \log(k+2) x^{\sigma_0},$$

because $x^2(1+x^2)^{-3/2} \leq x^{-1}$. Adding the three contributions and the tail, and using $\log(k+2) \geq 1$ and $x^{\sigma_0} \geq 1$,

$$2\pi|S(\xi; x)| \leq (560K_1 + 32) x^{\sigma_0} \log(k+2),$$

and $(560K_1 + 32)/(2\pi) \leq 90K_1 + 6$. $\square$

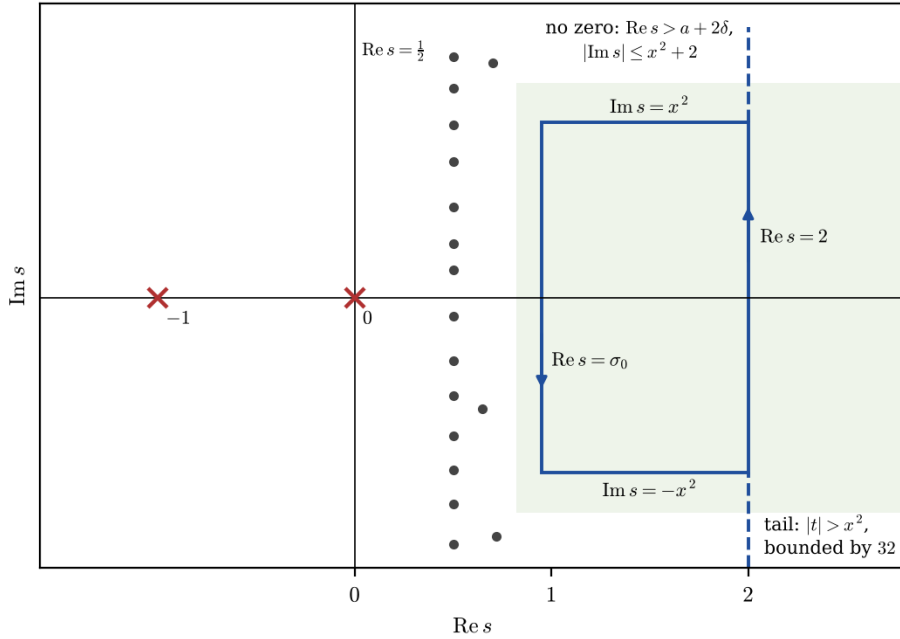


FIGURE 2. The contour shift in the proof of Proposition 3.7. The integral over the line $\operatorname{Re} s = 2$ is cut at the height $T = x^2$; the part beyond is bounded by the decay of the kernel, and the part within is moved to the line $\operatorname{Re} s = \sigma_0$ through the rectangle, which lies inside the region $\operatorname{Re} s > a + 2\delta$, $|\Im s| \leq x^2 + 2$ where $L(s, \xi)$ has no zero (dots: zeros outside that region; crosses: the poles of the kernel at 0 and -1). Horizontal spacing is schematic.

Parameters. For $1/2 \leq \theta < 1$ we fix once and for all

$$\delta_\theta = \min\left(\frac{1}{2000}, \frac{1-\theta}{100}\right), \quad a_\theta = \max\left(\frac{1}{2}, \theta - 2\delta_\theta\right), \quad \sigma_\theta = a_\theta + 8\delta_\theta, \quad \eta_\theta = \frac{1-\sigma_\theta}{4}, \quad (3.1)$$

and we write K_θ for the constant K of Proposition 3.7 with $\delta = \delta_\theta$. Then $0 < \delta_\theta < 1/1000$, $1/2 \leq a_\theta \leq 1$ and $a_\theta + 2\delta_\theta \geq \theta$, so that a character with no zero in $\operatorname{Re} s > \theta$, $|\Im s| \leq x^2 + 2$ satisfies the hypothesis of Proposition 3.7. Moreover

$$\theta < \sigma_\theta < 1, \quad 1 - \sigma_\theta \geq \frac{1}{2}(1 - \theta), \quad \frac{1}{8}(1 - \theta) \leq \eta_\theta \leq \frac{1}{8}. \quad (3.2)$$

Indeed, if $a_\theta = \theta - 2\delta_\theta$ then $\sigma_\theta = \theta + 6\delta_\theta \leq \theta + \frac{6}{100}(1 - \theta)$, and if $a_\theta = 1/2$ then $\theta < 1/2 + 2\delta_\theta$ and $\sigma_\theta = 1/2 + 8\delta_\theta \leq 1/2 + 8/2000$, while $1 - \theta \leq 1/2$; in both cases $\theta < \sigma_\theta$ and $1 - \sigma_\theta \geq \frac{94}{100}(1 - \theta)$. The upper bound $\eta_\theta \leq 1/8$ follows from $\sigma_\theta \geq 1/2$.

Corollary 3.8. *Let $1/2 \leq \theta < 1$, with the parameters (3.1), and put $C_3 = 8K_\theta + 8/\eta_\theta$. Let $x \geq 1$ and let ξ be nonprincipal modulo k with $L(s, \xi) \neq 0$ for $\operatorname{Re} s > \theta$, $|\Im s| \leq 4x^2 + 2$. Then:*

- (i) $|S(\xi; y)| \leq K_\theta y^{\sigma_\theta} \log(k+2)$ for $1 \leq y \leq 2x$;
- (ii) $|\sum_{p \leq x} \xi(p) \log p| \leq C_3 x^{1-\eta_\theta} \log(k+2)$;
- (iii) for integers $2 \leq A \leq B \leq x$ and $\gamma \in \mathbb{R}$,

$$\left| \sum_{A < p \leq B} \xi(p) e(p\gamma) \right| \leq 2C_3 \frac{B^{1-\eta_\theta} \log(k+2)(2 + 2\pi|\gamma|B)}{\log A}. \quad (3.3)$$

Proof. Write $\sigma = \sigma_\theta$, $\eta = \eta_\theta$ and $K = K_\theta$, so that $1 + \sigma = 2 - 4\eta$. (i) For $y \leq 2x$ we have $y^2 + 2 \leq 4x^2 + 2$, so Proposition 3.7 applies to y .

(ii) Put $T(y) = yS(\xi; y) = \sum_{n \leq y} \xi(n) \Lambda(n)(y - n)$ and $h = x^{1-2\eta}$, so that $1 \leq h \leq x$. Then

$$T(x+h) - T(x) = h \sum_{n \leq x} \xi(n) \Lambda(n) + \sum_{x < n \leq x+h} \xi(n) \Lambda(n)(x+h-n),$$

and the last sum has at most $h+1$ terms, each of modulus at most $h \log(2x)$. With $|T(y)| \leq Ky^{1+\sigma} \log(k+2)$ for $y \leq 2x$, from (i), this gives

$$\left| \sum_{n \leq x} \xi(n) \Lambda(n) \right| \leq \frac{2K(2x)^{1+\sigma} \log(k+2)}{h} + (h+1) \log(2x).$$

Here $(2x)^{1+\sigma}/h \leq 4x^{2-4\eta-1+2\eta} = 4x^{1-2\eta}$, so the first term is at most $8Kx^{1-\eta} \log(k+2)$. Since $\eta \leq 1/8$ we have $1 \leq x^\eta/\eta$, and $\log x = \eta^{-1} \log x^\eta \leq x^\eta/\eta$, so $\log(2x) \leq 1 + \log x \leq 2x^\eta/\eta$, and the second term is at most $2x^{1-2\eta} \cdot 2x^\eta/\eta = 4x^{1-\eta}/\eta$. The prime powers p^j with $j \geq 2$ contribute at most $\psi(x) - \theta(x) \leq 2\sqrt{x} \log x \leq 2x^{1/2+\eta}/\eta \leq 2x^{1-\eta}/\eta$ by Lemma 2.1(ii), since $\eta \leq 1/4$. As $\log(k+2) \geq 1$, (ii) follows with $C_3 = 8K + 8/\eta$.

(iii) Let $\Pi(u) = \sum_{p \leq u} \xi(p) \log p$ and $f(u) = e(u\gamma)/\log u$. Partial summation gives

$$\sum_{A < p \leq B} \xi(p) e(p\gamma) = f(B)\Pi(B) - f(A)\Pi(A) - \int_A^B \Pi(u) f'(u) du,$$

and $|f'(u)| \leq 2\pi|\gamma|/\log u + 1/(u \log^2 u)$ for $u \geq A$. By (ii), applied at every $u \leq x$, the right side is at most $C_3 B^{1-\eta} \log(k+2)(3 + 2\pi|\gamma|B)/\log A$, which gives (3.3). $\square$

4. EXPONENTIAL SUMS WITH MULTIPLICATIVE COEFFICIENTS

For $M > 0$ put $w_M(\theta) = M$ if $\|\theta\| < 1/M$ and $w_M(\theta) = 1/\|\theta\|$ otherwise; then $0 \leq w_M(\theta) \leq M$ and $w_M(-\theta) = w_M(\theta)$.

Theorem 4.1. *There is an absolute constant C_1 with the following property. Let f be completely multiplicative with $|f(n)| \leq 1$, let $\alpha \in \mathbb{R}$, and let a, r be coprime integers with $r \geq 1$, $|\alpha - a/r| \leq r^{-2}$ and $2 \leq R \leq r \leq N/R$. Then*

$$\left| \sum_{n \leq N} f(n) e(n\alpha) \right| \leq C_1 \left(\frac{N}{\log N} + \frac{N(\log N)^2}{R^{1/4}} \right).$$

Montgomery and Vaughan [22, Corollary 1] proved this for multiplicative f with $N(\log R)^{3/2}R^{-1/2}$ in place of $N(\log N)^2R^{-1/4}$. The proof below follows their approach, a bilinear decomposition together with the spacing of the multiples of α and an upper bound for prime pairs, with cruder estimates, which suffice for Theorem 1.1.

Lemma 4.2. *Let a, r be coprime with $r \geq 2$ and $|\alpha - a/r| \leq r^{-2}$, and let $H \geq 0$ be an integer.*

- (i) $\|k\alpha\| \geq 1/(2r)$ for every integer k with $0 < |k| \leq r/2$.
- (ii) $\#\{1 \leq h \leq H : \|h\alpha\| < t\} \leq (4H/r + 1)(4rt + 1)$ for every $t \geq 0$.
- (iii) For $M > 0$ and every integer $K \geq 0$ with $M \leq 2^K$,

$$\sum_{h \leq H} w_M(h\alpha) \leq \left(\frac{4H}{r} + 1\right)(3M + 4r + 8rK),$$

$$\sum_{h \leq H} w_M(h\alpha)^2 \leq \left(\frac{4H}{r} + 1\right)(20rM + 3M^2).$$

Proof. (i) Let z be the integer nearest to $k\alpha$. Since $0 < |k| < r$ and $(a, r) = 1$, $r \nmid ka$, so $|ka/r - z| \geq 1/r$ and $\|k\alpha\| \geq 1/r - |k| |\alpha - a/r| \geq 1/(2r)$.

(ii) Let $L = \lfloor r/2 \rfloor$, so $L \geq r/4$. For k in a block of L consecutive integers let $y_k = k\alpha - z_k$ with z_k the integer nearest to $k\alpha$. For $k \neq k'$ in the block, $0 < |k - k'| < L \leq r/2$ and $y_k - y_{k'} \equiv (k - k')\alpha \pmod{1}$, so $|y_k - y_{k'}| \geq 1/(2r)$ by (i). Points of $(-t, t)$ at mutual distances at least $1/(2r)$ number at most $4rt + 1$, and $[1, H]$ is covered by at most $H/L + 1 \leq 4H/r + 1$ blocks.

(iii) If $\|\theta\| \geq 1/M$, let $k \geq 1$ be least with $\|\theta\| < 2^k/M$; then $k \leq K$ because $\|\theta\| \leq 1/2 < 2^K/M$, and $\|\theta\| \geq 2^{k-1}/M$, so $w_M(\theta) \leq 2^{1-k}M$. Hence

$$w_M(\theta)^j \leq M^j [\|\theta\| < 1/M] + \sum_{k=1}^K (2^{1-k}M)^j [\|\theta\| < 2^k/M] \quad (j = 1, 2).$$

Summing over $h \leq H$ with (ii), the first sum is at most $(4H/r + 1)(4r + M + \sum_{k \leq K} (8r + 2^{1-k}M))$ and the second at most $(4H/r + 1)(4rM + M^2 + \sum_{k \geq 1} (16rM2^{-k} + 4M^24^{-k}))$. $\square$

Lemma 4.3. *Let $\theta, \alpha \in \mathbb{R}$, $U > 0$, and let A, B, N be nonnegative integers with $B \leq A + U$.*

- (i) $|\sum_{A < x \leq B} e(x\theta)| \leq w_U(\theta)$.
- (ii) For a finite set V of positive integers and $b : V \rightarrow \mathbb{C}$,

$$\sum_{A < u \leq B} \left| \sum_{\substack{v \in V \\ uv \leq N}} b(v)e(uv\alpha) \right|^2 \leq \sum_{v, v' \in V} |b(v)| |b(v')| w_U((v - v')\alpha).$$

- (iii) For $M > 0$ and an integer $H \geq 0$, $\sum_{1 \leq m, m' \leq H} w_M((m - m')\alpha) \leq H(M + 2 \sum_{h \leq H} w_M(h\alpha))$.

Proof. (i) If $\|\theta\| < 1/U$, the sum has at most U terms of modulus 1. Otherwise $\theta \notin \mathbb{Z}$, the sum has modulus at most $|e((B - A)\theta) - 1|/|e(\theta) - 1|$, and $|e(\theta) - 1| = 2|\sin \pi\theta| \geq 4\|\theta\|$, so the modulus is at most $1/(2\|\theta\|)$. (ii) Expanding the square gives $\sum_{v, v'} b(v)\bar{b}(v') \sum_u e(u(v - v')\alpha)$, where u runs over the integers of $(A, \min(B, N/v, N/v'))$, an interval of length at most U ; apply (i). (iii) For fixed m , the term $m' = m$ is $w_M(0) = M$, and each $h = |m - m'| \in [1, H]$ arises from at most two m' . $\square$

Lemma 4.4. *Let $C_5 = 1 + 4\zeta(3/2)$. For every completely multiplicative f with $|f(n)| \leq 1$, every $\alpha \in \mathbb{R}$ and every integer $N \geq 1$, the sum $S = \sum_{n \leq N} f(n)e(n\alpha)$ satisfies*

$$\left| S \log N - \sum_{p \leq N} \log p f(p) \sum_{m \leq N/p} f(m)e(pm\alpha) \right| \leq C_5 N.$$

Proof. Write $S \log N = S_1 + S_2$ with

$$S_1 = \sum_{n \leq N} f(n) e(n\alpha) \log n, \quad S_2 = \sum_{n \leq N} f(n) e(n\alpha) \log(N/n).$$

Then $|S_2| \leq N \log N - \log N! \leq N$, since $N^N/N! \leq e^N$. In S_1 write $\log n = \sum_{d|n} \Lambda(d)$ and $n = dm$; complete multiplicativity gives

$$S_1 = \sum_{d \leq N} \Lambda(d) f(d) \sum_{m \leq N/d} f(m) e(dm\alpha).$$

The terms with d prime form the sum in the statement, and the others have modulus at most

$$N \sum_p \sum_{k \geq 2} \frac{\log p}{p^k} \leq 2N \sum_p \frac{\log p}{p^2} \leq 4\zeta(3/2)N,$$

since $\log p \leq 2\sqrt{p}$. □

Let $G(h) = \prod_{p|h} p/(p-1)$.

Lemma 4.5. (i) For every integer $X \geq 2$ and every even $h \geq 2$,

$$\#\{0 \leq n < X : n \text{ and } n+h \text{ prime}\} \leq \frac{400 G(h)^2 X}{(\log X)^2} + 3X^{3/5}.$$

(ii) $\sum_{h \leq X} G(h)^4 \leq e^{30} X$ for every $X \geq 1$.

Proof. (i) Apply Lemma 2.2 to $a_n = n(n+h)$, $0 \leq n < X$, with $y = X^{1/5}$, $z = \lfloor y \rfloor$, P the product of the primes $p \leq z$ with $p \nmid h$, and $\nu(d) = 2^{\omega(d)}/d$. Every $p \mid P$ is odd, so $\nu(p) = 2/p < 1$. For $d \mid P$ the residues n with $d \mid n(n+h)$ form $2^{\omega(d)}$ classes modulo d (namely $n \equiv 0$ or $-h \pmod{p}$ for each $p \mid d$, distinct as $p \nmid h$), and each class contains between $X/d - 1$ and $X/d + 1$ of the integers $0 \leq n < X$; hence $|R_d| \leq 2^{\omega(d)}$. Since the prime factors of d are at least 3, $3^{\omega(d)} |R_d| \leq 6^{\omega(d)} \leq d^2 \leq y^2$, and the error term is at most $y^3 = X^{3/5}$.

Let $\mathcal{M}$ be the set of $m \leq \sqrt{y}$ whose prime factors divide P . Since $\nu(p)/(1 - \nu(p)) = \sum_{k \geq 1} \nu(p)^k$ and $\nu(m) = 2^{\Omega(m)}/m$ for these m , $g(l)$ is the sum of $2^{\Omega(m)}/m$ over the m whose set of prime factors is that of l . An $m \in \mathcal{M}$ is counted in $g(l)$ for l the product of its prime factors, and $l \leq m \leq \sqrt{y}$. Hence, with $\tau(m) \leq 2^{\Omega(m)}$ the number of divisors of m ,

$$S \geq \sum_{m \in \mathcal{M}} \frac{2^{\Omega(m)}}{m} \geq \sum_{m \in \mathcal{M}} \frac{\tau(m)}{m}.$$

Let $T = \lfloor y^{1/4} \rfloor$ and let B be the set of $b \leq T$ that are prime to h ; these b are odd. For $b, b' \in B$ the product bb' lies in $\mathcal{M}$, and a given m is a product bb' in at most $\tau(m)$ ways, so $S \geq (\sum_{b \in B} 1/b)^2$. If E is a set of primes and $p \notin E$, every $b \leq T$ prime to E is uniquely $p^k b'$ with b' prime to $E \cup \{p\}$, so the sum of $1/b$ over the $b \leq T$ prime to E is at most $p/(p-1)$ times the same sum for $E \cup \{p\}$. Adding the prime factors of h one at a time gives

$$\sum_{b \in B} \frac{1}{b} \geq \frac{1}{G(h)} \sum_{b \leq T} \frac{1}{b} \geq \frac{\log(T+1)}{G(h)}.$$

As $T+1 > y^{1/4} = X^{1/20}$, we get $S \geq (\log X)^2/(400G(h)^2)$, that is, $X/S \leq 400G(h)^2 X/(\log X)^2$. Finally, if $n < X$ and both n and $n+h$ are prime, then $n \leq z$ or $(n(n+h), P) = 1$, so the count is at most $z+1 + X/S + y^3 \leq 400G(h)^2 X/(\log X)^2 + 3X^{3/5}$.

(ii) For $t = p - 1 \geq 1$, $(t + 1)^5 \leq (t + 31)t^4$, that is, $(p/(p - 1))^4 \leq 1 + 30/p$. Hence $G(h)^4 \leq \sum_d 30^{\omega(d)}/d$ over the squarefree $d \mid h$, and

$$\begin{aligned} \sum_{h \leq X} G(h)^4 &\leq \sum_{d \leq X} \mu(d)^2 \frac{30^{\omega(d)}}{d} \cdot \frac{X}{d} \leq X \prod_p \left(1 + \frac{30}{p^2}\right) \\ &\leq X \exp\left(30 \sum_{n \geq 2} \frac{1}{n^2}\right) \leq e^{30} X. \end{aligned} \quad \square$$

Proof of Theorem 4.1. Let S be the sum in the theorem, $L = \log N$ and $\ell = \lfloor \log_2 N \rfloor$. Since $N \geq rR \geq 4$, we have $L \geq 1$. If $\ell < 64$, then $L < 64 \log 2 < 45$ and $|S| \leq N \leq 45N/L$. Let now $\ell \geq 64$, and put $I = \lfloor \log_2 \ell \rfloor + 1$ and $J = \ell + 1 - 4I$. Then $16^I > \ell^4$, $1 \leq J \leq \ell \leq 2L$, $I \leq \ell \leq 2L$, $2^I \leq 2\ell$, $N < 2^{\ell+1} = 2^J 16^I$ and $J^2 2^J \leq (\ell + 1)^2 2^{\ell+1} \ell^{-4} \leq 2N$. By Lemma 4.4, $|S|L \leq C_5 N + |T_1| + |T_2|$, where T_1 and T_2 are the parts of $\sum_p \log p f(p) \sum_{m \leq N/p} f(m) e(pm\alpha)$ with $p \leq 2^J$ and with $p > 2^J$.

The small primes. Let $0 \leq j < J$, $U = 2^j$, $H = \lfloor N/U \rfloor$, and for $U < x \leq 2U$ let $B(x) = \sum_{m \leq H, xm \leq N} f(m) e(xm\alpha)$; the inner sum for a prime $p \in (U, 2U]$ is $B(p)$. By Cauchy's inequality, Lemma 4.3(ii), (iii) and Lemma 4.2(iii) with $M = U$, $K = j$,

$$\sum_{U < p \leq 2U} \log p |B(p)| \leq (U \log^2(2U))^{1/2} Q_j^{1/2},$$

where $Q_j = H(U + 2(4H/r + 1)(3U + 4r + 8rj))$. Since $HU \leq N$, $UQ_j \leq 7NU + (8 + 16j)Nr + 24N^2/r + (32 + 64j)N^2/U$, and $\log(2U) \leq j + 1$. Summing over $j < J$ with $j + 1 \leq J$, $\sum_{j < J} 2^{j/2} \leq 3 \cdot 2^{J/2}$, $\sqrt{(32 + 64j)2^{-j}} \leq 8(j + 1)(3/4)^j$ and $\sum_{j \geq 0} (j + 1)^2 (3/4)^j = 112$,

$$|T_1| \leq 3\sqrt{7}J\sqrt{N2^J} + J^2N\sqrt{24/r} + 896N + J^2\sqrt{24JNr}.$$

Now $J^2 2^J \leq 2N$ gives $3\sqrt{7}J\sqrt{N2^J} \leq 12N$; $r \geq R \geq \sqrt{R}$ gives $\sqrt{24/r} \leq 5R^{-1/4}$; and $r\sqrt{R} \leq rR \leq N$ gives $\sqrt{24JNr} \leq 5JNR^{-1/4}$. With $J \leq 2L$ and $L \geq 1$, $|T_1| \leq 908N + 60L^3NR^{-1/4}$.

The large primes. For $p > 2^J$ we have $N/p < 16^I$, so the m in T_2 lie in $(0, 16^I - 1]$, the union of the blocks $(16^i - 1, 16^{i+1} - 1]$ with $0 \leq i < I$. Fix i , let $A = 16^i - 1$, $U = 15 \cdot 16^i$ and $X = \lfloor N/16^i \rfloor$, and assume $X \geq 1$, the block being empty otherwise. Only the primes $p \in (2^J, X]$ contribute, since $p \geq X + 1$ and $m \geq 16^i$ give $pm > N$. Let V be the set of these primes, which are odd, and $B(m) = \sum_{p \in V, pm \leq N} \log p f(p) e(pm\alpha)$. The block is $\sum_{A < m \leq A+U} f(m) B(m)$, and by Cauchy's inequality, Lemma 4.3(ii) and $\log p \leq \lambda = \log(X + 1)$, its square is at most $U\lambda^2 W$, where

$$W = \sum_{p, p' \in V} w_U((p - p')\alpha) = |V|U + 2 \sum_{h=1}^X w_U(h\alpha) \#\{p \in V : p + h \in V\}.$$

The count vanishes for odd h , and for even h Lemma 4.5(i), applied to $X + 1$, bounds it by $400G(h)^2(X + 1)\lambda^{-2} + 3(X + 1)^{3/5}$. By Cauchy's inequality, Lemma 4.5(ii) and Lemma 4.2(iii), $\sum_{h \leq X} G(h)^2 w_U(h\alpha) \leq e^{15} \sqrt{X} Q^{1/2}$ with $Q = (4X/r + 1)(20rU + 3U^2)$, and $\sum_{h \leq X} w_U(h\alpha) \leq XU$. Hence $U\lambda^2 W \leq A_1 + 800e^{15}Z + A_3$, where

$$A_1 = U^2 \lambda^2 X, \quad Z = U(X + 1) \sqrt{X} Q^{1/2}, \quad A_3 = 6U^2 \lambda^2 (X + 1)^{3/5} X.$$

Let $\kappa = \log(N + 1) \geq \lambda$ and $T_i = 4^i \kappa (N(2N)^{3/5})^{1/2}$. Since $U^2 X = 225 \cdot 16^i (16^i X) \leq 225 \cdot 16^i N$ and $X + 1 \leq 2N$, we get $A_1 \leq 225T_i^2$ and $A_3 \leq 1350T_i^2 \leq 37^2 T_i^2$. With $w = UX \leq 15N$,

$X \leq N/16^i$ and $R \leq r \leq N/R$,

$$\begin{aligned} Z^2 &\leq 4U^2 X^3 Q = 320w^3 X + \frac{48w^4}{r} + 80rw^3 + 12Uw^3 \\ &\leq \frac{320 \cdot 15^3 N^4}{16^i} + \frac{48 \cdot 15^4 N^4}{R} + \frac{80 \cdot 15^3 N^4}{R} \\ &\quad + 180 \cdot 15^3 \cdot 16^i N^3, \end{aligned}$$

and every coefficient is at most 40^4 , so $Z^{1/2} \leq 40(N2^{-i} + 2NR^{-1/4} + 2^i N^{3/4})$. The block is therefore at most

$$\begin{aligned} 15T_i + 40\sqrt{800e^{15}}(N2^{-i} + 2NR^{-1/4} + 2^i N^{3/4}) + 37T_i \\ \leq K_b(N2^{-i} + NR^{-1/4} + 2^i N^{3/4} + T_i), \end{aligned}$$

with $K_b = 80\sqrt{800e^{15}} + 52$. Summing over $i < I$,

$$|T_2| \leq K_b(2N + INR^{-1/4} + 2^I N^{3/4} + 4^I \kappa(N(2N)^{3/5})^{1/2}).$$

Here $I \leq 2L^3$; $2^I \leq 4L \leq 16N^{1/4}$; $4^I \leq 16L^2$ and $\kappa \leq 2L$; and $\log x \leq 15x^{1/15}$ gives $L^3(N(2N)^{3/5})^{1/2} \leq 2 \cdot 15^3 N$. Hence $|T_2| \leq K_b(216018N + 2L^3 NR^{-1/4})$.

Altogether $|S|L \leq (C_5 + 908 + 216018K_b)N + (60 + 2K_b)L^3 NR^{-1/4}$, and dividing by L gives the theorem with $C_1 = 45 + C_5 + 908 + 216018K_b + 60 + 2K_b$. $\square$

5. PROOF OF THEOREM 1.1

In this section χ is a primitive character modulo $q \geq 3$.

Lemma 5.1. *Let $B \geq 0$ be such that $|\sum_{m \leq H} \bar{\chi}(m)e(m\beta)/m| \leq B$ for every integer $H \leq q/2$ and every $\beta \in \mathbb{R}$. Then $|\sum_{n \leq x} \chi(n)| \leq \sqrt{q}(B + 6)$ for every $x \geq 0$.*

Proof. We may take $x = T$ to be an integer. Let $\tau = \sum_{a \bmod q} \bar{\chi}(a)e(a/q)$. Since $\bar{\chi}$ is primitive, $\sum_a \bar{\chi}(a)e(an/q) = \chi(n)\tau$ for every n , and summing the squares of the moduli of these identities over $n \bmod q$ gives $q\varphi(q) = |\tau|^2 \varphi(q)$, so $|\tau| = \sqrt{q}$. Summing over $1 \leq n \leq T$,

$$\tau \sum_{n \leq T} \chi(n) = \sum_{0 < |a| \leq q/2} \bar{\chi}(a) \frac{e(aT/q) - 1}{1 - e(-a/q)},$$

where a runs over the integers with $-q/2 < a \leq q/2$, $a \neq 0$. For $0 < |\vartheta| \leq 1/2$ put $r(\vartheta) = (1 - e(-\vartheta))^{-1} - (2\pi i \vartheta)^{-1}$; then $|1 - e(-\vartheta)| = 2|\sin \pi \vartheta| \geq 4|\vartheta|$. If $2\pi|\vartheta| \leq 1$, then $|e(-\vartheta) - 1 + 2\pi i \vartheta| \leq (2\pi \vartheta)^2$ and $|r(\vartheta)| \leq (2\pi \vartheta)^2 / (4|\vartheta| \cdot 2\pi|\vartheta|) < 3$; otherwise $|r(\vartheta)| \leq 1/(4|\vartheta|) + 1/(2\pi|\vartheta|) < 3$. The terms with $r(a/q)$ contribute at most $6q$. Writing $a = \pm m$ with $1 \leq m \leq q/2$, the remaining terms are $q/(2\pi i)$ times a combination, with coefficients of modulus at most 1, of four sums $\sum_{m \leq H} \bar{\chi}(m)e(m\beta)/m$ with $H \leq q/2$, so they contribute at most $2qB/\pi$. Dividing by $|\tau|$ gives the claim. $\square$

Lemma 5.2. *For $Y \geq 2$ and $H \geq 1$,*

$$\sum_{\substack{n \leq H \\ P(n) \leq Y}} \frac{1}{n} \leq e^{11} \log Y.$$

Proof. The sum is at most $\prod_{p \leq Y} (1 - 1/p)^{-1} \leq \exp(\sum_{p \leq Y} \frac{1}{p-1})$, and $\sum_p (\frac{1}{p-1} - \frac{1}{p}) \leq 1$; apply Lemma 2.1(iv). $\square$

Lemma 5.3. *Let $0 < r < q$ and let ψ be a character modulo r . Then the character $\chi\psi$ modulo qr is nonprincipal. Moreover, if χ_1, χ_2 are primitive characters modulo q with $\chi_1\psi = \chi_2\psi$, then $\chi_1 = \chi_2$.*

Proof. Suppose $\chi\psi$ is principal, so that $\chi(n) = \bar{\psi}(n)$ for every n prime to qr . Let $d = (q, r)$ and let n be prime to q with $n \equiv 1 \pmod{d}$. By the Chinese remainder theorem there is n' prime to qr with $n' \equiv n \pmod{q}$ and $n' \equiv 1 \pmod{r}$, and then $\chi(n) = \chi(n') = \bar{\psi}(n') = 1$. So χ is trivial on the units congruent to 1 modulo d , and is induced by a character modulo $d \leq r < q$, which contradicts primitivity. For the second claim, $\chi_1(n) = \chi_2(n)$ for every n prime to qr , and for n prime to q there is, again by the Chinese remainder theorem, an n' prime to qr with $n' \equiv n \pmod{q}$, so that $\chi_1(n) = \chi_1(n') = \chi_2(n') = \chi_2(n)$. $\square$

Lemma 5.4. *Let $1/2 \leq \theta < 1$, with the parameters (3.1) and the constant C_3 of Corollary 3.8, and put $C_2 = 2C_3/\eta_\theta$. Let $0 < r < q$, $b \in \mathbb{Z}$, $\beta \in \mathbb{R}$, let $Y \geq \max(2, r)$ and $t \geq 1$ be integers, and suppose that $L(s, \chi\psi) \neq 0$ for $\operatorname{Re} s > \theta$, $|\Im s| \leq 4t^2 + 2$ for every character ψ modulo r . Then*

$$\left| \sum_{\substack{n \leq t \\ P(n) > Y}} \chi(n) e\left(n\left(\frac{b}{r} + \beta\right)\right) \right| \leq C_2 \frac{r \log(qr + 2) (2 + 2\pi|\beta|t) t}{Y^{\eta_\theta} \log Y}.$$

Proof. Write $\eta = \eta_\theta$. Every $n \leq t$ with $P(n) > Y$ is uniquely $n = pm$ with $p = P(n)$ prime, $Y < p \leq t/m$ and $P(m) \leq p$, and then $m \leq t/Y$. For such p we have $p > r$, so $p \nmid r$, and on the units modulo r the function $u \mapsto e(umb/r)$ equals $\sum_\psi c_\psi \psi(u)$, the sum over the characters ψ modulo r , with $|c_\psi| \leq 1$. By Lemma 5.3 each $\chi\psi$ is nonprincipal modulo qr , and by hypothesis it satisfies the condition of Corollary 3.8 with $x = t$. Applying (3.3) to $\xi = \chi\psi$ with $\gamma = m\beta$, $A = \max(Y, P(m) - 1)$ and $B = \lfloor t/m \rfloor \leq t$ (the inner sum is empty if $A \geq B$) gives

$$\left| \sum_{\substack{n \leq t \\ P(n) > Y}} \chi(n) e\left(n\left(\frac{b}{r} + \beta\right)\right) \right| \leq 2C_3 r \frac{\log(qr + 2)(2 + 2\pi|\beta|t)}{\log Y} \sum_{m \leq t/Y} \left(\frac{t}{m}\right)^{1-\eta},$$

and $\sum_{m \leq M} m^{-1+\eta} \leq \int_0^M u^{\eta-1} du = M^\eta/\eta$, so the last sum is at most $t/(\eta Y^\eta)$. $\square$

Proposition 5.5. *Let $1/2 \leq \theta < 1$, with the parameters (3.1), the constant C_1 of Theorem 4.1 and the constant C_2 of Lemma 5.4, and put $A_\theta = \lceil 36/\eta_\theta \rceil$ and $C_4 = 1 + 5C_1 + C_2 + 4e^{11}A_\theta$. Suppose that $\log q \geq 200$, put $M = \lceil \log q \rceil$ and $R = M^{16}$, and suppose that (1.1) holds for every $r \leq R$ and every character ψ modulo r . Then $|\sum_{m \leq H} \chi(m) e(m\beta)/m| \leq C_4 \log \log q$ for every integer $H \leq q$ and every $\beta \in \mathbb{R}$.*

Proof. Let $L = \log q$, $\eta = \eta_\theta$, $A = A_\theta$ and $Y = M^A$. Since $e^{L/16} \geq (L/16)^4/24 > L + 1 \geq M$, we have $2 \leq R < q$, and $R \leq Y$ since $A \geq 288$. Split the sum according to $P(m) \leq Y$ or $P(m) > Y$. By Lemma 5.2 the first part is at most $e^{11} \log Y$. For the second put $A(t) = \sum_{n \leq t, P(n) > Y} \chi(n) e(n\beta)$ and $w(t) = 1/t - 1/(t+1)$. Partial summation and $|A(H)| \leq H$ give

$$\left| \sum_{\substack{m \leq H \\ P(m) > Y}} \frac{\chi(m) e(m\beta)}{m} \right| \leq 1 + \sum_{1 \leq t < H} |A(t)| w(t).$$

If $t \leq Y$, then $A(t) = 0$. Let $t > Y$ and $Q = \lfloor t/R \rfloor \geq 1$. By Dirichlet's theorem there are coprime a, r with $1 \leq r \leq Q$ and $|\beta - a/r| \leq 1/(r(Q+1))$; then $|\beta - a/r|t \leq R$ (Figure 3). If $r \leq R$, then $r < q$, and $4t^2 + 2 \leq 4q^2 + 2 \leq 5q^2$ because $t < H \leq q$ and $q \geq 3$, so (1.1) gives the hypothesis of Lemma 5.4 for this r ; that lemma with $b = a$ gives $|A(t)| \leq C_2 kt$, where

$$k = \frac{R \log(qR + 2)(2 + 2\pi R)}{Y^\eta \log Y}.$$

If $r > R$, then $R \leq r \leq t/R$ and $|\beta - a/r| \leq r^{-2}$, so Theorem 4.1 for $f = \chi$ bounds the full sum $\sum_{n \leq t} \chi(n) e(n\beta)$, and since $\log t \leq L$, removing the $\Psi(t) = \#\{n \leq t : P(n) \leq Y\}$ smooth terms gives

$$|A(t)| \leq C_1 \left(\frac{t}{\log t} + t\rho \right) + \Psi(t), \quad \rho = \frac{L^2}{R^{1/4}} = \frac{L^2}{M^4}.$$

Since $tw(t) \leq 1/t$ and $w(t) \leq 1/t^2$,

$$\sum_{1 \leq t < H} |A(t)| w(t) \leq C_1 \sum_{2 \leq t \leq H} \frac{1}{t \log t} + (C_1 \rho + C_2 k) \sum_{t \leq H} \frac{1}{t} + \sum_{1 \leq t < H} \Psi(t) w(t).$$

The first sum is at most $3 + \log L$, the second at most $1 + L$, and by partial summation the third is at most $\sum_{n \leq H, P(n) \leq Y} 1/n \leq e^{11} \log Y$. Finally $L \leq M \leq 2L$ gives $\rho(1 + L) \leq 2/M \leq 1$; $\log(qR + 2) \leq 18M$, $2 + 2\pi R \leq 8M^{16}$, $Y^\eta = M^{A\eta} \geq M^{36}$ and $\log Y = A \log M \geq \log M$ give $k(1 + L) \leq 288M^{34}/(M^{36} \log M) \leq 1$; and $\log Y = A \log M \leq 2A \log L$, since $\log M \leq \log 2 + \log L \leq 2 \log L$. Collecting terms, the sum is at most $1 + 4C_1 + C_2 + (C_1 + 4e^{11}A) \log L \leq (1 + 5C_1 + C_2 + 4e^{11}A) \log L$, as $\log L \geq 1$. $\square$

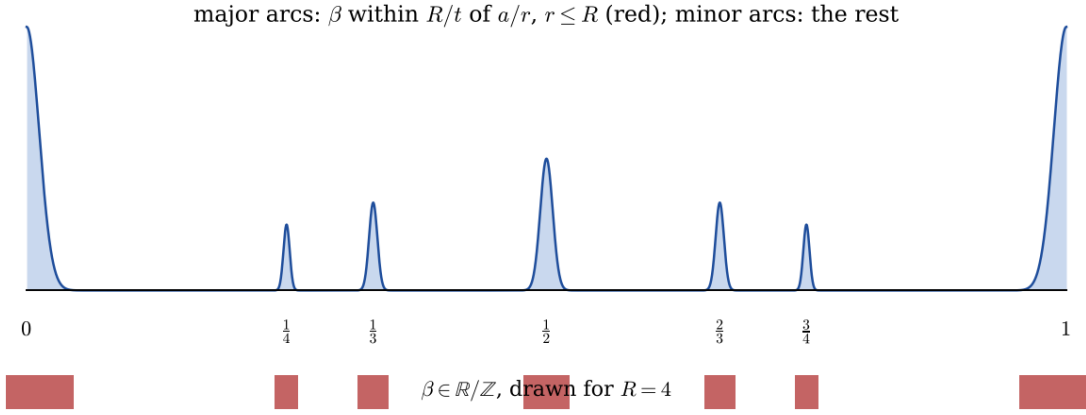


FIGURE 3. The dissection of $\mathbb{R}/\mathbb{Z}$ in the proof of Proposition 5.5, drawn for $R = 4$. Near a fraction a/r with $r \leq R$ (major arcs) the sum $A(t)$ is controlled by the twisted prime sums of Corollary 3.8; elsewhere (minor arcs) by Theorem 4.1. The heights only indicate that the major arcs carry the large values.

Proof of Theorem 1.1. Put $C_\theta = \max(C_4 + 6, 207/\log \log 3)$ with C_4 as in Proposition 5.5. If $\log q < 200$, then every sum $\sum_{m \leq H} \bar{\chi}(m)e(m\beta)/m$ with $H \leq q$ has modulus at most $1 + \log q < 201$, and Lemma 5.1 gives

$$\left| \sum_{n \leq x} \chi(n) \right| \leq 207\sqrt{q} \leq \frac{207}{\log \log 3} \sqrt{q} \log \log q$$

without any hypothesis. Let $\log q \geq 200$. The character $\bar{\chi}$ is primitive modulo q , and it satisfies (1.1) for every $r \leq R$ and ψ modulo r : the zeros of $L(s, \bar{\chi}\psi)$ are the complex conjugates of the zeros of $L(s, \chi\bar{\psi})$, $\bar{\psi}$ is again a character modulo r , and the rectangle in (1.1) is symmetric about the real axis. Hence Proposition 5.5 applies to $\bar{\chi}$, and Lemma 5.1 gives

$$\left| \sum_{n \leq x} \chi(n) \right| \leq \sqrt{q} (C_4 \log \log q + 6) \leq (C_4 + 6) \sqrt{q} \log \log q. \quad \square$$

6. THE EXCEPTIONAL CHARACTERS

Proof of Theorem 1.2. It suffices to treat $0 < \varepsilon \leq 1$, since for larger ε the statement is weaker. Put $\theta = 1 - \varepsilon/48$, so that $47/48 \leq \theta < 1$, and let C_θ be the constant of Theorem 1.1. Let q satisfy $\log q \geq 200$, put $L = \log q$, $M = \lceil L \rceil$, $R = M^{16}$, and let E_q be the set of primitive characters χ modulo q for which (1.1) fails for some $r \leq R$ and some ψ modulo r . By Theorem 1.1, every primitive $\chi \notin E_q$ has $M(\chi) \leq C_\theta \sqrt{q} \log \log q$, so it remains to bound $|E_q|$.

For $\chi \in E_q$ choose $r \leq R$, ψ modulo r and a zero ρ of $L(s, \chi\psi)$ with $\operatorname{Re} \rho > \theta$ and $|\Im \rho| \leq 5q^2$. The character $\chi\psi$ is a character modulo qr , and ρ is counted in $N(\theta, 5q^2, \chi\psi)$.

For fixed r and ψ the map $\chi \mapsto \chi\psi$ is injective by Lemma 5.3, so each term $N(\theta, 5q^2, \xi)$, ξ modulo qr , is charged by at most one χ for each pair (r, ψ) . Each $\chi\psi$ is nonprincipal by the same lemma. Hence, by Theorem 2.3 with $k = qr \geq 3$ and $T = 5q^2$, and since $\theta \geq 11/12$,

$$\begin{aligned} |E_q| &\leq \sum_{r \leq R} \sum_{\psi \bmod r} \sum_{\substack{\xi \bmod qr \\ \xi \neq \xi_0}} N(\theta, 5q^2, \xi) \leq c_0 \sum_{r \leq R} \varphi(r) (5q^3 r)^{12(1-\theta)} (\log(5q^3 r))^9 \\ &\leq c_0 R^2 (5q^3 R)^{\varepsilon/4} (\log(5q^3 R))^9. \end{aligned}$$

Now $M \leq L + 1 \leq 2L$, so $R \leq 2^{16} L^{16}$ and $(5R)^{1/4} \leq (5 \cdot 2^{16})^{1/4} L^4 \leq 2^5 L^4$; and $\log(5q^3 R) \leq \log(5 \cdot 2^{16}) + 3L + 16 \log(2L) \leq 20L$, since $L \geq 200$. Therefore

$$|E_q| \leq c_0 2^{37} 20^9 L^{45} q^{3\varepsilon/4}. \quad (6.1)$$

Since $(\log q)^{45} q^{-\varepsilon/4} \rightarrow 0$ as $q \rightarrow \infty$, there is $q_0 = q_0(\varepsilon) \geq e^{200}$ such that $c_0 2^{37} 20^9 (\log q)^{45} \leq q^{\varepsilon/4}$ for $q \geq q_0$, and then $|E_q| \leq q^{\varepsilon}$.

For $3 \leq q < q_0$ no character is exceptional: every sum $\sum_{m \leq H} \bar{\chi}(m) e(m\beta)/m$ with $H \leq q$ has modulus at most $1 + \log q$, so Lemma 5.1 gives $M(\chi) \leq (7 + \log q_0) \sqrt{q} \leq ((7 + \log q_0)/\log \log 3) \sqrt{q} \log \log q$. The theorem holds with $C_\varepsilon = \max(C_\theta, (7 + \log q_0)/\log \log 3)$. $\square$

The bound (6.1) is where the exponent ε comes from: the zero-density theorem gives a power $q^{36(1-\theta)}$ of the modulus, times a power of $\log q$, and θ is chosen so that $36(1-\theta) = 3\varepsilon/4$. Any zero-density theorem of the shape (2.1), with $12(1-\sigma)$ replaced by $c(1-\sigma)$ for an absolute c , gives the same conclusion with $\theta = 1 - \varepsilon/(4c)$; with the exponents of Montgomery's Theorem 12.2 one may take $\theta = 1 - \varepsilon/8$.

Proof of Corollary 1.3. Let χ be primitive modulo $q \geq 3$, $r \leq R < q$ and ψ a character modulo r . By Lemma 5.3 the character $\chi\psi$ is nonprincipal, so by hypothesis $L(s, \chi\psi) \neq 0$ for $\operatorname{Re} s > \theta$, and in particular (1.1) holds. Theorem 1.1 gives the claim. $\square$

7. THE LOGARITHMIC DERIVATIVE AT $s = 1$

In this section χ is a nonprincipal character modulo $q \geq 3$, not necessarily primitive.

Proposition 7.1. *Let $1/2 \leq \theta < 1$, with the parameters (3.1), let K_1 be the constant of Lemma 3.5 with $\delta = \delta_\theta$, and put*

$$T_1 = (\log(q+2))^{4/(1-\theta)} + 2, \quad C_5 = \frac{6}{1-\sigma_\theta} + 429K_1 \left(1 + \frac{1}{1-\sigma_\theta}\right) + 99.$$

If $L(s, \chi) \neq 0$ for $\operatorname{Re} s > \theta$, $|\Im s| \leq T_1$, then $|L'/L(1, \chi)| \leq C_5 \log \log q$.

Proof. Write $\sigma = \sigma_\theta$, $a = a_\theta$, $\delta = \delta_\theta$, and put

$$x = (\log(q+2))^{1/(1-\sigma)}, \quad T_0 = x^2,$$

so that $x > 1$, $x^{1-\sigma} = \log(q+2)$, and $T_0 + 2 \leq T_1$ by (3.2). Note that $L(1, \chi) \neq 0$ by hypothesis, since $\operatorname{Re} 1 > \theta$.

Let $a(n) = \chi(n)\Lambda(n)/n$ and $F(x) = S(a; x) = \sum_{n \leq x} \chi(n)\Lambda(n)n^{-1}(1 - n/x)$. The Dirichlet series of a is $-L'/L(s+1, \chi)$, so Lemma 3.1(ii) gives $2\pi F(x) = \int_{-\infty}^{\infty} \Phi(2+it) dt$ with

$$\Phi(s) = -\frac{L'}{L}(s+1, \chi) \frac{x^s}{s(s+1)}.$$

For $|t| > T_0$ Lemmas 3.3 and 3.6 give $|\Phi(2+it)| \leq 16x^2(1+|t|)^{-2}$, so this part of the integral has modulus at most $32x^2/(1+T_0) \leq 32$.

Consider the rectangle $\sigma - 1 \leq \operatorname{Re} s \leq 2$, $|\Im s| \leq T_0$. On a neighbourhood of it the function $-L'/L(s+1, \chi)$ is holomorphic except for a simple pole at $s = 0$ with residue 1 (the pole of $L(s+1, \chi)$ is excluded since χ is nonprincipal): indeed $\operatorname{Re}(s+1) \geq \sigma > \theta$ by (3.2) and $|\Im(s+1)| \leq T_0 < T_1$, so $L(s+1, \chi) \neq 0$ there except at $s = 0$, where $L(1, \chi) \neq 0$ as well,

and the pole of $-L'/L(s+1, \chi)$ at $s=0$ is the pole of the kernel $1/(s(s+1))$. The pole of the kernel at $s=-1$ lies outside the rectangle, since $\sigma-1 > -1/2$. Thus Φ has in the rectangle exactly one pole, a simple pole at $s=0$ with residue $-L'/L(1, \chi)$. By the residue theorem the integral of Φ over the right side of the rectangle equals $-2\pi L'/L(1, \chi)$ plus the integral over the left side plus the integrals over the two horizontal sides, with the appropriate orientations. Consequently

$$2\pi \left| \frac{L'}{L}(1, \chi) \right| \leq 2\pi |F(x)| + 32 + I_{\text{left}} + I_{\text{top}} + I_{\text{bottom}}, \quad (7.1)$$

where the three I are the moduli of the integrals of Φ over the left and horizontal sides.

By Lemma 2.1(v), $|F(x)| \leq \sum_{n \leq x} \Lambda(n)/n \leq \log x + 3$.

On the left side $s = \sigma - 1 + it$ with $|t| \leq T_0$. Lemma 3.5 applies to $L'/L(\sigma + it, \chi)$, since $a + 8\delta = \sigma \leq 2$ and the hypothesis gives $L(s, \chi) \neq 0$ for $\text{Re } s > a + 2\delta$ (as $a + 2\delta \geq \theta$) and $|\Im s - t| \leq 2$ (as $T_0 + 2 \leq T_1$); with Lemma 3.6 it gives $|L'/L(\sigma + it, \chi)| \leq 10K_1 \log(q+2)(1+|t|)^{1/2}$. Moreover $|s| \geq \max(1-\sigma, |t|) \geq (1-\sigma)(1+|t|)/2$ and $|s+1| = |\sigma + it| \geq \max(1/2, |t|) \geq (1+|t|)/3$, so $|s(s+1)| \geq (1-\sigma)(1+|t|)^2/6$. Hence

$$I_{\text{left}} \leq \int_{-\infty}^{\infty} \frac{60K_1 \log(q+2) x^{\sigma-1}}{(1-\sigma)(1+|t|)^{3/2}} dt = \frac{240K_1}{1-\sigma},$$

by the choice of x .

On the horizontal sides $s = \sigma' \pm iT_0$ with $\sigma - 1 \leq \sigma' \leq 2$, and $T_0 \geq 1$. For $\sigma' \leq 1$ Lemmas 3.5 and 3.6 give $|L'/L(s+1, \chi)| \leq 10K_1 \log(q+2)(1+T_0)^{1/2}$, and for $\sigma' > 1$ the same bound holds by Lemma 3.3, since $K_1 \geq 1$. Also $|x^s| \leq x^2$ and $|s(s+1)| \geq T_0^2 \geq (1+T_0)^2/4$. Each horizontal side has length at most 3, so

$$I_{\text{top}}, I_{\text{bottom}} \leq 3 \cdot 40K_1 \log(q+2) x^2 (1+T_0)^{-3/2} \leq 120K_1 \log(q+2) x^{-1} = 120K_1 x^{-\sigma} \leq 120K_1.$$

Inserting these bounds in (7.1) and dividing by 2π ,

$$\left| \frac{L'}{L}(1, \chi) \right| \leq \log x + 3 + 6 + \frac{39K_1}{1-\sigma} + 39K_1 = \frac{\log \log(q+2)}{1-\sigma} + 39K_1 \left(1 + \frac{1}{1-\sigma}\right) + 9.$$

Finally $\log \log(q+2) \leq 6 \log \log q$ for $q \geq 3$, since the ratio of the two sides decreases in q and is below 5.1 at $q=3$, and $1 \leq 11 \log \log q$ since $\log \log 3 > 1/11$. This gives the constant C_5 . $\square$

Proof of Theorem 1.4. As before let $0 < \varepsilon \leq 1$, $\theta = 1 - \varepsilon/48$, and let C_5 and $T_1 = (\log(q+2))^{192/\varepsilon} + 2$ be as in Proposition 7.1. Let E'_q be the set of nonprincipal characters χ modulo q for which $L(s, \chi)$ has a zero with $\text{Re } s > \theta$ and $|\Im s| \leq T_1$. By Proposition 7.1 every nonprincipal $\chi \notin E'_q$ satisfies $|L'/L(1, \chi)| \leq C_5 \log \log q$, and by Theorem 2.3, as E'_q contains only nonprincipal characters and $\theta \geq 11/12$,

$$|E'_q| \leq \sum_{\substack{\xi \bmod q \\ \xi \neq \xi_0}} N(\theta, T_1, \xi) \leq c_0 (qT_1)^{\varepsilon/4} (\log qT_1)^9.$$

Here $T_1 \leq 2(\log(q+2))^{192/\varepsilon}$, so $T_1^{\varepsilon/4} \leq 2(\log(q+2))^{48}$, and $\log(qT_1) \leq \log q + 1 + 192\varepsilon^{-1} \log \log(q+2)$. Hence $|E'_q| \leq q^{\varepsilon/4} P_\varepsilon(\log q)$ for a polynomial P_ε , and there is $q_1 = q_1(\varepsilon)$ with $|E'_q| \leq q^\varepsilon$ for $q \geq q_1$. For $3 \leq q < q_1$ there are finitely many nonprincipal characters, each with $L(1, \chi) \neq 0$ by Dirichlet's theorem [6, Chapter 4], so $|L'/L(1, \chi)|/\log \log q$ is bounded on them by a constant C' depending on ε ; no character is exceptional there. The theorem holds with $C_\varepsilon = \max(C_5, C')$. $\square$

APPENDIX A. PROOF OF THE ZERO-DENSITY THEOREM

This appendix proves Theorem 2.3. The argument is the one of Ingham and Montgomery [16, 21]: a zero of $L(s, \chi)$ makes a mollified Dirichlet polynomial, or a mollified integral of L , large at that zero, and the mean values of these objects over all characters to the modulus and

over the height are small. Every estimate below is proved with an explicit constant. The exponent $12(1 - \sigma)$ obtained is far from the best known, which is why the theorem is stated for $11/12 \leq \sigma \leq 1$ only; the proofs of Theorems 1.2 and 1.4 use nothing more.

Throughout, $k \geq 3$, $T \geq 2$, χ runs over the nonprincipal characters modulo k , and

$$\mathcal{L} = \log(kT) \geq \log 6, \quad \sigma_1 = \frac{5}{6}, \quad X = kT, \quad Y = (kT)^3,$$

and σ satisfies $11/12 \leq \sigma \leq 1$. Note that $\mathcal{L} \geq 1.79$, $Y \geq 216$ and $\sigma - \sigma_1 \geq 1/12$.

A.1. Lemmas.

Lemma A.1. *Let χ be nonprincipal modulo k , let $N \geq 1$ be an integer and $\operatorname{Re} s = \sigma > 0$. Then*

$$L(s, \chi) = \sum_{n \leq N} \chi(n) n^{-s} + E_N(s, \chi), \quad |E_N(s, \chi)| \leq k \left(1 + \frac{|s|}{\sigma}\right) N^{-\sigma}.$$

Proof. Put $A(t) = \sum_{n \leq t} \chi(n)$, so that $|A(t)| \leq k$ because the sum of χ over a period vanishes. For $\operatorname{Re} s > 1$ partial summation gives $\sum_{n > N} \chi(n) n^{-s} = -A(N)N^{-s} + s \int_N^\infty A(t) t^{-s-1} dt$; the right side is holomorphic for $\sigma > 0$ and bounded by $kN^{-\sigma} + |s|kN^{-\sigma}/\sigma$ there, and so is the analytic continuation of the left side, since $L(s, \chi)$ is entire.⁵ $\square$

Lemma A.2. *Let χ be nonprincipal modulo k and $t \in \mathbb{R}$. The number of zeros $\rho = \beta + i\gamma$ of $L(s, \chi)$ with $\beta \geq 3/4$ and $|\gamma - t| \leq 1/2$, counted with multiplicity, is at most $7 \log(8k(|t| + 4))$.*

Proof. Let $c = 2 + it$. Jensen's formula [27, Theorem 15.18] for $L(s, \chi)$ on the disc $|s - c| \leq 7/4$ reads

$$\log |L(c, \chi)| + \sum_{|\rho - c| < 7/4} \log \frac{7/4}{|\rho - c|} = \frac{1}{2\pi} \int_0^{2\pi} \log \left| L\left(c + \frac{7}{4} e^{i\phi}, \chi\right) \right| d\phi,$$

the sum running over the zeros in the open disc with multiplicity. A zero with $\beta \geq 3/4$ and $|\gamma - t| \leq 1/2$ has $|\rho - c|^2 \leq (5/4)^2 + (1/2)^2 = 29/16 < 9/4$, so it lies in the disc $|\rho - c| \leq 3/2$ and contributes at least $\log(7/6)$ to the sum. On the circle $|s - c| = 7/4$ we have $\operatorname{Re} s \geq 1/4$ and $|s| \leq |t| + 4$, so $|L(s, \chi)| \leq 4k(|t| + 4)$ by Lemma 3.2. At c the Euler product gives $|L(c, \chi)| = \prod_p |1 - \chi(p)p^{-c}|^{-1} \geq \prod_p (1 + p^{-2})^{-1} = \zeta(4)/\zeta(2) = \pi^2/15 > 1/2$. Hence the number of zeros in question is at most $(\log(4k(|t| + 4)) + \log 2)/\log(7/6) \leq 7 \log(8k(|t| + 4))$, as $1/\log(7/6) < 6.49$. $\square$

Lemma A.3. *Let $k \geq 1$, $T' \geq 1$, $N \geq 1$ and let $b_1, \dots, b_N$ be complex numbers. Then*

$$\sum_{\chi \bmod k} \int_{-T'}^{T'} \left| \sum_{n \leq N} b_n \chi(n) n^{-it} \right|^2 dt \leq (43N + 22kT') \sum_{n \leq N} |b_n|^2,$$

the sum on the left running over all characters modulo k .

Proof. Let $K(t) = \max(0, 1 - |t|/(2T'))$, so that $K \geq 1/2$ on $[-T', T']$ and its Fourier transform is

$$\hat{K}(\xi) = \int_{-\infty}^{\infty} K(t) e^{i\xi t} dt = 2T' \left(\frac{\sin(T'\xi)}{T'\xi} \right)^2 \leq 2T' \min(1, (T'\xi)^{-2}),$$

which is nonnegative.⁶ Writing $S_\chi(t)$ for the inner sum, the integral on the left is at most $2 \int K(t) |S_\chi(t)|^2 dt$, and

$$\int_{-\infty}^{\infty} K(t) |S_\chi(t)|^2 dt = \sum_{m, n \leq N} b_m \bar{b}_n \chi(m) \bar{\chi}(n) \hat{K}(\log(n/m)).$$

⁵The continuation of $\sum_{n \leq N} \chi(n) n^{-s}$ is itself, and that of $L(s, \chi)$ is given by the same partial summation from $n = 1$, as in the proof of Lemma 3.2.

⁶For a triangle of half-width a and height 1 the transform is $a(\sin(a\xi/2)/(a\xi/2))^2$; here $a = 2T'$.

Summing over the characters, $\sum_{\chi} \chi(m) \bar{\chi}(n)$ equals $\varphi(k)$ if $m \equiv n \pmod{k}$ and $(mn, k) = 1$, and 0 otherwise. Since $\hat{K} \geq 0$ and $|b_m \bar{b}_n| \leq (|b_m|^2 + |b_n|^2)/2$, the total is at most

$$\varphi(k) \sum_{m \leq N} |b_m|^2 \sum_{\substack{n \leq N \\ n \equiv m \pmod{k}}} \hat{K}(\log(n/m)).$$

Fix m . The integers $n \equiv m \pmod{k}$ with $|\log(n/m)| \leq 1/T'$ lie in an interval of length $m(e^{1/T'} - e^{-1/T'}) \leq 2m(e^{1/T'} - 1) \leq 4N/T'$, since $e^x - 1 \leq (e - 1)x \leq 2x$ for $0 \leq x \leq 1$; there are at most $4N/(kT') + 1$ of them, and each has $\hat{K} \leq 2T'$. For $j \geq 1$, the $n \equiv m \pmod{k}$ with $j/T' < |\log(n/m)| \leq (j+1)/T'$ lie in two intervals, $(me^{j/T'}, me^{(j+1)/T'})$ and $[me^{-(j+1)/T'}, me^{-j/T'}]$; the first is empty unless $me^{j/T'} \leq N$, in which case its length is at most $N(e^{1/T'} - 1) \leq 2N/T'$, and the second has length at most $2m/T' \leq 2N/T'$. So there are at most $4N/(kT') + 2$ such n , each with $\hat{K} \leq 2T'/j^2$. Altogether

$$\sum_{n \equiv m \pmod{k}} \hat{K}(\log(n/m)) \leq 2T' \left(\frac{4N}{kT'} + 2 \right) \left(1 + \sum_{j \geq 1} \frac{1}{j^2} \right) \leq 2T' \left(\frac{4N}{kT'} + 2 \right) \cdot 2.65 \leq \frac{21.2N}{k} + 10.6T',$$

and the lemma follows on multiplying by $2\varphi(k) \leq 2k$. $\square$

Lemma A.4. For $X \geq 1$ and $U \geq 1$, with $\tau_X(n) = \#\{d \leq X : d \mid n\}$,

$$\sum_{n \leq U} \tau_X(n)^2 \leq U(1 + \log X)^3.$$

In particular $|\sum_{d \mid n, d \leq X} \mu(d)| \leq \tau_X(n)$, and $\tau_X(n) \leq d(n) \leq 2\sqrt{n}$.

Proof. Expanding the square, the sum is

$$\sum_{d_1, d_2 \leq X} \#\{n \leq U : [d_1, d_2] \mid n\} \leq U \sum_{d_1, d_2 \leq X} \frac{(d_1, d_2)}{d_1 d_2}.$$

Writing $g = (d_1, d_2)$, $d_i = ge_i$, this is at most $U \sum_{g \leq X} g^{-1} (\sum_{e \leq X/g} e^{-1})^2 \leq U(1 + \log X)^3$. The bound $d(n) \leq 2\sqrt{n}$ holds because the divisors pair off as $d \leftrightarrow n/d$ with $\min(d, n/d) \leq \sqrt{n}$. $\square$

Lemma A.5. Let f be holomorphic on a neighbourhood of the closed disc $|z - z_0| \leq r$. Then

$$|f(z_0)|^2 \leq \frac{1}{\pi r^2} \iint_{|z - z_0| \leq r} |f(z)|^2 dA(z).$$

Proof. For $0 < \rho \leq r$ the mean of f over the circle $|z - z_0| = \rho$ is $f(z_0)$. Multiplying by ρ and integrating over $0 \leq \rho \leq r$ gives $f(z_0) r^2/2 = (2\pi)^{-1} \iint f dA$, so $f(z_0)$ is the mean of f over the disc, and the claim is the Cauchy–Schwarz inequality. $\square$

Lemma A.6. For $u > 0$ and $c > 0$ put

$$\Phi(u) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{w^2} u^{-w} \frac{dw}{w}.$$

The integral does not depend on c , and (i) $|\Phi(u)| \leq e^{-(\log u)^2/4}/(\sqrt{\pi} \log u)$ for $u > 1$; (ii) $|\Phi(u) - 1| \leq e^{-(\log u)^2/4}/(\sqrt{\pi} |\log u|)$ for $0 < u < 1$; (iii) $|\Phi(u)| \leq 2$ for every $u > 0$.

Proof. On the line $\operatorname{Re} w = c$ we have $|e^{w^2}| = e^{c^2 - v^2}$, $w = c + iv$, so the integral converges absolutely, and by Cauchy's theorem it does not change when $c > 0$ is moved, the integrand being holomorphic for $\operatorname{Re} w > 0$ and tending to zero as $|\Im w| \rightarrow \infty$ uniformly in bounded vertical strips. Moving the line to $\operatorname{Re} w = -c < 0$ crosses the simple pole at $w = 0$, of residue 1, so

$$\Phi(u) = 1 + \frac{1}{2\pi i} \int_{-c-i\infty}^{-c+i\infty} e^{w^2} u^{-w} \frac{dw}{w} \quad (c > 0). \quad (\text{A.1})$$

Now $|u^{-w}| = u^{-\operatorname{Re} w}$ and $|e^{w^2}| = e^{(\operatorname{Re} w)^2 - v^2}$, and $\int e^{-v^2} dv = \sqrt{\pi}$. For $u > 1$ take $c = (\log u)/2$ in the definition: then $u^{-c} e^{c^2} = e^{-c^2} = e^{-(\log u)^2/4}$ and $|w| \geq c$, which gives (i). For $u < 1$ take

$c = |\log u|/2$ in (A.1): then $u^c e^{c^2} = e^{-c^2}$ and (ii) follows in the same way. For (iii), if $u \geq 1$ take $c = 1/2$ in the definition: $u^{-c} \leq 1$, $e^{c^2} = e^{1/4}$ and $1/|w| \leq 2$ give $|\Phi(u)| \leq e^{1/4}/\sqrt{\pi} < 0.73$; if $u < 1$ take $c = 1/2$ in (A.1) and get $|\Phi(u) - 1| < 0.73$. $\square$

For $X \geq 1$ and a character χ modulo k put

$$M_X(s, \chi) = \sum_{d \leq X} \mu(d) \chi(d) d^{-s}, \quad a(n) = \sum_{\substack{d|n \\ d \leq X}} \mu(d),$$

so that $a(1) = 1$, $a(n) = 0$ for $1 < n \leq X$, $|a(n)| \leq \tau_X(n)$, and $L(s, \chi) M_X(s, \chi) = \sum_n a(n) \chi(n) n^{-s}$ for $\operatorname{Re} s > 1$.

Lemma A.7. *Let χ be nonprincipal modulo k , $X, Y \geq 1$, $0 < \sigma_1 < 1$, and let $\rho = \beta + i\gamma$ be a zero of $L(s, \chi)$ with $\sigma_1 < \beta \leq 1$. Then*

$$\sum_{n \geq 1} a(n) \chi(n) n^{-\rho} \Phi(n/Y) = \frac{1}{2\pi i} \int_{\sigma_1 - \beta - i\infty}^{\sigma_1 - \beta + i\infty} L(\rho + w, \chi) M_X(\rho + w, \chi) Y^w e^{w^2} \frac{dw}{w},$$

and the series on the left converges absolutely.

Proof. Take $c = 2 - \beta > 0$ in the definition of $\Phi(n/Y)$ and sum over n ; the double integral converges absolutely because $\operatorname{Re}(\rho + w) = 2$ and $|a(n)| \leq d(n) \leq 2\sqrt{n}$, so the sum and the integral may be exchanged and the left side equals $(2\pi i)^{-1} \int_{\operatorname{Re} w = 2 - \beta} L(\rho + w, \chi) M_X(\rho + w, \chi) Y^w e^{w^2} dw/w$. The integrand is holomorphic in the strip $\sigma_1 - \beta \leq \operatorname{Re} w \leq 2 - \beta$ except for the simple pole at $w = 0$, whose residue $L(\rho, \chi) M_X(\rho, \chi)$ vanishes; on horizontal segments $\Im w = \pm V$ it is bounded by $e^{4 - V^2}$ times a polynomial in V , by Lemma 3.2 and the trivial bound for M_X , so the line may be moved to $\operatorname{Re} w = \sigma_1 - \beta$. $\square$

A.2. The proof. Let χ be nonprincipal modulo k and let $\mathcal{Z}$ be a set of zeros $\rho = \beta + i\gamma$ of $L(s, \chi)$ with $\beta \geq \sigma$ and $|\gamma| \leq T$ whose ordinates are pairwise at distance at least 1; we call such a set *1-spaced*.

Step 1: reduction to 1-spaced sets. For $j \in \mathbb{Z}$ let $I_j = (j - 1/2, j + 1/2]$. If I_j contains an ordinate of a zero counted by $N(\sigma, T, \chi)$, choose one such zero ρ_j ; the chosen zeros with j even form a 1-spaced set, and so do those with j odd, since ordinates in I_j and I_{j+2} differ by more than 1. By Lemma A.2, applied with $t = j$ and $\beta \geq \sigma \geq 3/4$, each I_j with $|j| \leq T + 1/2$ holds at most $7 \log(8k(T + 5))$ zeros. Hence

$$N(\sigma, T, \chi) \leq 14 \log(8k(T + 5)) \max_{\mathcal{Z}} |\mathcal{Z}| \leq 42\mathcal{L} \max_{\mathcal{Z}} |\mathcal{Z}|, \quad (\text{A.2})$$

the maximum over 1-spaced sets, because $8(T + 5) \leq 28T$ and $\log 28 < 1.87\mathcal{L}$.

Step 2: the two classes. Let $\mathcal{Z}$ be 1-spaced and $\rho \in \mathcal{Z}$. By Lemma A.7 with $\sigma_1 = 5/6$ (note $\beta \geq \sigma > \sigma_1$),

$$\Phi(1/Y) + \sum_{X < n \leq Y^2} a(n) \chi(n) n^{-\rho} \Phi(n/Y) + \sum_{n > Y^2} a(n) \chi(n) n^{-\rho} \Phi(n/Y) = I(\rho), \quad (\text{A.3})$$

where $I(\rho)$ is the integral of that lemma. By Lemma A.6(ii) and $Y \geq 216$, $|\Phi(1/Y)| \geq 1 - e^{-(\log Y)^2/4}/(\sqrt{\pi} \log Y) \geq 1/2$. For $n > Y^2$ we have $\log(n/Y) \geq \log Y$, so Lemma A.6(i) gives $|\Phi(n/Y)| \leq (Y/n)^{(\log Y)/4}/(\sqrt{\pi} \log Y)$, and with $|a(n)| \leq 2\sqrt{n}$, $\beta \geq 11/12$, the third sum in (A.3) has modulus at most

$$\frac{2}{\sqrt{\pi} \log Y} Y^{(\log Y)/4} \int_{Y^2}^{\infty} u^{-5/12 - (\log Y)/4} du = \frac{2}{\sqrt{\pi} \log Y} \cdot \frac{Y^{7/6 - (\log Y)/4}}{(\log Y)/4 - 7/12} \leq \frac{1}{8},$$

the last step because the middle expression decreases in Y and is below 0.11 at $Y = 216$.⁷ Consequently every $\rho \in \mathcal{Z}$ is of *class I*, $|\sum_{X < n \leq Y^2} a(n) \chi(n) n^{-\rho} \Phi(n/Y)| \geq 1/4$, or of *class II*, $|I(\rho)| \geq 1/8$.

⁷The summand decreases in n , so the sum over $n > Y^2$ is at most the integral from Y^2 .

Step 3: class I. Split $(X, Y^2]$ into the blocks $(U, 2U]$ with $U = 2^j X$, $j \geq 0$, the last block cut at Y^2 ; there are at most $2 \log(Y^2)/\log 2 + 1 \leq 4 \log Y$ blocks. For a class I zero some block has $|D_U(\rho)| \geq 1/(16 \log Y)$, where

$$D_U(s) = \sum_{U < n \leq \min(2U, Y^2)} a(n) \Phi(n/Y) \chi(n) n^{-s},$$

a Dirichlet polynomial whose coefficients do not depend on ρ . Fix U and put $r = 1/(2 \log(2U))$, so that $r \leq 1/4$. The discs $|s - \rho| \leq r$, $\rho \in \mathcal{Z}$, are disjoint and lie in the rectangle $\sigma - r \leq \operatorname{Re} s \leq 1 + r$, $|\Im s| \leq T'$ with $T' = T + 1/2$, so Lemma A.5 gives

$$\sum_{\rho \in \mathcal{Z}} |D_U(\rho)|^2 \leq \frac{4(\log 2U)^2}{\pi} \int_{\sigma-r}^{1+r} \int_{-T'}^{T'} |D_U(x+it)|^2 dt dx.$$

For fixed x in that range, $n^{-2x} \leq n^{-2\sigma}(2U)^{2r} = e n^{-2\sigma} \leq eU^{-2\sigma}$ for $n \leq 2U$, so by Lemmas A.3 and A.4 and $|\Phi| \leq 2$,

$$\sum_{\chi} \int_{-T'}^{T'} |D_U(x+it)|^2 dt \leq (86U + 22kT') \cdot 4e U^{-2\sigma} \cdot 2U(1 + \log X)^3.$$

The x -range has length $1 - \sigma + 2r \leq 7/12$. With $T' \leq 5T/4$ and $U \geq X = kT$ we have $(86U + 22kT')U^{1-2\sigma} \leq 114 \max(kT X^{1-2\sigma}, U^{2-2\sigma}) \leq 114(kT)^{12(1-\sigma)}$, because $2 - 2\sigma \leq 12(1 - \sigma)$ and $U \leq Y^2 = (kT)^6$. Also $\log(2U) \leq \log(2Y^2) \leq 6.4\mathcal{L}$ and $1 + \log X \leq 1.56\mathcal{L}$. Therefore

$$\sum_{\chi} \sum_{\rho \in \mathcal{Z}_{\chi}} |D_U(\rho)|^2 \leq \frac{4(6.4\mathcal{L})^2}{\pi} \cdot \frac{7}{12} \cdot 4e \cdot 2 \cdot 114 (1.56\mathcal{L})^3 (kT)^{12(1-\sigma)} \leq 3 \cdot 10^5 \mathcal{L}^5 (kT)^{12(1-\sigma)}.$$

The class I zeros attached to the block U number at most $(16 \log Y)^2 = (48\mathcal{L})^2$ times this, and there are at most $4 \log Y = 12\mathcal{L}$ blocks, so

$$\sum_{\chi} \#\{\rho \in \mathcal{Z}_{\chi} \text{ of class I}\} \leq 10^{10} \mathcal{L}^8 (kT)^{12(1-\sigma)}. \quad (\text{A.4})$$

Step 4: class II. Write $F(s) = L(s, \chi) M_X(s, \chi)$. For ρ of class II, with $w = \sigma_1 - \beta + iv$, we have $-1/6 \leq \operatorname{Re} w \leq -1/12$, so $|e^{w^2}| \leq e^{1/36} e^{-v^2}$, $|w|^{-1} \leq 12$ and $Y^{\operatorname{Re} w} \leq Y^{\sigma_1 - \sigma}$; thus

$$\begin{aligned} \frac{1}{8} \leq |I(\rho)| &\leq \frac{12e^{1/36}}{2\pi} Y^{\sigma_1 - \sigma} \int_{-\infty}^{\infty} e^{-v^2} |F(\sigma_1 + i\gamma + iv)| dv \\ &\leq 2\pi^{1/4} Y^{\sigma_1 - \sigma} \left(\int_{-\infty}^{\infty} e^{-v^2} |F(\sigma_1 + i\gamma + iv)|^2 dv \right)^{1/2} \end{aligned}$$

by the Cauchy–Schwarz inequality. Squaring and summing over the class II zeros of $\mathcal{Z}$,

$$\#\{\rho \in \mathcal{Z} \text{ of class II}\} \leq 454 Y^{2(\sigma_1 - \sigma)} \int_{-\infty}^{\infty} |F(\sigma_1 + it)|^2 w(t) dt, \quad w(t) = \sum_{\rho \in \mathcal{Z}} e^{-(t-\gamma)^2}.$$

Since the ordinates are 1-spaced, at most one of them lies in each interval $[t+j-1/2, t+j+1/2)$, so $w(t) \leq 1 + 2 \sum_{j \geq 1} e^{-(j-1/2)^2} < 3$. Let $T_2 = T + 4\mathcal{L}$. For $|t| > T_2$ the distances $|t - \gamma|$ exceed $u = |t| - T \geq 4\mathcal{L}$ and at most one lies in each $[u+j, u+j+1)$, so $w(t) \leq \sum_{j \geq 0} e^{-(u+j)^2} \leq 1.4e^{-u^2}$; and $|F(\sigma_1 + it)| \leq \frac{6}{5}k(|t| + 1) \cdot 6X^{1/6}$ by Lemma 3.2 and $\sum_{d \leq X} d^{-5/6} \leq 6X^{1/6}$. Hence, using $(u + T + 1)^2 \leq 2u^2 + 8e^{2\mathcal{L}}$, $\int_a^{\infty} u^2 e^{-u^2} du \leq ae^{-a^2}$ and $\int_a^{\infty} e^{-u^2} du \leq e^{-a^2}/(2a)$ for $a \geq 1$,

$$\int_{|t| > T_2} |F(\sigma_1 + it)|^2 w(t) dt \leq 2 \cdot 52k^2 X^{1/3} \cdot 1.4(8\mathcal{L} + e^{2\mathcal{L}})e^{-16\mathcal{L}^2} \leq 292 e^{13\mathcal{L}/3 - 16\mathcal{L}^2} \leq 1.$$

On $|t| \leq T_2$ apply Lemma A.1 with $N = \lceil (kT_2)^2 \rceil$: $|E_N(\sigma_1 + it, \chi)| \leq k(1 + \frac{6}{5}(T_2 + 1))(kT_2)^{-5/3} \leq 2.3(kT_2)^{-2/3} \leq 1$, as $kT_2 \geq 27$. So $|F|^2 \leq 2|P_N M_X|^2 + 2|M_X|^2$ with $P_N(s) = \sum_{n \leq N} \chi(n) n^{-s}$. By Lemma A.3, $\sum_{\chi} \int_{-T_2}^{T_2} |M_X(\sigma_1 + it, \chi)|^2 dt \leq (43X + 22kT_2) \sum_{d \leq X} d^{-5/3} \leq$

$163kT_2$. The product $P_N M_X$ is $\sum_{m \leq NX} b_m \chi(m) m^{-s}$ with $|b_m| \leq \tau_X(m)$; cut it into the blocks $m \in (U, 2U]$, $U = 2^j$, $0 \leq j \leq J$, $2^J \leq NX \leq 2(kT_2)^3$, so that $J+1 \leq 11\mathcal{L}$ (we use $\mathcal{L}_2 = \log(kT_2) \leq 2\mathcal{L}$, since $T_2 \leq 3\mathcal{L}T$ and $\log(3\mathcal{L}) \leq \mathcal{L}$). By the Cauchy–Schwarz inequality $|P_N M_X|^2 \leq (J+1) \sum_j |B_j|^2$, and for each block, by Lemmas A.3 and A.4,

$$\begin{aligned} \sum_{\chi} \int_{-T_2}^{T_2} |B_j(\sigma_1 + it)|^2 dt &\leq (86U + 22kT_2) U^{-5/3} \cdot 2U(1 + \log X)^3 \\ &\leq 2(1 + \log X)^3 (86 \cdot 1.26 + 22)kT_2 \leq 10^3 \mathcal{L}^3 kT_2, \end{aligned}$$

using $U^{1/3} \leq (2(kT_2)^3)^{1/3} \leq 1.26kT_2$. Collecting,

$$\sum_{\chi} \int_{-T_2}^{T_2} |F(\sigma_1 + it)|^2 dt \leq 2(11\mathcal{L})^2 \cdot 10^3 \mathcal{L}^3 kT_2 + 2 \cdot 163kT_2 \leq 2.5 \cdot 10^5 \mathcal{L}^5 kT_2,$$

and with $kT_2 \leq 3\mathcal{L}kT$ and $Y^{2(\sigma_1 - \sigma)} = (kT)^{5-6\sigma}$,

$$\sum_{\chi} \#\{\rho \in \mathcal{Z}_{\chi} \text{ of class II}\} \leq 454(3 \cdot 2.5 \cdot 10^5 \mathcal{L}^5 \cdot 3\mathcal{L}kT + k)(kT)^{5-6\sigma} \leq 1.1 \cdot 10^9 \mathcal{L}^6 (kT)^{6(1-\sigma)}. \quad (\text{A.5})$$

Step 5: conclusion. By (A.2), (A.4) and (A.5), choosing for each χ a 1-spaced set of largest size,

$$\sum_{\chi \neq \chi_0} N(\sigma, T, \chi) \leq 42\mathcal{L}(10^{10}\mathcal{L}^8 + 1.1 \cdot 10^9 \mathcal{L}^6)(kT)^{12(1-\sigma)} \leq 10^{12} (kT)^{12(1-\sigma)} (\log kT)^9,$$

which is Theorem 2.3.

APPENDIX B. THE ZERO-FREE HALF-PLANE $\operatorname{Re} s > 11/12$

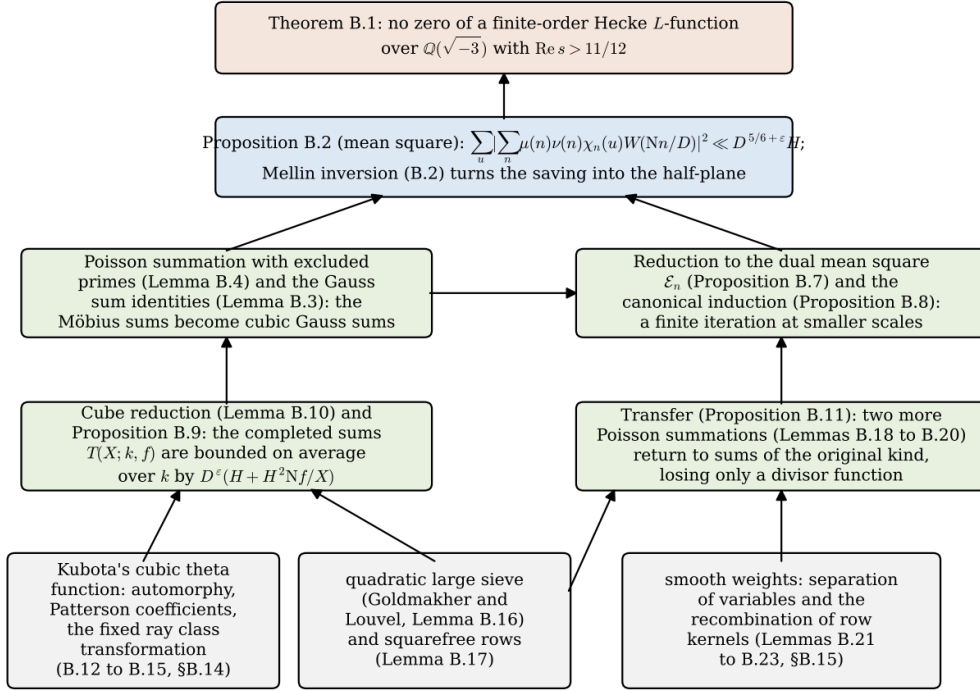
This appendix proves the zero-free half-plane that Theorem 1.5 uses, so that the paper is complete in itself. The theorem and its proof are due to OpenAI [24]: the argument below is their argument, reproduced step by step in our notation, with the published results it rests on cited at the places where they enter and with the computations that the preprint leaves to the reader written out.⁸ We claim no originality for anything in this appendix.

Theorem B.1 (OpenAI). *Let $K = \mathbb{Q}(\sqrt{-3})$. Every finite-order Hecke L -function over K has no zero in the half-plane $\operatorname{Re} s > 11/12$. In particular every Dirichlet L -function, the Riemann zeta function included, has no zero with $\operatorname{Re} s > 11/12$.*

The strategy is the following. A zero of $L_K(s, \nu)$ with $\operatorname{Re} s > 11/12$ would prevent a power saving $D^{11/12+\varepsilon}$ in the smoothed Möbius sums $\sum \mu(\mathbf{n}) \nu(\mathbf{n}) W(\mathbf{N}\mathbf{n}/D)$ of the field, by a Mellin transform argument of Littlewood’s kind. That power saving is extracted from a mean square over a family of such sums twisted by sextic residue symbols $\chi_{\mathbf{n}}(u)$, $u \in \mathcal{O}_K$, because the twist by $u = p^6$ differs from the untwisted sum by $O(D/Np)$ only. The mean square over u is opened by Poisson summation; the Gauss sums that appear absorb the Möbius function and become cubic Gauss sums, which are Fourier coefficients of Kubota’s cubic theta function. The automorphy of that theta function turns the sextic twists into quadratic ones, to which the quadratic large sieve of Goldmakher and Louvel applies; the cube factors that the theta function carries are removed by Möbius inversion, and the sums that this leaves are of the

⁸Two preprints of OpenAI prove such a half-plane. The first [23] proves $\operatorname{Re} s > 7/8$ and is written in Lean 4; its kernel check is the one recorded in the data availability section. The second [24] isolates the weaker half-plane $\operatorname{Re} s > 11/12$, which is an intermediate step of the first, and gives it a proof of ordinary length; it is this second proof that we reproduce. The transfer theorem gives Theorem 1.5 from either half-plane, with a different constant, so the choice of the shorter proof costs nothing but the constant.

original kind at smaller scales, so that a finite iteration closes the argument.⁹ Figure 4 shows the parts.



The argument of OpenAI, Appendix B; the arrows point from a result to the result that uses it.

FIGURE 4. The structure of the proof of Theorem B.1. The half-plane follows from the mean-square estimate of Proposition B.2, which is proved by Poisson summation, the canonical induction of Proposition B.8 and the transfer of Proposition B.11; the cubic theta function and the quadratic large sieve enter through Proposition B.9.

B.1. Notation. Write $\mathcal{O} = \mathcal{O}_K = \mathbb{Z}[\omega]$ with $\omega = e^{2\pi i/3}$, and $N(a) = a\bar{a} = |a|^2$ for $a \in K$, $N(\mathfrak{a}) = \#(\mathcal{O}/\mathfrak{a})$ for a nonzero ideal. The ring $\mathcal{O}$ is Euclidean, so every ideal is principal. An element is *primary* if it is congruent to 1 modulo 3; since the six units represent the six invertible classes modulo 3, an ideal prime to 3 has a unique primary generator, and we write $\mu(n)$, $\nu(n)$ and so on for the values at the ideal (n) . Put $\lambda = 1 + 2\omega$, so that (λ) is the prime above 3 and λ^2 is a unit times 3. For a primary prime $p \nmid 6$ and $p \nmid u$, $(u/p)_6$ is the sixth root of unity congruent to $u^{(Np-1)/6}$ modulo p , and $(u/p)_6 = 0$ for $p \mid u$; this is extended multiplicatively in the denominator to

$$\chi_n(u) = \left(\frac{u}{n}\right)_6 \quad (n \text{ primary, } (n, 6) = 1),$$

with $\chi_1 \equiv 1$, including at $u = 0$. The quadratic and cubic symbols are χ_n^3 and χ_n^2 .¹⁰ Define the additive character

$$e(z) = \exp(4\pi i \operatorname{Im} z / \sqrt{3}) \quad (z \in \mathbb{C}),$$

⁹The sketch is that of [24, §2]. The role of the cubic theta function goes back to the work of Heath-Brown and Patterson on Kummer sums [14], and the completion by cubes and its inversion to Dunn and Radziwiłł [8]; the iteration is the admissible-exponent method of Heath-Brown's quadratic large sieve [12].

¹⁰For j an integer, χ_p^j means the j th power on $(\mathcal{O}/(p))^\times$ extended by zero on the multiples of p ; in particular $\chi_p^0(x) = \mathbf{1}_{p \nmid x}$, which is not the constant 1. This convention matters in the theta transformation of Section B.14.

which is trivial on $\mathcal{O}$ and, for primary squarefree n prime to 6 and an integer j , the normalised Gauss sums

$$\gamma_j(n) = \frac{1}{\sqrt{\mathbb{N}n}} \sum_{x \bmod n} \chi_n(x)^j e(x/n), \quad \alpha(n) = \frac{n}{|n|}. \quad (\text{B.1})$$

For $j = -1$ the character is $\bar{\chi}_n$. We use $\mathbf{1}_C$ for the indicator of a condition, v_p for the exponent of a prime, $\tau_{\text{div}}(b)$ for the number of ideal divisors of (b) , and $\text{rad } \mathfrak{r}$ for the product of the distinct prime divisors of $\mathfrak{r}$.

The notation $f \ll g$ means $|f| \leq Cg$ with an absolute constant unless subscripts list the parameters the constant may depend on; $f \asymp g$ means $f \ll g \ll f$. A dyadic range is $R \leq Na < 2R$. Throughout, $D \geq 2$ is the norm scale of the original Möbius sum and all other scales are bounded by fixed powers of D ; we write

$$A \preccurlyeq B \quad \text{if} \quad A \ll_{\varepsilon} D^{\varepsilon} B \text{ for every } \varepsilon > 0,$$

uniformly in the varying parameters, the constant depending also on the fixed data of each statement. For a compact interval $I \subset (0, \infty)$ and a smooth W supported in I , $\|W\|_{C^J(I)} = \max_{j \leq J} \sup_I |W^{(j)}|$; for functions of several variables all partial derivatives of total order at most J are taken. A star on a sum restricts the index to squarefree primary elements, prime to the finite set S of primes fixed in the next section.

Finally, the *ray class group* modulo an ideal $\mathfrak{m}$ is the group of fractional ideals prime to $\mathfrak{m}$ modulo the principal ideals with a generator congruent to 1 modulo $\mathfrak{m}$; a ray class character is a character of this finite group, and the ray class of an element prime to $\mathfrak{m}$ is the class of the ideal it generates.

B.2. From the mean-square estimate to the zero-free half-plane. Fix a finite-order Hecke character ν of K , extended by zero to the ideals not prime to its conductor, and a finite set S of prime ideals containing the primes above 2 and 3 and the primes dividing the conductor of ν . Write $(a, S) = 1$ if no prime of S divides a . For $W \in C_c^{\infty}((0, \infty); \mathbb{C})$ and $u \in \mathcal{O}$ define the family

$$A_u(D) = \sum_{(n, S)=1} \mu(n) \nu(n) \chi_n(u) W(\mathbb{N}n/D), \quad (\text{B.2})$$

where n runs over the ideals prime to S , represented by their primary generators. The untwisted sum is $A_1(D)$. The whole appendix is devoted to the following estimate.

Proposition B.2. *For every fixed $0 < \vartheta \leq 1/10$ and $\varepsilon > 0$ there is an integer $k = k(\vartheta, \varepsilon) \geq 1$ such that for every compact interval $I \subset (0, \infty)$, every smooth W supported in I and every $D \geq 2$,*

$$\sum_{0 < \mathbb{N}u \leq D^{1+\vartheta}} |A_u(D)|^2 \ll_{\nu, S, I, \vartheta, \varepsilon} \left(\max_{0 \leq j \leq k} \|W^{(j)}\|_{\infty} \right)^2 D^{2+\vartheta+\varepsilon}. \quad (\text{B.3})$$

The proposition is proved in Section B.9. We first deduce the theorem from it.

Proof of Theorem B.1 from Proposition B.2. Fix $0 < \vartheta \leq 1/10$ and put $H = D^{1+\vartheta}$, $Y = H^{1/6}$. For a prime ideal $\mathfrak{p} \notin S$ with $Y/2 < \mathbb{N}\mathfrak{p} \leq Y$ let p be its primary generator. Then $\chi_n(p^6) = \mathbf{1}_{\mathfrak{p} \nmid n}$, so $A_1(D)$ and $A_{p^6}(D)$ differ only in the terms with $\mathfrak{p} \mid n$:

$$|A_1(D) - A_{p^6}(D)| \leq \|W\|_{\infty} \#\{n : (n, S) = 1, \mathfrak{p} \mid n, \mathbb{N}n \asymp D\} \ll_W D/Y.$$

By Landau's prime ideal theorem [20] there are $J \asymp Y/\log Y$ such primes, and their sixth powers are distinct rows u of norm at most H .¹¹ Apply (B.3) with the loss $\varepsilon/2$ and use

¹¹In a mean square we call the variable of the outer sum the *row* and the variable of the inner sum the *column*; in (B.3) these are u and n .

$\log Y \ll_\varepsilon D^{\varepsilon/2}$:

$$\begin{aligned} |A_1(D)|^2 &\leq \frac{2}{J} \sum_{\substack{Y/2 < Np \leq Y \\ (p,S)=1}} |A_{p^6}(D)|^2 + O_W\left(\frac{D^2}{Y^2}\right) \\ &\ll_{\nu,S,W,\vartheta,\varepsilon} \frac{D^{2+\vartheta+\varepsilon}}{Y} + \frac{D^2}{Y^2} = D^{11/6+5\vartheta/6+\varepsilon} + D^{5/3-\vartheta/3}. \end{aligned} \quad (\text{B.4})$$

Hence $A_1(D) \ll D^{11/12+5\vartheta/12+\varepsilon}$. For a requested exponent loss choose ϑ and then ε small, and rename the total loss ε :

$$A_1(D) \ll_{\nu,S,W,\varepsilon} D^{11/12+\varepsilon}. \quad (\text{B.5})$$

Suppose now that $L_K(s, \nu)$ has a zero ϱ with $\text{Re } \varrho > 11/12$. Take $0 \neq \phi \in C_c^\infty((1, 2))$, $\phi \geq 0$, and $W(y) = y^{-\varrho} \phi(y)$. With $\widehat{W}(s) = \int_0^\infty W(y) y^s dy/y$ we have $\widehat{W}(\varrho) = \int \phi(y) dy/y > 0$. The function

$$\mathcal{M}_W(s) = \int_0^\infty A_1(D) D^{-s} \frac{dD}{D}$$

is holomorphic for $\text{Re } s > 11/12$, because (B.5) controls the integral at infinity and $A_1(D)$ vanishes for small D by the compact support of W . For $\text{Re } s > 1$ termwise integration gives

$$\mathcal{M}_W(s) = \widehat{W}(s) \sum_{(n,S)=1} \frac{\mu(n)\nu(n)}{Nn^s} = \frac{\widehat{W}(s)}{L_K^S(s, \nu)}, \quad L_K^S(s, \nu) = L_K(s, \nu) \prod_{\mathfrak{p} \in S} (1 - \nu(\mathfrak{p})N\mathfrak{p}^{-s}).$$

By Hecke's continuation theorem [15] and the identity theorem, $L_K^S(s, \nu)\mathcal{M}_W(s) = \widehat{W}(s)$ on $\text{Re } s > 11/12$. The omitted Euler factors do not vanish there, so at $s = \varrho$ the left side is 0 and the right side is $\widehat{W}(\varrho) > 0$, a contradiction.

For a Dirichlet character χ , let χ_{-3} be the nontrivial character modulo 3. Quadratic base change gives, up to Euler factors without zeros in $\text{Re } s > 0$,

$$L_K(s, \chi \circ N) = L(s, \chi)L(s, \chi\chi_{-3}),$$

so the Hecke statement excludes zeros of either factor off $s = 1$; at $s = 1$ a zero of one factor could only cancel a pole of the other when χ is principal, and then it would be a zero of $L(1, \chi_{-3})$, which is positive by Dirichlet's theorem [6, Chapters 4 and 6]. $\square$

B.3. The arithmetic identities. Poisson summation in the row variable will produce the Gauss sums (B.1), and the identities of this section convert the Möbius coefficients into the cubic coefficients

$$a_\xi(n) = \overline{\alpha(n)} \gamma_2(n) \xi(n) \quad (n \text{ squarefree primary, } (n, S) = 1), \quad (\text{B.6})$$

where ξ is a ray class character of a fixed modulus. That modulus is chosen once and for all: supported on S , divisible by the conductor of ν , and large enough that every reciprocity factor below depends only on the ray class; a function on the ray class group is then a finite linear combination of characters ξ by orthogonality, which is how ν and the factors G , $\mathcal{R}$ below are absorbed into the twist ξ .

Lemma B.3. *On the fixed ray class group there are a function G with values in the unit circle and a symmetric $\{\pm 1\}$ -valued bicharacter $\mathcal{R}$, that is, $\mathcal{R}(a, b) = \mathcal{R}(b, a)$ and $\mathcal{R}$ is multiplicative in each argument, with the following properties. For primary a, b, n prime to S with $(a, b) = 1$ and n squarefree,*

$$\chi_b(a) = \mathcal{R}(a, b) \chi_a(b), \quad G(ab) = G(a)G(b)\mathcal{R}(a, b), \quad (\text{B.7})$$

$$\gamma_2(n)^3 = \mu(n)\alpha(n), \quad \gamma_1(n)\gamma_2(n) = \mu(n)\alpha(n)G(n), \quad (\text{B.8})$$

$$G(n) = \overline{\chi_n(4)} \gamma_3(n), \quad \gamma_1(n)\gamma_{-1}(n) = \chi_n(-1). \quad (\text{B.9})$$

Consequently, for such n and for coprime squarefree a, b ,

$$\overline{\alpha(n)}\gamma_2(n)\gamma_1(n) = \mu(n)G(n), \quad (\text{B.10})$$

$$\mu(n)\gamma_{-1}(n) = \chi_n(-1)G(n)^{-1}\overline{\alpha(n)}\gamma_2(n), \quad (\text{B.11})$$

$$a_\xi(ab) = a_\xi(a)a_\xi(b)\chi_b(a)^4, \quad (\text{B.12})$$

$$\chi_a(-1)\overline{G(a)}G(b)\mathcal{R}(a, b) = G(ba^{-1}), \quad (\text{B.13})$$

where the identity for $G(ab)$ extends to all classes of the ray class group and the quotient in (B.13) is taken there.

Proof. The Gauss–Jacobi relations (B.8). Set $G(n) = \overline{\chi_n(4)}\gamma_3(n)$ for squarefree n ; that it factors through the ray class group is shown below. Fix a primary prime $p \notin S$ and, for characters A, B of $(\mathcal{O}/(p))^\times$ extended by zero, the Jacobi sum $J(A, B) = \sum_{x \bmod p} A(x)B(1-x)$. The substitution $u = 2x - 1$ shows that $4x(1-x) = y$ has $1 + \chi_p^3(1-y)$ solutions x modulo p , so

$$\chi_p(4)J(\chi_p, \chi_p) = \sum_x \chi_p(4x(1-x)) = \sum_y \chi_p(y)(1 + \chi_p^3(1-y)) = J(\chi_p, \chi_p^3).$$

The Gauss–Jacobi identity $\gamma_1\gamma_3/\gamma_4 = J(\chi_p, \chi_p^3)/\sqrt{\mathbb{N}p}$ (see [17, (3.18)]) and the same identity for $J(\chi_p, \chi_p)$ give

$$\frac{\gamma_1(p)\gamma_3(p)}{\gamma_4(p)} = \chi_p(4)\frac{\gamma_1(p)^2}{\gamma_2(p)}.$$

For $j \not\equiv 0 \pmod{6}$ one has $\gamma_j(p)\gamma_{-j}(p) = \chi_p(-1)^j$; with $j = 2$ this is $\gamma_2(p)\gamma_4(p) = 1$, as χ_p^2 is cubic. Cancelling $\gamma_1(p)$ and rearranging,

$$\gamma_1(p)\gamma_2(p) = \overline{\chi_p(4)}\gamma_3(p)\gamma_2(p)^3.$$

It remains to evaluate $\gamma_2(p)^3 = \gamma_2(p)^2/\gamma_4(p) = J(\chi_p^2, \chi_p^2)/\sqrt{\mathbb{N}p}$. Put $q = \mathbb{N}p$. The Jacobi sum lies in $\mathcal{O}$ and has norm q . Reducing modulo p and expanding the binomial,

$$J(\chi_p^2, \chi_p^2) \equiv \sum_x x^{\frac{q-1}{3}}(1-x)^{\frac{q-1}{3}} = \sum_{\ell=0}^{(q-1)/3} (-1)^\ell \binom{(q-1)/3}{\ell} \sum_x x^{\frac{q-1}{3}+\ell} \equiv 0 \pmod{p},$$

since each exponent $(q-1)/3 + \ell$ lies strictly between 0 and $q-1$. So $J(\chi_p^2, \chi_p^2)$ is a unit times p . Write $\chi_p^2(x(1-x)) = \omega^{j_x}$ for $x \neq 0, 1$; the product of all $x(1-x)$ is 1, so $\sum j_x \equiv 0 \pmod{3}$, and since $(\omega-1)^2$ generates (3) ,

$$J(\chi_p^2, \chi_p^2) = \sum_{x \neq 0,1} \omega^{j_x} \equiv (q-2) + (\omega-1) \sum_{x \neq 0,1} j_x \equiv -1 \pmod{3}.$$

With $p \equiv 1 \pmod{3}$ this fixes the unit: $J(\chi_p^2, \chi_p^2) = -p$, so $\gamma_2(p)^3 = -\alpha(p)$. The Chinese remainder theorem extends both prime identities to squarefree n , one sign -1 for each prime factor giving $\mu(n)$.¹²

The identity (B.12). For coprime squarefree primary a, b prime to S , $\gamma_2(ab) = \gamma_2(a)\gamma_2(b)\chi_a(b)^2\chi_b(a)^2$, and cubic reciprocity gives $\chi_a(b)^2 = \chi_b(a)^2$; multiplying by $\alpha(ab)\xi(ab)$ gives (B.12).

Reciprocity (B.7) and the first identity of (B.9). For $c \in \mathcal{O}$ prime to 2 we show

$$\Gamma_{\text{quad}}(c) := |c|^{-1} \sum_{x \bmod c} e(x^2/c) = \frac{1}{2} \sum_{y \bmod 2\mathcal{O}} e(-cy^2/4). \quad (\text{B.14})$$

¹²The Chinese remainder theorem gives $\gamma_j(ab) = \gamma_j(a)\gamma_j(b)\chi_a(b)^j\chi_b(a)^j$ for coprime a, b ; for $j = 2$ cubic reciprocity makes the two symbols equal, and for the products of two sums with $j + j' \equiv 0$ the symbols cancel.

Apply Poisson summation over $z \in \mathcal{O}$ to $e(z^2/c)e^{-\pi\eta|z|^2}$, $\eta > 0$. The Fourier transform of this product is

$$\frac{|c|}{2\sqrt{r_\eta}} e^{-\pi\eta N(c)|y|^2/(4r_\eta)} e(-cy^2/(4r_\eta)), \quad r_\eta = 1 + 3N(c)\eta^2/16.$$

Divide both sides by the Gaussian mass $\sum_z e^{-\pi\eta|z|^2} \asymp \eta^{-1}$. Replacing r_η by 1 on the Fourier side changes it by $O_c(\eta^2 \sum_y |y|^2 e^{-C_c\eta|y|^2}) = O_c(1)$, which is $O_c(\eta)$ after the division. Group the original sum modulo c and the Fourier sum modulo $2\mathcal{O}$ and let $\eta \downarrow 0$ to get (B.14). For $c = a + b\omega$ it reads

$$\Gamma_{\text{quad}}(a + b\omega) = \frac{1}{2}(1 + i^{-b} + i^a + i^{b-a}),$$

so $\Gamma_{\text{quad}}(c)$ depends only on c modulo $4\mathcal{O}$ and is invariant under multiplication by squares of $(\mathcal{O}/4\mathcal{O})^\times$; on the four square classes $1, -1, \lambda, -\lambda$ it takes the values $1, 1, i, -i$. Hence

$$\mathcal{R}(c_1, c_2) := \frac{\Gamma_{\text{quad}}(c_1 c_2)}{\Gamma_{\text{quad}}(c_1) \Gamma_{\text{quad}}(c_2)},$$

$$\mathcal{R}((-1)^e \lambda^f, (-1)^g \lambda^h) = (-1)^{eh+fg+fh} \quad (e, f, g, h \in \{0, 1\}),$$

is a symmetric bicharacter on these classes. For a prime p every y has $1 + \chi_p^3(y)$ square roots, so grouping by $y = x^2$ and using $\sum_y e(y/p) = 0$ gives $\Gamma_{\text{quad}}(p) = \gamma_3(p)$, and the Chinese remainder theorem extends this to squarefree n . With cubic reciprocity this identifies $\mathcal{R}$ with the quotient $\chi_b(a)/\chi_a(b)$ for coprime primary a, b . Since $\chi_c(4) = (-2/c)_3 = (c/(-2))_3$ is a character modulo 2, the function $G(c) = \overline{\chi_c(4)} \Gamma_{\text{quad}}(c)$ factors through the fixed ray class group and satisfies (B.7), with $\mathcal{R}(c, c) = \chi_c(-1)$.

The quotient (B.13). In the ray class group the multiplicativity of G gives $G(a^{-1}) = \chi_a(-1) \overline{G(a)}$, and then $G(ba^{-1}) = G(b)G(a^{-1})\mathcal{R}(b, a^{-1}) = \chi_a(-1) \overline{G(a)} G(b) \mathcal{R}(a, b)$ by the symmetry and multiplicativity of $\mathcal{R}$.

The conversions. The second identity of (B.9) is the inverse-character Gauss identity. Multiplying the second identity of (B.8) by $\overline{\alpha(n)}$ gives (B.10), and combining it with $\gamma_1 \gamma_{-1} = \chi_n(-1)$ gives (B.11). $\square$

Under Poisson summation, (B.11) converts the Möbius coefficients into $a_\xi(n)$, up to fixed ray class factors, and (B.10) converts them back.

B.4. Poisson summation with excluded primes. The row sums to which Poisson summation is applied carry a coprimality condition, and we record the formula with that condition included, so that its effect on the frequencies and on the normalisation is visible. In quotients and Gauss sums every ideal is represented by a fixed generator, primary when the ideal is prime to 3.

Lemma B.4. *Let $\mathcal{H} > 0$, let $\mathfrak{r}$ be a nonzero ideal, let $\Phi(Nk/\mathcal{H})$ be a smooth radial Schwartz weight on $\mathcal{O}$, and let χ be a primitive character of $(\mathcal{O}/\mathfrak{m})^\times$, viewed as a function on $\mathcal{O}$ by reduction modulo $\mathfrak{m}$ and extension by zero. Write $\gamma(\chi) = N\mathfrak{m}^{-1/2} \sum_{x \bmod \mathfrak{m}} \chi(x) e(x/\mathfrak{m})$ for its normalised Gauss sum. Then*

$$\sum_{k \in \mathcal{O}} \chi(k) \mathbf{1}_{(k, \mathfrak{r})=1} \Phi(Nk/\mathcal{H}) = \frac{\mathcal{H}\gamma(\chi)}{\sqrt{N\mathfrak{m}}} \sum_{d \mid \text{rad } \mathfrak{r}} \frac{\mu(d)\chi(d)}{Nd} \sum_{h \in \mathcal{O}} \overline{\chi(h)} \widehat{\Phi}\left(\frac{\mathcal{H}Nh}{NdN\mathfrak{m}}\right), \quad (\text{B.15})$$

where $\widehat{\Phi}(|w|^2) = \frac{2}{\sqrt{3}} \int_{\mathbb{C}} \Phi(|z|^2) e(-zw) dx dy$, the measure giving $\mathcal{O}$ covolume one. If χ is nonprincipal then $\chi(0) = 0$ and the zero frequency contributes nothing; for the principal primitive character ($\mathfrak{m} = 1$) the zero frequency contributes $\mathcal{H}\widehat{\Phi}(0) \prod_{\mathfrak{p} \mid \mathfrak{r}} (1 - N\mathfrak{p}^{-1})$.

Proof. Inclusion–exclusion and $k = d\ell$ give

$$\sum_k \chi(k) \mathbf{1}_{(k, \mathfrak{r})=1} \Phi(Nk/\mathcal{H}) = \sum_{d \mid \text{rad } \mathfrak{r}} \mu(d) \chi(d) \sum_{\ell} \chi(\ell) \Phi\left(\frac{N\ell}{\mathcal{H}/Nd}\right).$$

Apply Poisson summation on the lattice $\mathcal{O}$ in ℓ , on the residue classes modulo $\mathfrak{m}$, at the scale $\mathcal{H}/Nd$ [17, Theorem 4.5]; the finite Fourier transform of a primitive character is $\sqrt{Nm} \gamma(\chi) \overline{\chi(h)}$ for every h [17, (3.12)], which gives (B.15). The statements about the zero frequency follow by summing $\mu(d)/Nd$ over $d \mid \text{rad } \mathfrak{r}$. $\square$

B.5. The dual mean square.

Definition B.5. Let $\mathcal{H}, X, F > 0$ be the norm scales of the row $k \in \mathcal{O}$, the column n and an auxiliary index f . For a ray class character ξ and a smooth compactly supported W on $(0, \infty)$ the *dual mean square* is

$$\mathcal{E}(\mathcal{H}, X, F; \xi, W) = \frac{1}{XF} \sum_{\substack{F \leq Nf < 2F \\ (f, S)=1}}^* \sum_{0 < Nk \leq \mathcal{H}(n, S)=1} \left| \sum_{(n, S)=1}^* a_{\xi}(n) \chi_n(k) \chi_n(f)^4 W(Nn/X) \right|^2. \quad (\text{B.16})$$

The twist $\chi_n(f)^4$ will arise from the common factors of two columns; note that $\chi_n(f)^4$ enforces $(n, f) = 1$ since $\chi_n(f) = 0$ otherwise. Write $\mathcal{E}_{\mathfrak{r}}$ for the same expression with the further restriction $(n, \mathfrak{r}) = 1$ on the column. The restriction is removed at the cost of a divisor factor, without changing the row range or the product XF .

Lemma B.6. *Let r_0 be the product of the primes dividing $\mathfrak{r}$ outside S . For $\mathcal{H}, X > 0$, $F \geq 1$ and smooth compactly supported W ,*

$$\mathcal{E}_{\mathfrak{r}}(\mathcal{H}, X, F; \xi, W) \leq \tau_{\text{div}}(r_0) \sum_{d \mid r_0} \mathcal{E}\left(\mathcal{H}, \frac{X}{Nd}, FNd; \xi, W\right). \quad (\text{B.17})$$

Proof. Let $C_{\mathfrak{r}}(X; k, f)$ be the inner column sum of $\mathcal{E}_{\mathfrak{r}}$. Inclusion–exclusion over $d \mid r_0$, the substitution $n = dm$ and (B.12) give

$$\begin{aligned} C_{\mathfrak{r}}(X; k, f) &= \sum_{d \mid r_0} \mu(d) \sum_{\substack{(n, S)=1 \\ d \mid n}}^* a_{\xi}(n) \chi_n(k) \chi_n(f)^4 W(Nn/X) \\ &= \sum_{\substack{d \mid r_0 \\ (d, f)=1}} \mu(d) a_{\xi}(d) \chi_d(k) \chi_d(f)^4 C_1(X/Nd; k, df), \end{aligned}$$

because $\chi_m(d)^4$ enforces $(m, d) = 1$ and combines with $\chi_m(f)^4$ into $\chi_m(df)^4$, while the terms with $(d, f) \neq 1$ vanish. Each coefficient in front of C_1 has modulus at most one. Cauchy–Schwarz in d gives the factor $\tau_{\text{div}}(r_0)$, and enlarging the injective image $f \mapsto df$ to the whole squarefree range $FNd \leq N(df) < 2FNd$ gives the terms on the right of (B.17); the normalisation is unchanged since $(X/Nd)(FNd) = XF$. $\square$

B.6. Reduction to the dual mean square.

Proposition B.7. *Fix $0 < \vartheta \leq 1/10$ and put $H = D^{1+\vartheta}$. For each fixed $C \geq 1$ and all real $B, F \geq 1$ consider the ranges*

$$X = \frac{D}{BF}, \quad 0 < \mathcal{H} \leq \frac{CD^2}{HB^2}, \quad XF = \frac{D}{B}. \quad (\text{B.18})$$

Suppose that for every $\varepsilon > 0$ there is an integer $J = J(\vartheta, \varepsilon) \geq 1$ such that, for every compact interval $I \subset (0, \infty)$ and every smooth W supported in I ,

$$\mathcal{E}(\mathcal{H}, X, F; \xi, W) \ll_{\nu, S, I, C, \vartheta, \varepsilon} \|W\|_{C^J(I)}^2 D^{\varepsilon} XF. \quad (\text{B.19})$$

Then Proposition B.2 holds.

Proof. Write $I = [a_0, b_0]$ and $N = N$. Choose a nonnegative radial Schwartz weight Φ with $\Phi \geq 1$ on $[0, 1]$ and $\text{supp } \hat{\Phi} \subset [0, C_{\Phi}]$ for a fixed $C_{\Phi} > 0$.¹³ It suffices to prove

¹³Such a weight is the square of a real radial Schwartz function with compactly supported Fourier transform, rescaled; its Fourier transform is then a convolution of compactly supported functions and is itself compactly supported.

$\mathcal{M}_D \ll \|W\|_{C^{J'}(I)}^2 HD^\varepsilon$ for some $J' = J'(\vartheta, \varepsilon)$, where

$$\mathcal{M}_D = \frac{1}{D} \sum_{u \in \mathcal{O}} \Phi(N(u)/H) \left| \sum_{(n,S)=1} \mu(n) \nu(n) \chi_n(u) W(N(n)/D) \right|^2, \quad (\text{B.20})$$

since the left side of (B.3) is at most $D\mathcal{M}_D$.

Expanding the square. Conjugate the inner sum, expand, and put $g = (n_1, n_2)$, $n_j = gz_j$; the squarefree indices satisfy $(z_1, z_2) = (z_1 z_2, g) = 1$ and $\overline{\chi_{n_1}(u)} \chi_{n_2}(u) = \mathbf{1}_{(u,g)=1} \chi_{z_1}(u) \chi_{z_2}(u)$. The character $\chi = \overline{\chi_{z_1}} \chi_{z_2}$ is primitive modulo $z_1 z_2$, so Lemma B.4 with the exclusion g gives

$$\sum_u \Phi(N(u)/H) \mathbf{1}_{(u,g)=1} \chi(u) = \sum_{e|g} \frac{H \mu(e) \chi(e) \gamma(\chi)}{N(e) \sqrt{N(z_1) N(z_2)}} \sum_h \overline{\chi(h)} \hat{\Phi}\left(\frac{HN(h)}{N(e)N(z_1)N(z_2)}\right). \quad (\text{B.21})$$

The zero frequency occurs only for $z_1 = z_2 = 1$ and contributes

$$Z = \frac{H}{D} \hat{\Phi}(0) \sum_{(g,S)=1}^* |W(N(g)/D)|^2 \prod_{p|g} \left(1 - \frac{1}{N(p)}\right) \ll_{\Phi, I} H \|W\|_\infty^2.$$

All $h \neq 0$ are retained, including those with $z_1 = z_2 = 1$.

Converting the coefficients. Expand the fixed ray class function $\overline{\nu(t)} G(t^{-1}) = \sum_\xi c_\xi \xi(t)$. The Chinese remainder theorem and (B.11), (B.13) give¹⁴

$$\mu(z_1) \mu(z_2) \overline{\nu(z_1)} \nu(z_2) \gamma(\overline{\chi_{z_1} \chi_{z_2}}) = \sum_\xi c_\xi a_\xi(z_1) \overline{a_\xi(z_2)}. \quad (\text{B.22})$$

Put $W_0(x) = x^{-1/2} \overline{W(x)}$. Substituting (B.22) into (B.21), with the factor $1/D$ of (B.20),

$$\mathcal{M}_D - Z = \sum_\xi c_\xi \mathcal{S}_\xi, \quad |\mathcal{M}_D - Z| \ll \max_\xi |\mathcal{S}_\xi|,$$

the sum over ξ being finite and fixed. Fix ξ . Using $\overline{\chi_z(e)} = \chi_z(e^5)$,

$$\begin{aligned} \mathcal{S}_\xi &= \frac{H}{D^2} \sum_g^* \sum_{e|g} \frac{\mu(e) N(g)}{N(e)} \sum_{h \neq 0} \sum_{\substack{z_1, z_2 \\ (z_1, z_2)=1 \\ (z_1 z_2, g)=1}}^* a_\xi(z_1) \overline{a_\xi(z_2)} \chi_{z_1}(he^5) \overline{\chi_{z_2}(he^5)} \\ &\quad \times W_0\left(\frac{N(gz_1)}{D}\right) \overline{W_0\left(\frac{N(gz_2)}{D}\right)} \hat{\Phi}\left(\frac{HN(h)}{N(e)N(z_1)N(z_2)}\right), \end{aligned}$$

where starred variables are squarefree, primary and prime to S , and h, k range over nonzero elements of $\mathcal{O}$. We must show $|\mathcal{S}_\xi| \ll \|W\|_{C^{J'}(I)}^2 HD^\varepsilon$.

Removing the coprimality of the columns. Insert $\mathbf{1}_{(z_1, z_2)=1} = \sum_{v|z_1, v|z_2} \mu(v)$ and put $z_j = vm_j$. By (B.12) the common factor satisfies $a_\xi(vm) = a_\xi(v) a_\xi(m) \chi_m(v)^4$ and $|a_\xi(v) \chi_v(he^5)|^2 = \mathbf{1}_{(v,h)=1}$, with $(v, e) = 1$. Thus

$$\begin{aligned} \mathcal{S}_\xi &= \frac{H}{D^2} \sum_{\substack{g, v \\ (g, v)=1}}^* \sum_{e|g} \frac{\mu(e) \mu(v) N(g)}{N(e)} \sum_{\substack{h \neq 0 \\ (h, v)=1}} \sum_{\substack{m_1, m_2 \\ (m_1 m_2, gv)=1}}^* a_\xi(m_1) \overline{a_\xi(m_2)} \\ &\quad \times \chi_{m_1}(he^5 v^4) \overline{\chi_{m_2}(he^5 v^4)} W_0\left(\frac{N(gvm_1)}{D}\right) \overline{W_0\left(\frac{N(gvm_2)}{D}\right)} \hat{\Phi}\left(\frac{HN(h)}{N(e)N(v)^2 N(m_1)N(m_2)}\right), \end{aligned}$$

with no restriction $(m_1, m_2) = 1$ left. Make the bijective change of variables

$$b = g/e, \quad f = ev, \quad k = eh; \quad e = (f, k), \quad v = f/e, \quad g = be, \quad h = k/e.$$

¹⁴By the Chinese remainder theorem the Gauss sum factors as $\gamma(\overline{\chi_{z_1} \chi_{z_2}}) = \gamma_{-1}(z_1) \gamma_1(z_2) \overline{\chi_{z_1}(z_2)} \chi_{z_2}(z_1)$. Convert $\mu(z_1) \gamma_{-1}(z_1)$ by (B.11) and $\mu(z_2) \gamma_1(z_2)$ by (B.10). The ray class factors then collect, by (B.13) with $t = z_1 z_2^{-1}$, into $\chi_{z_1}(-1) G(z_1)^{-1} G(z_2) \mathcal{R}(z_1, z_2) \overline{\nu(z_1)} \nu(z_2)$, which equals $\overline{\nu(t)} G(t^{-1})$.

Then b, f are coprime and squarefree, $k \neq 0$ is arbitrary, $gv = bf$, $he^5v^4 = kf^4$, $\mu(e)\mu(v) = \mu(f)$ and $N(g)/N(e) = N(b)$, so

$$\begin{aligned} \mathcal{S}_\xi = \frac{H}{D^2} \sum_{\substack{b,f \\ (b,f)=1}}^* \mu(f)N(b) \sum_{k \neq 0} \sum_{\substack{m_1, m_2 \\ (m_1 m_2, b)=1}}^* a_\xi(m_1) \overline{a_\xi(m_2)} \chi_{m_1}(kf^4) \overline{\chi_{m_2}(kf^4)} \\ \times W_0\left(\frac{N(bfm_1)}{D}\right) \overline{W_0\left(\frac{N(bfm_2)}{D}\right)} \widehat{\Phi}\left(\frac{HN(k)}{N(f)^2 N(m_1)N(m_2)}\right). \end{aligned} \quad (\text{B.23})$$

The characters enforce $(m_j, f) = 1$, so only the exclusion $(m_j, b) = 1$ is displayed. The supports give $N(b)N(f) \leq b_0 D$ and $0 < N(k) \leq C_\Phi b_0^2 D^2 / (HN(b)^2)$.

Dyadic ranges and the smooth weights. Partition $B \leq N(b) < 2B$ and $F \leq N(f) < 2F$ dyadically, write $\mathcal{S}_{\xi; B, F}$ for the corresponding part of (B.23), and put

$$X = \frac{D}{BF}, \quad \mathcal{H} = \frac{C_{I, \Phi} D^2}{HB^2}, \quad C_{I, \Phi} = \max(2, C_\Phi b_0^2),$$

which satisfy (B.18). In the variables $x_j = N(m_j)/X$ the coupled weight is

$$\begin{aligned} \mathcal{K}_{b, f, k}(x_1, x_2) &= W_0(rx_1) \overline{W_0(rx_2)} \widehat{\Phi}\left(\frac{a}{x_1 x_2}\right), \\ r = \frac{N(b)N(f)}{BF} &\in [1, 4), \quad a = \frac{HN(k)}{N(f)^2 X^2} \leq C_{I, \Phi}, \end{aligned}$$

and for every integer $q \geq 0$ these kernels satisfy $\sup_{b, f, k} \|\mathcal{K}_{b, f, k}\|_{C^q([a_0/4, b_0]^2)} \ll_{I, \Phi, q} \|W\|_{C^q(I)}^2$. For $U \in C_c^\infty(I_*)$, $I_* = [a_0/8, 2b_0]$, put

$$S_U(b, f, k) = \sum_{\substack{(n, S)=1 \\ (n, b)=1}}^* a_\xi(n) \chi_n(k) \chi_n(f)^4 U(N(n)/X).$$

The shifts $(X, F) \mapsto (X/N(d), FN(d))$ preserve (B.18), so Lemma B.6 and the hypothesis (B.19) with $J = J(\vartheta, \varepsilon/4)$ give

$$\begin{aligned} \sum_{f, k} |S_U(b, f, k)|^2 &= XF \mathcal{E}_{(b)}(\mathcal{H}, X, F; \xi, U) \\ &\leq XF \tau_{\text{div}}(b) \sum_{d|b} \mathcal{E}(\mathcal{H}, X/N(d), FN(d); \xi, U) \ll D^{\varepsilon/2} (XF)^2 \|U\|_{C^J(I_*)}^2, \end{aligned}$$

the divisor factor being absorbed in $D^{\varepsilon/4}$ since $N(b) \ll_I D$. Now Lemma B.22 below turns a mean-square bound for every common test function U into a bound for the quadratic form with the row-dependent kernel $\mathcal{K}_{b, f, k}$, at the cost of the C^{2J+4} norm of the kernel. Apply it for each fixed b with row index (f, k) , coefficient $\mu(f)\mathbf{1}_{(f, b)=1}$ and kernel $\mathcal{K}_{b, f, k}$; the larger row range $0 < N(k) \leq \mathcal{H}$ only adds rows whose kernel vanishes. With $q = 2J + 4$,

$$\begin{aligned} |\mathcal{S}_{\xi; B, F}| &\ll D^{\varepsilon/2} \|W\|_{C^q(I)}^2 \sum_{B \leq N(b) < 2B}^* \frac{HN(b)}{D^2} (XF)^2 \\ &\ll D^{\varepsilon/2} \|W\|_{C^q(I)}^2 \frac{HB^2 (XF)^2}{D^2} = D^{\varepsilon/2} \|W\|_{C^q(I)}^2 H, \end{aligned} \quad (\text{B.24})$$

since $XF = D/B$ and there are $O(B)$ ideals b in the range. Summing the $O((\log D)^2)$ nonempty dyadic ranges bounds $\mathcal{S}_\xi$, summing over the finite set of ξ and adding Z bounds $\mathcal{M}_D$, and multiplying by D gives (B.3); the derivative order depends only on ϑ and ε . $\square$

B.7. The completed sums and the two inputs. Put $\Sigma = XF$. The dual mean square is bounded by the following proposition, whose proof occupies Sections B.7 to B.13.

Proposition B.8. *Fix $\kappa > 0$ and $C_0 \geq 1$, and suppose that*

$$\mathcal{H}, X, F \geq 1, \quad \Sigma = XF \leq D^{C_0}, \quad \mathcal{H} \leq \Sigma D^{-\kappa}. \quad (\text{B.25})$$

For every $\varepsilon > 0$ there is an integer $J = J(\kappa, C_0, \varepsilon) \geq 1$ such that

$$\mathcal{E}(\mathcal{H}, X, F; \xi, W) \ll_{I, \nu, S, \kappa, C_0, \varepsilon} \|W\|_{C^J(I)}^2 D^\varepsilon \Sigma \quad (\text{B.26})$$

for every compact interval $I \subset (0, \infty)$ and every smooth W supported in I .

The completed sums. Let Ψ be a completely multiplicative function on the ideals prime to 3, vanishing on the ideals divisible by a prime of S , and write $\Psi(n) = \Psi((n))$ for primary n . Define

$$T(X; \Psi) = \sum_{(n, S)=1}^* \sum_{\substack{b \equiv 1 \pmod{3} \\ (b, S)=1}} \frac{\overline{\alpha(n)} \gamma_2(n) \Psi(n) \overline{\alpha(b)}^3 \Psi(b)^3}{\sqrt{Nn} Nb} V_*(Nn Nb^3/X), \quad V_*(y) = \sqrt{y} W(y), \quad (\text{B.27})$$

where n is squarefree and b is any primary element prime to S . The index b supplies the cubes that occur in the Fourier expansion of the cubic theta function, and the $b = 1$ part is $X^{-1/2} \sum_n^* \overline{\alpha(n)} \gamma_2(n) \Psi(n) W(Nn/X)$. For a ray class character ξ , a nonzero row $k \in \mathcal{O}$ and a squarefree primary f prime to S put

$$\Psi_k(n) = \xi(n) \chi_n(k) \chi_n(f)^4 \text{ for } (n, S) = 1, \quad \Psi_k(n) = 0 \text{ otherwise,} \quad T(X; k, f) = T(X; \Psi_k), \quad (\text{B.28})$$

so that the $b = 1$ part of $T(X; k, f)$ is $X^{-1/2}$ times the column sum of (B.16).

The completed mean-square estimate. The theta transformation turns the sextic twists of Ψ_k into quadratic characters, and the quadratic large sieve then gives the following estimate, proved in Section B.12.

Proposition B.9. *Fix $\varepsilon > 0$, $C_0 \geq 1$ and a character ξ of the fixed ray class group. There is $J = J(\varepsilon, C_0) \geq 1$ such that*

$$\sum_{0 < Nk \leq \mathcal{H}} |T(X; k, f)|^2 \ll_{I, \xi, S, \varepsilon, C_0} D^\varepsilon \|W\|_{C^J(I)}^2 \left(\mathcal{H} + \frac{\mathcal{H}^2 N f}{X} \right) \quad (\text{B.29})$$

whenever $1 \leq \mathcal{H}, X, Nf \leq D^{C_0}$, for f squarefree primary with $(f, S) = 1$, W smooth and supported in a compact $I \subset (0, \infty)$, and T as in (B.28).

Cube inversion. Set

$$H_c^3 = \min\left(X, \frac{X^2}{\mathcal{H}^2}\right), \quad L_b = \frac{X}{Nb^3}. \quad (\text{B.30})$$

Lemma B.10. *Fix $C_0 \geq 1$ and $\varepsilon > 0$, and suppose $1 \leq \mathcal{H}, X, F \leq D^{C_0}$ and $\mathcal{H} \leq \Sigma = XF$. There is an integer $J = J(\varepsilon, C_0) \geq 1$ such that*

$$\mathcal{E}(\mathcal{H}, X, F; \xi, W) \ll_{I, \nu, S, C_0, \varepsilon} D^\varepsilon \left(\Sigma \|W\|_{C^J(I)}^2 + \sup_{\substack{b \equiv 1 \pmod{3}, (b, S)=1 \\ Nb > H_c, L_b > 1}} \mathcal{E}(\mathcal{H}, L_b, F; \xi, W) \right) \quad (\text{B.31})$$

for every compact $I \subset (0, \infty)$ and $W \in C_c^\infty(I)$. An empty supremum is zero, and it is empty when $\mathcal{H}^2 \leq X$.

Proof. Write $N = N$ and $I = [u, v]$. Complete multiplicativity, including at the zeros of Ψ_k , gives the Möbius inversion

$$X^{-1/2} \sum_{(n, S)=1}^* a_\xi(n) \chi_n(k) \chi_n(f)^4 W(N(n)/X) = \sum_{(h, S)=1} \frac{\mu(h) \overline{\alpha(h)}^3 \Psi_k(h)^3}{N(h)} T(X/N(h)^3; k, f) : \quad (\text{B.32})$$

after substituting (B.27) the coefficient of the total cube index b contains $\sum_{h|b} \mu(h) = \mathbf{1}_{b=1}$.¹⁵

Split the right side into $P_{\text{short}}(k, f)$, the terms with $N(h) \leq H_c$, and $P_{\text{long}}(k, f)$, the rest, and let $\mathcal{E}_{\text{short}}$, $\mathcal{E}_{\text{long}}$ be the corresponding mean squares $F^{-1} \sum_f^* \sum_k |P|^2$. Weighted Cauchy–Schwarz for each k, f and Proposition B.9, with the exponent $\varepsilon/4$ and the range $C_0 + 1$ (since $N(f) < 2D^{C_0}$), give

$$\begin{aligned} \mathcal{E}_{\text{short}} &\leq \left(\sum_{N(h) \leq H_c} \frac{1}{N(h)} \right) \sum_{N(h) \leq H_c} \frac{1}{N(h)} \frac{1}{F} \sum_{\substack{F \leq N(f) < 2F \\ (f, S)=1}}^* \sum_{0 < N(k) \leq \mathcal{H}} |T(X/N(h)^3; k, f)|^2 \\ &\ll D^{\varepsilon/2} \|W\|_{C^J(I)}^2 \left(\mathcal{H} + \frac{\mathcal{H}^2 F H_c^3}{X} \right) \ll D^{\varepsilon/2} \Sigma \|W\|_{C^J(I)}^2, \end{aligned} \quad (\text{B.33})$$

where $X/N(h)^3 \geq 1$, the reciprocal norm sum is a logarithm, and the last step uses $\mathcal{H} \leq \Sigma$ and $H_c^3 \leq X^2/\mathcal{H}^2$. If $H_c < 1$ the short sum is empty.

In the long part expand T by (B.27) with the cube index c , put $b = hc$, and note that the pairs giving b are the $(h, b/h)$ with $h \mid b$, $N(h) > H_c$, and that $\Psi_k(b)^3 = \xi(b)^3 \chi_b(k)^3 \mathbf{1}_{(b, f)=1}$. This gives

$$\begin{aligned} P_{\text{long}}(k, f) &= \sum_{\substack{(b, S)=1 \\ N(b) > H_c}} \frac{\beta_0(b, f) \chi_b(k)^3}{N(b)} \frac{1}{\sqrt{L_b}} \sum_{(n, S)=1}^* a_\xi(n) \chi_n(k) \chi_n(f)^4 W(N(n)/L_b), \\ \beta_0(b, f) &= \overline{\alpha(b)}^3 \xi(b)^3 \mathbf{1}_{(b, f)=1} \sum_{\substack{h|b \\ N(h) > H_c}} \mu(h), \end{aligned}$$

with $|\beta_0(b, f)| \leq \tau_{\text{div}}(b)$; the indices are primary and prime to S , b need not be squarefree, and the support of W restricts to $N(b)^3 \leq vX$. Weighted Cauchy–Schwarz gives

$$\mathcal{E}_{\text{long}} \leq \left(\sum_{\substack{H_c < N(b) \leq (vX)^{1/3} \\ (b, S)=1}} \frac{\tau_{\text{div}}(b)}{N(b)} \right) \sum_{\substack{H_c < N(b) \leq (vX)^{1/3} \\ (b, S)=1}} \frac{\tau_{\text{div}}(b)}{N(b)} \mathcal{E}(\mathcal{H}, L_b, F; \xi, W),$$

and the reciprocal divisor sum is $\ll_I \log^2(2 + X)$, because $\sum_{N(b) \leq R} \tau_{\text{div}}(b)/N(b)$ is at most $(\sum_{N(c) \leq R} 1/N(c))^2 \ll \log^2(2R)$. If $L_b \leq 1$ the column sum is empty unless $L_b \geq 1/v$, and then it has $O_I(1)$ terms, so $\mathcal{E}(\mathcal{H}, L_b, F; \xi, W) \ll_I (\mathcal{H}/L_b) \|W\|_\infty^2 \ll_I \Sigma \|W\|_\infty^2$. Absorb the logarithms in D^ε and use $\mathcal{E} \leq 2\mathcal{E}_{\text{short}} + 2\mathcal{E}_{\text{long}}$. If $\mathcal{H}^2 \leq X$ then $H_c = X^{1/3}$ and $N(b) > H_c$ forces $L_b < 1$, so the supremum is empty. $\square$

The transfer estimate. Fix a nonnegative radial Schwartz weight Φ with $\Phi \geq 1$ on $[0, 1]$ and $\text{supp } \hat{\Phi} \subset [0, C_\Phi]$, as in the proof of Proposition B.7. For fixed $\mathcal{H}, L, F, \xi$ put

$$\mathcal{A}(W) = \frac{1}{LF} \sum_{\substack{F \leq Nf < 2F \\ (f, S)=1}}^* \sum_{k \in \mathcal{O}} \Phi(Nk/\mathcal{H}) \left| \sum_{(n, S)=1}^* a_\xi(n) \chi_n(k) \chi_n(f)^4 W(Nn/L) \right|^2, \quad (\text{B.34})$$

so that $\mathcal{E}(\mathcal{H}, L, F; \xi, W) \leq \mathcal{A}(W)$. In the next proposition Σ is an independent target scale; it will be applied at the column scale L_b with $\Sigma = XF$.

Proposition B.11. *Assume $1 \leq \mathcal{H}, L, F, \Sigma \leq D^{C_0}$ and $\max\{\mathcal{H}, LF\} \leq \Sigma$, for a fixed $C_0 \geq 1$. For every integer $m \geq 0$ and every $\varepsilon > 0$,*

$$\mathcal{A}(W) \ll D^\varepsilon \Sigma \|W\|_{C^{4m+12}(I)}^2 \left(1 + \sup \frac{\mathcal{E}(\mathcal{H}', X', F'; \xi', U)}{\Sigma'} \right), \quad (\text{B.35})$$

¹⁵The weight $w_k(b) = \overline{\alpha(b)}^3 \Psi_k(b)^3$ is completely multiplicative, so $w_k(h)w_k(b/h) = w_k(b)$ even when h and b/h share prime factors, which is what makes the inversion exact; compare [17, (1.18)] and [8, (8.2)]. One cannot simply discard the terms $b \neq 1$ of T , because the contributions of different b can cancel.

where the supremum is over the family (B.16) with $\Sigma' = X'F'$, $\mathcal{H}', X', F' \geq 1$,

$$\mathcal{H}' \leq \frac{\mathcal{H}L}{\Sigma F}, \quad \frac{\mathcal{H}'}{\Sigma'} \leq \frac{\mathcal{H}}{\Sigma}, \quad \Sigma' \leq L, \quad (\text{B.36})$$

and, for $I = [u, v]$, test functions $U \in C_c^\infty(I')$, $I' = [u/16, 4v]$, with $\|U\|_{C^m(I')} \leq 1$; an empty supremum is zero. The implied constant depends on $m, C_0, \varepsilon, I, \nu, S$ and the fixed ray class group.

The proposition is proved in Section B.13.

B.8. Proof of Proposition B.8.

Proof. Fix $\kappa > 0$ and $C_0 \geq 1$. We prove by induction on $j \geq 0$ that the proposition holds under the further restriction $\mathcal{H} \leq D^{j\kappa}$, with a derivative order $J = J(j, \kappa, C_0, \varepsilon)$ and an implied constant depending also on j ; the derivative order is independent of I , which is what allows the induction hypothesis to be applied on the enlarged interval of the transfer estimate.

For $j = 0$ we have $\mathcal{H} \leq 1$. The support of W restricts n to the $O_I(X)$ ideals with $Nn \in [uX, vX]$, the factors $a_\xi(n)$, $\chi_n(k)$, $\chi_n(f)$ ⁴ have modulus at most one, so each inner sum is $O_I(X\|W\|_\infty)$; there are $O(F)$ choices of f and $O(\mathcal{H})$ of k , whence $\mathcal{E} \ll_I (F\mathcal{H}/XF)(X\|W\|_\infty)^2 = \mathcal{H}X\|W\|_\infty^2 \leq \Sigma\|W\|_\infty^2$, as $\mathcal{H} \leq 1 \leq F$. This is the case $j = 0$ with $J = 1$.

Suppose the assertion holds for j , fix $\varepsilon > 0$, and let m be the derivative order supplied by the induction hypothesis with the exponent $\varepsilon/3$. For $\mathcal{H} \leq D^{(j+1)\kappa}$ apply Lemma B.10 with the exponent $\varepsilon/3$. If its supremum is nonempty then $\mathcal{H}^2 > X$ and $H_c^3 = X^2/\mathcal{H}^2$. For each b in it, $1 < L_b \leq X$ and $\Sigma \geq \max\{\mathcal{H}, L_b F\}$, so Proposition B.11 applies at the column scale L_b with the derivative order m and the exponent $\varepsilon/3$, and it bounds $\mathcal{E}(\mathcal{H}, L_b, F; \xi, W) \leq \mathcal{A}(W)$. For any parameters $\mathcal{H}', X', F'$ in its supremum, (B.36) and $Nb > H_c$ give

$$\begin{aligned} \mathcal{H}' &\leq \frac{\mathcal{H}L_b}{\Sigma F} = \frac{\mathcal{H}}{Nb^3 F^2} < \mathcal{H} \left(\frac{\mathcal{H}}{\Sigma} \right)^2 \leq D^{-2\kappa} \mathcal{H} \leq D^{j\kappa}, \\ \frac{\mathcal{H}'}{\Sigma'} &\leq \frac{\mathcal{H}}{\Sigma} \leq D^{-\kappa}, \quad \Sigma' \leq L_b \leq \Sigma \leq D^{C_0}. \end{aligned}$$

Each of these mean squares is therefore covered by the induction hypothesis, on the interval $[u/16, 4v]$ with a weight of C^m norm at most one, which gives $\mathcal{E}(\mathcal{H}', X', F'; \xi', U)/\Sigma' \ll D^{\varepsilon/3}$ with the permitted dependence on I . Substituting into (B.35) gives, uniformly in b , $\mathcal{E}(\mathcal{H}, L_b, F; \xi, W) \ll D^{2\varepsilon/3} \Sigma\|W\|_{C^{4m+12}(I)}^2$, and substituting this into (B.31) gives the remaining factor $D^{\varepsilon/3}$. Take J at least $4m+12$ and at least the order required by Lemma B.10; then $\mathcal{E}(\mathcal{H}, X, F; \xi, W) \ll D^\varepsilon \Sigma\|W\|_{C^J(I)}^2$, with J depending only on $j, \kappa, C_0, \varepsilon$. If the supremum is empty the bound follows from the lemma alone. This completes the induction, and $j = \lceil C_0/\kappa \rceil$ covers every $\mathcal{H} \leq \Sigma \leq D^{C_0}$.¹⁶ $\square$

B.9. Proof of Proposition B.2.

Proof. Fix $0 < \vartheta \leq 1/10$ and $H = D^{1+\vartheta}$. The ranges (B.18) give

$$\mathcal{H} \leq \frac{CD^{1-\vartheta}}{B^2}, \quad \Sigma = XF = \frac{D}{B}, \quad \frac{\mathcal{H}}{\Sigma} \ll D^{-\vartheta}.$$

For large D , Proposition B.8 applies with $\kappa = \vartheta/2$ and $C_0 = 2$. If $X < 1$ a nonempty support forces $X \gg_I 1$ and counting gives $\mathcal{E} \ll_I \mathcal{H}\|W\|_\infty^2 \ll \Sigma\|W\|_\infty^2$; if $\mathcal{H} < 1$ the sum is empty. So (B.19) holds with a derivative order depending only on ϑ, ε , and Proposition B.7 gives (B.3); bounded D is covered by counting. $\square$

¹⁶The contraction is the point of the whole construction: each passage through the cube inversion and the transfer estimate divides the row range by at least $D^{2\kappa}$ while keeping the ratio $\mathcal{H}/\Sigma$, so that after $O(C_0/\kappa)$ steps the row range is at most 1 and trivial counting finishes. This is the admissible-exponent induction of [12, Lemma 8 and §8], [13, Lemma 9]; see also [9, Theorem 4.1].

B.10. The cubic theta function and its transformation. We now express the completed sum $T(X; k, f)$ through the Fourier coefficients of Kubota's cubic theta function, whose automorphy will give a transformation formula for it. Write $(z, v) \in \mathbb{C} \times \mathbb{R}_{>0}$ for the coordinates of upper half-space, and use the theta function in the normalisation of Dunn and Radziwiłł [8, §5.1, (5.6)]:

$$\theta(z, v) = \frac{3^{5/2}}{2} v^{2/3} + \sum_{\substack{\ell \in \lambda^{-3}\mathcal{O} \\ \ell \neq 0}} \tau(\ell) v K_{1/3}(4\pi|\ell|v) \exp(2\pi i(\ell z + \bar{\ell} \bar{z})), \quad (\text{B.37})$$

where $K_{1/3}$ is the modified Bessel function of the second kind and $\tau(\ell)$ is the coefficient sequence given explicitly in [8, (5.7)–(5.8)] from Patterson's calculation [26, Theorem 8.1].¹⁷ For primary n, b with n squarefree and $(nb, 6) = 1$ write $c_\theta(nb^3)$ for the coefficient of $\bar{\theta}$ at the index $\lambda^{-3}nb^3$; Patterson's formula, as recorded in [8, (5.7)], is

$$c_\theta(nb^3) = 3^{5/2} |b| \overline{\chi_n(\lambda)^2} \gamma_2(n). \quad (\text{B.38})$$

Define $\phi_k(n) = \chi_n(\lambda)^2 \Psi_k(n)$ for primary n prime to S and $\phi_k(n) = 0$ otherwise, and the twisted series

$$\Theta_k(z, v) = \sum_{\substack{\ell \in \lambda^{-3}\mathcal{O} \\ \ell \neq 0}} \overline{\tau(-\ell)} \phi_k(\lambda^3 \ell) v K_{1/3}(4\pi|\ell|v) \exp(2\pi i(\ell z + \bar{\ell} \bar{z})), \quad (\text{B.39})$$

whose coefficients are those of $\bar{\theta}$, namely $\overline{\tau(-\ell)}$ at ℓ , multiplied by $\phi_k(\lambda^3 \ell)$. Choose $q \in \mathcal{O} \setminus \{0\}$ such that ϕ_k is periodic modulo (q) and put $\hat{\phi}_k(h) = Nq^{-1} \sum_{n \bmod q} \phi_k(n) e(-hn/q)$.

Lemma B.12. *With these definitions,*

$$\begin{aligned} T(X; k, f) &= \frac{1}{3^{5/2} \sqrt{X}} \sum_{\substack{n, b \equiv 1 \pmod{3} \\ n \text{ squarefree} \\ (nb, S) = 1}} c_\theta(nb^3) \phi_k(nb^3) \overline{\alpha(nb^3)} W\left(\frac{N(nb^3)}{X}\right), \\ \Theta_k(z, v) &= \sum_{h \in \mathcal{O}/(q)} \hat{\phi}_k(h) \overline{\theta(z + \lambda^2 h/q, v)}. \end{aligned} \quad (\text{B.40})$$

Proof. By (B.38) and complete multiplicativity, $c_\theta(nb^3) \phi_k(nb^3) = 3^{5/2} |b| \gamma_2(n) \Psi_k(n) \Psi_k(b)^3$, and $\overline{\alpha(nb^3)} = \overline{\alpha(n)} \overline{\alpha(b)}^3$; substitution into (B.27) with $V_*(y) = \sqrt{y} W(y)$ gives the first identity. For the second, finite Fourier inversion gives $\phi_k(n) = \sum_h \hat{\phi}_k(h) e(nh/q)$ and $\sum_h \hat{\phi}_k(h) = \phi_k(0) = 0$; multiplying the ℓ th Fourier mode by $e(\lambda^3 \ell h/q)$ translates z by $\lambda^2 h/q$, and the second identity cancels the constant terms of the translates. $\square$

The local factors. For a primary prime $p \notin S$, $j \in \{0, \dots, 5\}$ and $x \in \mathcal{O}$ define

$$B_{p,j}(x) = \begin{cases} \chi_p(x)^{-j-2}, & j \neq 0, 4, \\ Np^{-1/2}(-1 + Np \mathbf{1}_{p|x}), & j = 4, \\ Np^{-1/2} \chi_p(x)^{-2}, & j = 0. \end{cases} \quad (\text{B.41})$$

The case that matters most is $B_{p,1}(x) = \chi_p(x)^3$, a quadratic character. For the weight V_* of (B.27) put $\hat{V}_*(s) = \int_0^\infty V_*(x) x^s dx/x$, write $\int_{(\sigma)}$ for the integral up the line $\operatorname{Re} t = \sigma$, and define the transformed weight

$$V_*^\sharp(x) = \frac{1}{2\pi i} \int_{(0)} \hat{V}_*(-t) \frac{\Gamma(7/6+t)\Gamma(5/6+t)}{\Gamma(7/6-t)\Gamma(5/6-t)} \left(\frac{(2\pi)^4 x}{27}\right)^{-t} dt \quad (x > 0). \quad (\text{B.42})$$

¹⁷The theta function is the residue at its first pole of Kubota's cubic metaplectic Eisenstein series on $\operatorname{SL}_2(\mathbb{C})$ [19]; Patterson computed its Fourier coefficients and found cubic Gauss sums at the squarefree indices. Everything we use about it is the coefficient formula (B.38), the automorphy under the congruence subgroup $\Gamma(3)$ with Kubota's character κ , the invariance under $\operatorname{SL}_2(\mathbb{Z})$ and under the translations by $\mathbb{Z} + 3\mathcal{O}$, and the cusp expansions (B.44), all taken from [8] and [26].

The dual sums use the Fourier coefficients of $\bar{\theta}$ at three cusps. With the representatives

$$\gamma_0 = I, \quad \gamma_+ = \begin{pmatrix} 1 & 0 \\ \omega & 1 \end{pmatrix}, \quad \gamma_- = \begin{pmatrix} 1 & 0 \\ \omega^2 & 1 \end{pmatrix} \quad (\text{B.43})$$

acting on upper half-space, define $d_\sigma(\ell)$, $\sigma \in \{0, +, -\}$, by the expansions [8, (5.9), (5.15)]

$$\overline{\theta(\gamma_\sigma(z, v))} = \frac{3^{5/2}}{2} \mathbf{1}_{\sigma=0} v^{2/3} + \sum_{0 \neq \ell \in \lambda^{-4}\mathcal{O}} d_\sigma(\ell) v K_{1/3}(4\pi|\ell|v) \exp(2\pi i(\ell z + \bar{\ell}z)), \quad (\text{B.44})$$

so that $d_0(\ell) = \overline{\tau(-\ell)}$ with τ extended by zero outside $\lambda^{-3}\mathcal{O}$, and $c_\theta(nb^3) = d_0(\lambda^{-3}nb^3)$. The formulas for the three sequences are (B.67) and (B.68) below.

We state the transformation for a general product of local twists. Fix a ray class character with conductor supported on S and let Ψ_0 be its extension by zero at the primes of S ; let $\mathcal{P}$ be a finite set of primes outside S , represented by primary generators, with exponents $0 \leq j_p \leq 5$, and put $\Psi(n) = \Psi_0(n) \prod_{p \in \mathcal{P}} \chi_p^{j_p}(n)$ for primary n . In the transformed sums, $\mathcal{A}$ ranges over the subsets with $\{p : j_p \neq 0\} \subseteq \mathcal{A} \subseteq \mathcal{P}$; the primes of $\mathcal{A}$ are called *active*, the others *inactive*, so only primes with $j_p = 0$ can be inactive. For such $\mathcal{A}$ and $c_0 \in \mathcal{O} \setminus \{0\}$ write $c = c_0 \prod_{p \in \mathcal{A}} p$.

Proposition B.13. *The sum $T(X; \Psi)$ of (B.27) is a sum of $O_{\Psi_0, S}(2^{|\mathcal{P}|})$ terms of the form*

$$C \sum_{0 \neq \ell \in \lambda^{-4}\mathcal{O}} \frac{d(\ell)\alpha(\ell)}{\sqrt{N\ell}} \psi(\lambda^4\ell) \prod_{p \in \mathcal{A}} B_{p, j_p}(\lambda^4\ell) V_*^\# \left(\frac{N\ell X}{Nc^2} \right), \quad (\text{B.45})$$

with $|C| \ll_{\Psi_0, S} 1$, where the triples (d, ψ, c_0) belong to a fixed finite family depending only on Ψ_0, S : $d \in \{d_0, d_+, d_-\}$, $c_0 \in \mathcal{O} \setminus \{0\}$, and ψ is a unit-modulus additive character of $\mathcal{O}$.

To average over the row we need the data to be chosen consistently as the row varies. Fix a finite set $\mathcal{P}_{\text{fix}}$ of primary primes outside S with exponents j_p , and for k_0 primary, squarefree and prime to $S \prod_{p \in \mathcal{P}_{\text{fix}}} p$ put

$$\Psi_{k_0}(n) = \Psi_0(n) \prod_{p \in \mathcal{P}_{\text{fix}}} \chi_p^{j_p}(n) \prod_{p|k_0} \chi_p(n), \quad \mathcal{A} = \mathcal{A}_{\text{fix}} \cup \{p : p \mid k_0\},$$

where $\mathcal{A}_{\text{fix}}$ ranges over the subsets with $\{p \in \mathcal{P}_{\text{fix}} : j_p \neq 0\} \subseteq \mathcal{A}_{\text{fix}} \subseteq \mathcal{P}_{\text{fix}}$; every prime dividing k_0 is active since its exponent is 1.

Lemma B.14 (uniformity in the twist). *For the family Ψ_{k_0} the summands of Proposition B.13 may be indexed by $(h, \mathcal{A}_{\text{fix}})$ with h in a fixed finite set depending only on Ψ_0, S , zero coefficients allowed, and for each index the triple (d, ψ, c_0) depends on k_0 only through its ray class modulo a fixed ideal supported on S .*

Lemma B.15. *For each $d \in \{d_0, d_+, d_-\}$ the coefficient $d(\ell)$ vanishes unless*

$$\ell = u\lambda^m nb^3, \quad u \in \mathcal{O}^\times, \quad m \in \mathbb{Z}, \quad m \geq -4, \quad n, b \text{ primary}, \quad n \text{ squarefree}, \quad (\text{B.46})$$

and then $|d(\ell)| \leq 27 \cdot 3^{m/6} |b|$. For $A > 0$, integers $j \geq 0$ and V_* supported in a fixed compact $I \subset (0, \infty)$ there is $J = J(A, j)$ with

$$|(x\partial_x)^j V_*^\#(x)| \ll_{A, j, I} \|V_*\|_{C^j(I)} \min(x^{1/4}, x^{-A}) \quad (x > 0). \quad (\text{B.47})$$

Proposition B.13 and the two lemmas are proved in Section B.14.

B.11. The quadratic large sieve.

Lemma B.16. *Let $\mathcal{H}, U \geq 1$ and let $\beta(n)$ be complex coefficients on squarefree primary n with $Nn \asymp U$. For every $\varepsilon > 0$,*

$$\sum_{\substack{k \equiv 1 \pmod{3}, \\ Nk \leq \mathcal{H}}}^* \left| \sum_{\substack{(k, S)=1 \\ Nn \asymp U}} \beta(n) \chi_k(n)^3 \right|^2 \ll_{S, \varepsilon} (\mathcal{H}U)^\varepsilon (\mathcal{H} + U) \sum_n |\beta(n)|^2, \quad (\text{B.48})$$

the star restricting k to squarefree primary elements.

Proof. This is the quadratic large sieve of Goldmakher and Louvel over number fields [9, Theorem 1.1], a generalisation of Heath-Brown’s [12], after fixing the product of the prime factors of n in S and finitely many ray classes. To match their hypotheses, let $e_k \in \{0, 1\}$ according as $Nk \equiv 1, 3 \pmod{4}$ and let κ_λ be the nontrivial character modulo λ ; the character $x \mapsto (x/k)_2 \kappa_\lambda(x)^{e_k}$ is trivial on the units and has primitive conductor $k\lambda^{e_k}$. Classes modulo $24\mathcal{O}$ fix the supplementary characters and the reciprocity factors, and for coprime k_1, k_2 in the same class the product character has conductor $k_1 k_2$; these are the hypotheses of [9, Definition 1 and §2].¹⁸ $\square$

B.12. Squarefree rows and the proof of Proposition B.9. Write $N = N$. In (B.28) allow any $g \in \mathcal{O} \setminus \{0\}$ in place of f , so $\Psi_k(n) = \xi(n)\chi_n(k)\chi_n(g)^4$ with the zero extension at S retained. Choose generators of the prime ideals, primary away from 3, and extend multiplicatively to all ideals. We first bound the squarefree rows.

Lemma B.17. *For every $\varepsilon > 0$ and $C_0 \geq 1$ there is $J = J(\varepsilon, C_0) \geq 1$ such that*

$$\sum_{\substack{0 < N(s) \leq \mathcal{H} \\ s \text{ squarefree}}} |T(X; u_0 s, g)|^2 \ll_{I, \xi, S, \varepsilon, C_0} D^\varepsilon \|W\|_{C^J(I)}^2 \left(\mathcal{H} + \frac{\mathcal{H}^2 N(g)}{X} \right) \quad (\text{B.49})$$

for $1 \leq \mathcal{H}, X, N(g) \leq D^{C_0}$, $u_0 \in \mathcal{O}^\times$, every compact $I \subset (0, \infty)$ and $W \in C_c^\infty(I)$, the sum running over the chosen generators of the squarefree ideals, including those meeting S .

Proof. By homogeneity assume $\|W\|_{C^J(I)} \leq 1$ with J chosen below. Write $s = tk_0$ with t the product of the primes $p \mid s$ that divide g or lie in S ; then k_0 is squarefree, primary, prime to g and to S , and $N(k_0) \leq \mathcal{H}_0 := \mathcal{H}/N(t)$. Fix t and discard empty ranges, so $\mathcal{H}_0 \geq 1$. For a coefficient function $A(n, b)$, a scale $Y > 0$ and an integer $m \geq -4$ put

$$\mathcal{S}_m[A, Y](k_0) = \sum_{\substack{n, b \equiv 1 \pmod{3} \\ n \text{ squarefree}}} \frac{A(n, b) \chi_{k_0}(nb)^3}{3^{m/3} \sqrt{N(n)} N(b)} V_*^\# \left(\frac{3^m N(n) N(b)^3 \mathcal{H}_0^2}{Y N(k_0)^2} \right). \quad (\text{B.50})$$

We shall show, after partitioning k_0 into fixed ray classes, that

$$T(X; u_0 t k_0, g) = \sum_{\iota \in \mathcal{I}} \sum_{m \geq -4} c_{\iota, m}(k_0) \mathcal{S}_m[A_{\iota, m}, Y_\iota](k_0),$$

where $\mathcal{I}$ is independent of k_0 , $|\mathcal{I}| \ll 1$, $|c_{\iota, m}(k_0)| \ll 1$, $A_{\iota, m}$ is independent of k_0 , and for some $a_\iota, Y_\iota > 0$

$$|A_{\iota, m}(n, b)| \ll a_\iota, \quad a_\iota^2 \leq 1, \quad a_\iota^2 Y_\iota \ll \frac{\mathcal{H}_0^2}{X} N(t)^2 N(g); \quad (\text{B.51})$$

all column restrictions are included by extending $A_{\iota, m}$ by zero.

The transformed terms. Put $k = u_0 t k_0$ and $\mathcal{P} = \{p \notin S : p \mid kg\}$. Sextic reciprocity (B.7) gives

$$\Psi_k(n) = \Psi_0(n) \prod_{p \in \mathcal{P}} \chi_p^{j_p}(n), \quad j_p \equiv v_p(k) + 4v_p(g) \pmod{6}, \quad 0 \leq j_p \leq 5, \quad (\text{B.52})$$

where even for $j_p = 0$ the factor $\chi_p^0(n) = \mathbf{1}_{p \nmid n}$ keeps the zero extension at $p \mid kg$, and Ψ_0 contains ξ , the factors at S , the unit factors and $n \mapsto \mathcal{R}(n, \prod_{p \notin S} p^{v_p(k)})$. The fourth power at g contributes no reciprocity sign and the factors at S depend only on exponents modulo six, so Ψ_0 ranges over a fixed finite family. In the notation of Lemma B.14 take $\mathcal{P}_{\text{fix}} = \{p \notin S : p \mid tg\}$, so $\mathcal{P} = \mathcal{P}_{\text{fix}} \sqcup \{p : p \mid k_0\}$ with the exponents on $\mathcal{P}_{\text{fix}}$ fixed by t, g . Partitioning k_0 into fixed ray classes fixes Ψ_0 and, by the lemma, the data d, ψ, c_0 of each transformed term.

¹⁸It is here that the passage from sextic to quadratic twists pays: the large sieve for quadratic characters has the factor $\mathcal{H} + U$, while the general higher-order large sieve of Blomer, Goldmakher and Louvel [5, Theorem 1.3] carries the additional term $(\mathcal{H}U)^{2/3}$, which would not give the mean-square estimate needed.

Fix one transformed term, with active primes $\mathcal{A}_{\text{fix}}$ away from k_0 , and put $c_* = c_0 \prod_{p \in \mathcal{A}_{\text{fix}}} p$, so $c = c_* k_0$ in (B.45). Write $\ell = u\lambda^m n b^3$ as in (B.46). Every prime $p \mid k_0$ has $j_p = 1$, so

$$B_{p,1}(u\lambda^{m+4}nb^3) = \chi_p(u\lambda^{m+4})^3 \chi_p(nb)^3, \quad (\text{B.53})$$

because the exponent $-1 - 2 \equiv 3 \pmod{6}$ and $\chi_p(b)^9 = \chi_p(b)^3$, also for $p \mid b$ by the zero extension. The coefficient bound and the decay of Lemma B.15 make the transformed series absolutely convergent, so its terms may be regrouped. Set

$$A_{u,m}(n, b) = \frac{d(u\lambda^m n b^3) \alpha(u\lambda^m n b^3) \psi(u\lambda^{m+4} n b^3)}{3^{m/6} \sqrt{N(b)}}, \quad |A_{u,m}(n, b)| \leq 27.$$

The transformed term is then

$$\begin{aligned} C(k_0) \sum_{u \in \mathcal{O}^\times} \sum_{m \geq -4} \chi_{k_0}(u\lambda^{m+4})^3 \sum_{\substack{n, b \equiv 1 \pmod{3} \\ n \text{ squarefree}}} \frac{A_{u,m}(n, b) \chi_{k_0}(nb)^3}{3^{m/3} \sqrt{N(n)} N(b)} \\ \times \prod_{p \in \mathcal{A}_{\text{fix}}} B_{p,j_p}(u\lambda^{m+4}nb^3) V_*^\sharp \left(\frac{3^m N(n) N(b)^3 X}{N(c_*)^2 N(k_0)^2} \right), \quad |C(k_0)| \ll 1. \end{aligned}$$

The representation of ℓ is unique, since away from λ the prime exponents of ℓ are $3v_p(b)$ or $1 + 3v_p(b)$. Apart from the quadratic character and the weight, the dependence on k_0 is a bounded scalar.

The active primes away from k_0 . Fix u, m and let $p \in \mathcal{A}_{\text{fix}}$, $q = N(p)$. For $j_p = 4$ the local factor is $B_{p,4}(u\lambda^{m+4}nb^3) = -q^{-1/2} + q^{1/2} \mathbf{1}_{p|n} + q^{1/2} \mathbf{1}_{p \nmid n, p|b}$. Let $A^{(p)}(n, b)$ contain all the other local factors and set $A^{(p,n)}(n, b) = \mathbf{1}_{p \nmid n} A^{(p)}(pn, b)$, $A^{(p,b)}(n, b) = \mathbf{1}_{p \nmid n} A^{(p)}(n, pb)$. The substitutions $n = pn'$ and $b = pb'$ give the exact identity

$$\begin{aligned} \mathcal{S}_m[A^{(p)} B_{p,4}(u\lambda^{m+4} \cdot), q^2 Y](k_0) &= -q^{-1/2} \mathcal{S}_m[A^{(p)}, q^2 Y](k_0) \\ &+ \chi_{k_0}(p)^3 \mathcal{S}_m[A^{(p,n)}, qY](k_0) + q^{-1/2} \chi_{k_0}(p)^3 \mathcal{S}_m[A^{(p,b)}, Y/q](k_0), \end{aligned}$$

where in the second term $p \nmid n'$ preserves squarefreeness, in the third $p \nmid n$ is kept and b' is unrestricted at p , both extracted phases have modulus one, and the factors $q^{1/2}$ cancel against $\sqrt{N(pn')}$ or leave $q^{-1/2}$ after the division by $N(pb')$. Let a track a bound $|A(n, b)| \leq 27a$ and Y the scale in (B.50); before inserting the active primes, $a = 1$ and $Y = N(c_0)^2 \mathcal{H}_0^2 / X$. Each active prime contributes q^2 to the squared conductor norm, and with this factor the updates are

$$(a^2, Y) \mapsto \begin{cases} (a^2, q^2 Y), & j_p \notin \{0, 4\}, \\ (a^2/q, q^2 Y), & j_p = 0, \\ (a^2/q, q^2 Y), (a^2, qY), (a^2/q, Y/q), & j_p = 4. \end{cases}$$

So a^2 never increases. If $p \mid t$ and $p \notin S$ then j_p is odd and $a^2 Y$ costs at most $N(p)^2$; if $p \mid g$ and $p \nmid t$ the cost is $N(p)$ for $j_p = 0, 4$ and $N(p)^2$ for $j_p = 2$, which needs $v_p(g) \geq 2$. Hence

$$a^2 Y \ll \frac{\mathcal{H}_0^2}{X} \prod_{\substack{p \mid t \\ p \notin S}} N(p)^2 \prod_{\substack{p \mid g \\ p \nmid t, p \notin S}} N(p)^{v_p(g)} \leq \frac{\mathcal{H}_0^2 N(t)^2 N(g)}{X}.$$

Reindexing at distinct primes preserves the form and the zero extensions. By Lemma B.14, the transformed terms, the units and the at most three choices at each $p \in \mathcal{A}_{\text{fix}}$ with $j_p = 4$ form an index set whose size is divisor-bounded in tg , meaning bounded by $C \tau_{\text{div}}(tg)^A$ for fixed C, A , hence $\ll_\varepsilon N(tg)^\varepsilon$. This is the representation announced, with (B.51).

The mean square of one prepared sum. Fix ι and abbreviate $a = a_\iota$, $Y = Y_\iota$, $A_m = A_{\iota, m}$; the local updates give $Y \ll D^{C_1}$ with $C_1 = C_1(C_0)$. Split b into $B \leq N(b) < 2B$ and n by a

smooth dyadic partition $V(N(n)/U)$, $U, B \geq 1$, with uniformly bounded cutoffs, and write for this part

$$\begin{aligned}\mathcal{S}_{m;U,B}(k_0) &= 3^{-m/3} \sum_{B \leq N(b) < 2B} \frac{\chi_{k_0}(b)^3}{N(b)} F_b(k_0), \\ F_b(k_0) &= \sum_{n \text{ squarefree}} \frac{A_m(n, b) \chi_{k_0}(n)^3}{\sqrt{N(n)}} V(N(n)/U) V_*^\# \left(\frac{3^m N(n) N(b)^3 \mathcal{H}_0^2}{Y N(k_0)^2} \right).\end{aligned}$$

Put $z_* = 3^m U B^3 / Y$ and fix a decay exponent $A > 0$. With m, U, B, Y fixed, Lemma B.21, applied in the single variable $N(n)/U$, separates the weight from the column:

$$\begin{aligned}F_b(k_0) &= \int_{\mathbb{R}} c_{b,k_0}(s) G_{b,s}(k_0) ds, \quad G_{b,s}(k_0) = \sum_{\substack{n \text{ squarefree} \\ N(n) \asymp U}} \frac{A_m(n, b) \chi_{k_0}(n)^3}{\sqrt{N(n)}} (N(n)/U)^{is}, \\ |c_{b,k_0}(s)| &\ll_{I,A} (1 + z_*)^{-A} (1 + |s|)^{-2},\end{aligned}\tag{B.54}$$

since the scale R of (B.79) is $3^m U N(b)^3 \mathcal{H}_0^2 / (Y N(k_0)^2) \geq z_*$ and Lemma B.15 supplies the derivative bounds, independently of b, k_0, U, B, m . Since $\sum_{N(n) \asymp U} N(n)^{-1} \ll 1$, Lemma B.16 gives, for any $\sigma > 0$, $\sum_{N(k_0) \leq \mathcal{H}_0}^* |G_{b,s}(k_0)|^2 \ll_{S,\sigma} a^2 (\mathcal{H}_0 U)^\sigma (\mathcal{H}_0 + U)$. Weighted Cauchy–Schwarz in the Mellin integral (Lemma B.23) with the common majorant of (B.54), and then in b with $\sum_{B \leq N(b) < 2B} N(b)^{-1} \ll 1$, give

$$\sum_{N(k_0) \leq \mathcal{H}_0}^* |\mathcal{S}_{m;U,B}(k_0)|^2 \ll_{I,S,A,\sigma} a^2 3^{-2m/3} (\mathcal{H}_0 U)^\sigma (\mathcal{H}_0 + U) (1 + z_*)^{-2A}.\tag{B.55}$$

Since $m \geq -4$ we have $U \leq 81Y z_*$, so $\mathcal{H}_0 U \ll D^{C_0+C_1} (1 + z_*)$. Choose $0 < \sigma \leq 1$ small in terms of $\varepsilon_1 > 0$ and C_0 , take $A = 3$, and take square roots:

$$\left(\sum_{N(k_0) \leq \mathcal{H}_0}^* |\mathcal{S}_{m;U,B}(k_0)|^2 \right)^{1/2} \ll a D^{\varepsilon_1} 3^{-m/3} (\sqrt{\mathcal{H}_0} + \sqrt{U}) (1 + 3^m U B^3 / Y)^{-2}.$$

The dyadic sums converge: for dyadic $U, B \geq 1$ and $Z > 0$, $\sum_{U,B} (1 + U B^3 / Z)^{-2} \ll \log^2(2 + Z)$ and $\sum_{U,B} \sqrt{U} (1 + U B^3 / Z)^{-2} \ll \sqrt{Z}$. The triangle inequality in $\ell^2(k_0)$ and the geometric sum over $m \geq -4$ give

$$\left(\sum_{N(k_0) \leq \mathcal{H}_0}^* \left| \sum_{m \geq -4} c_{\iota,m}(k_0) \mathcal{S}_m[A_m, Y](k_0) \right|^2 \right)^{1/2} \ll a D^{\varepsilon_1} (\sqrt{\mathcal{H}_0} \log^2(2 + Y) + \sqrt{Y}).$$

Combining the divisor-bounded choices of ι with (B.51),

$$\sum_{N(k_0) \leq \mathcal{H}_0}^* |T(X; u_0 t k_0, g)|^2 \preceq \max_{\iota} a_{\iota}^2 (\mathcal{H}_0 + Y_{\iota}) \ll \mathcal{H}_0 + \frac{\mathcal{H}_0^2 N(t)^2 N(g)}{X} \leq \mathcal{H} + \frac{\mathcal{H}^2 N(g)}{X}.$$

Sum over the fixed ray classes and over the divisor-bounded choices $t \mid \text{rad}(g \prod_{\mathfrak{p} \in S} \mathfrak{p})$, choosing ε_1 and the other small losses in terms of ε . The weight estimates use a fixed number $J = J(\varepsilon, C_0)$ of derivatives, and homogeneity restores $\|W\|_{C^J(I)}^2$. $\square$

Proof of Proposition B.9. Write $k = u_0 s v^2$ uniquely with s squarefree and s, v among the chosen generators, no condition $(s, v) = 1$ imposed. Since $8 \equiv 2 \pmod{6}$, with the zero extensions, $\chi_n(u_0 s v^2) \chi_n(f)^4 = \chi_n(u_0 s) \chi_n(f v^2)^4$, so $T(X; u_0 s v^2, f) = T(X; u_0 s, f v^2)$. Apply Lemma B.17 with the row bound $\mathcal{H}/N(v)^2$ and $g = f v^2$, where $N(g) \leq \mathcal{H} N(f) \leq D^{2C_0}$, so

with $2C_0$ in place of C_0 :

$$\begin{aligned} \sum_{0 < N(k) \leq \mathcal{H}} |T(X; k, f)|^2 &= \sum_{u_0 \in \mathcal{O}^\times} \sum_{N(v)^2 \leq \mathcal{H}} \sum_{\substack{N(s) \leq \mathcal{H}/N(v)^2 \\ s \text{ squarefree}}} |T(X; u_0 s, f v^2)|^2 \\ &\ll D^\varepsilon \|W\|_{C^J(I)}^2 \left(\mathcal{H} + \frac{\mathcal{H}^2 N(f)}{X} \right) \sum_v \frac{1}{N(v)^2}, \end{aligned}$$

and the last sum is $\zeta_K(2) < \infty$. $\square$

B.13. Proof of the transfer proposition. Throughout this section $\mathcal{H}, L, F, \Sigma, \xi$ satisfy the hypotheses of Proposition B.11, Φ and C_Φ are the weight and support bound fixed before (B.34), $N = N$, $I = [u, v]$ and $I_* = [u/2, 2v]$. The proof applies Poisson summation twice. The first application turns the Gauss-sum coefficients a_ξ into Möbius coefficients and produces a mean square of the shape (B.34) with the roles of rows and columns exchanged; the second turns the Möbius coefficients back into Gauss-sum coefficients and returns to the family (B.16), with a row range at most $\mathcal{H}L/(\Sigma F)$. The enlargement by the factor $\Sigma/(LF) \geq 1$ in the intermediate mean square is what makes the second step land in the required ranges.¹⁹ All ideal indices are prime to S and are represented by their primary generators; a star on a sum additionally requires squarefreeness; the variables k, h, y range over $\mathcal{O}$.

First application of Poisson summation. For squarefree coprime C, t and $d \mid C$ put

$$\ell = \frac{L}{N(C)N(t)}, \quad w_{C,d} = \frac{\mathcal{H}N(C)}{N(d)L^2F}, \quad Y_{C,d} = c_I \frac{\Sigma L F N(d)}{\mathcal{H}N(C)^2},$$

where $c_I \geq \max(1, 4C_\Phi v^2)$ is fixed, and retain only $\ell \geq 1/(2v)$. For $U \in C_c^\infty(I_*)$ and a character ξ_1 of the fixed ray class group define

$$p_y(n) = \mu(n) \xi_1(n) \mathbf{1}_{(n,t)=1} \overline{\chi_n(y)} \chi_n(C)^4 \chi_n(d), \quad P_{C,d,t}(y; U) = \sum_n^* p_y(n) U(N(n)/\ell), \quad (\text{B.56})$$

the notation p_y suppressing the dependence on C, d, t, ξ_1 , and the nonnegative form

$$\mathcal{Q}_{\xi_1}(U) = \sum_{\substack{C, t \text{ squarefree} \\ (Ct, S)=1, (C, t)=1 \\ N(C)N(t) \leq 2vL}} \sum_{d \mid C} w_{C,d} \sum_y \Phi(N(y)/Y_{C,d}) |P_{C,d,t}(y; U)|^2. \quad (\text{B.57})$$

The outer sums are finite and the sum over y converges absolutely; $Y_{C,d}$ may be smaller than 1.

Lemma B.18. *Fix an integer $j \geq 0$ and $\varepsilon_0 > 0$. If $M \geq 0$ satisfies $\mathcal{Q}_{\xi_1}(U) \leq M \|U\|_{C^j(I_*)}^2$ for every $U \in C_c^\infty(I_*)$ and every ξ_1 , then $\mathcal{A}(W) \ll D^{\varepsilon_0} (\Sigma + M) \|W\|_{C^{2j+4}(I)}^2$.*

Proof. Expand the square in (B.34) and write $n_i = Cu_i$ with $C = (n_1, n_2)$, so that $(u_1, u_2) = (u_1 u_2, C) = 1$. The row character is $\chi_{u_1} \overline{\chi_{u_2}}$, primitive modulo $u_1 u_2$, with the extra restriction $(k, C) = 1$. Lemma B.4 introduces $d \mid C$ and a frequency h . The zero frequency requires $u_1 = u_2 = 1$, so its contribution Z satisfies

$$|Z| \ll \frac{\mathcal{H}}{LF} \sum_{F \leq N(f) < 2F}^* \sum_C^* |W(N(C)/L)|^2 \ll_I \mathcal{H} \|W\|_\infty^2.$$

For the other frequencies, the Chinese remainder theorem, (B.10) and (B.13) give the paired identity

$$a_\xi(u_1) \overline{a_\xi(u_2)} \gamma(\chi_{u_1} \overline{\chi_{u_2}}) = \mu(u_1) \mu(u_2) (\xi G)(u_1 u_2^{-1}). \quad (\text{B.58})$$

¹⁹This is the mechanism behind the second inequality of (B.36): the new ratio $\mathcal{H}'/\Sigma'$ equals $\mathcal{H}/(\Sigma F R)$ for a dyadic $R \geq 1$, so it never exceeds the old ratio $\mathcal{H}/\Sigma$. The induction of Section B.8 needs exactly this monotonicity.

Expand $(\xi G)(z) = \sum_{\xi_1} \ell_{\xi_1} \xi_1(z)$ on the fixed ray class group, insert $\mathbf{1}_{(u_1, u_2)=1} = \sum_{t|u_1, t|u_2} \mu(t)$ and put $u_i = tx_i$; squarefreeness imposes $(x_i, t) = 1$ but no condition $(x_1, x_2) = 1$ remains. The factors at C, t cancel against their conjugates, leaving $\mu(d)\mu(t)$ and the restrictions $(f, Ct) = (h, t) = 1$. Thus

$$\begin{aligned} \mathcal{A}(W) - Z &= \sum_{\xi_1} \ell_{\xi_1} \sum_{\substack{C, t \text{ squarefree} \\ (Ct, S)=1, (C, t)=1 \\ N(C)N(t) \leq 2vL}} \sum_{d|C} \mu(d)\mu(t) w_{C,d} \sum_{\substack{F \leq N(f) < 2F \\ (f, Ct)=1}}^* \sum_{\substack{h \neq 0, (h, t)=1 \\ N(h) \leq C_{\Phi} v^2 L^2 N(d)/(\mathcal{H}N(C)^2)}} \\ &\quad \times \sum_{x_1, x_2}^* p_{hf^2}(x_1) \overline{p_{hf^2}(x_2)} \mathcal{K}_h \left(\frac{N(x_1)}{\ell}, \frac{N(x_2)}{\ell} \right), \end{aligned}$$

where, for every $j_0 \geq 0$,

$$\begin{aligned} \mathcal{K}_h(z_1, z_2) &= (z_1 z_2)^{-1/2} W(z_1) \overline{W(z_2)} \widehat{\Phi} \left(\frac{\mathcal{H}N(h)N(C)^2}{L^2 N(d) z_1 z_2} \right), \\ \sup_{C, d, h} \|\mathcal{K}_h\|_{C^{j_0}(I^2)} &\ll_{I, \Phi, j_0} \|W\|_{C^{j_0}(I)}^2. \end{aligned}$$

Here $\chi_{x_i}(C)^4$ supplies $(x_i, C) = 1$, the other zeros are supplied by $p_{hf^2}(x_i)$, and the range of h comes from the support of $\widehat{\Phi}$. For fixed C, d, t the map $(f, h) \mapsto y = hf^2$ has divisor-bounded multiplicity, and since $\Sigma/(LF) \geq 1$ its nonzero image satisfies $N(y) \leq 4C_{\Phi} v^2 L^2 F^2 N(d)/(\mathcal{H}N(C)^2) \leq Y_{C,d}$. As $\Phi \geq 1$ on $[0, 1]$ and $\Phi \geq 0$,

$$\sum_{C, t} \sum_{d|C} w_{C,d} \sum_{f, h} |P_{C,d,t}(hf^2; U)|^2 \preceq \mathcal{Q}_{\xi_1}(U) \leq M \|U\|_{C^j(I_*)}^2$$

with the ranges above on the left. Lemma B.22, applied to the displayed expression for $\mathcal{A}(W) - Z$ with the kernel bound, gives $|\mathcal{A}(W) - Z| \ll D^{\varepsilon_0/2} M \|W\|_{C^{2j+4}(I)}^2$; together with $\mathcal{H} \leq \Sigma$ this proves the lemma, the divisor losses being allocated within ε_0 . $\square$

Second application of Poisson summation. Let C, t be coprime squarefree primary elements prime to S , $d | C$, ξ_1 a character of the ray class group, $\ell, Y > 0$ arbitrary scales and $U \in C_c^\infty(I_*)$ with $I_* = [a, b] \subset (0, \infty)$. Put

$$\begin{aligned} P(y) &= \sum_{(n, S)=1}^* \mu(n) \xi_1(n) \mathbf{1}_{(n, t)=1} \overline{\chi_n(y)} \chi_n(C)^4 \chi_n(d) U(N(n)/\ell), \\ \mathcal{M} &= \sum_{y \in \mathcal{O}} \Phi(N(y)/Y) |P(y)|^2, \end{aligned} \tag{B.59}$$

so that $P(y) = P_{C,d,t}(y; U)$ with the scale ℓ now arbitrary. Put $U_0(x) = x^{-1/2} U(x)$ and expand $\xi_1(z) G(z^{-1}) = \sum_{\xi'} c_{\xi'} \xi'(z)$ on the fixed ray class group. On the transformed side g is the common divisor of the original columns, $e | g$ comes from the row exclusion, w removes the remaining coprimality condition and h is the nonzero Fourier frequency: g, w range over squarefree primary elements prime to S , e over the divisors of g and h over $\mathcal{O} \setminus \{0\}$, subject to

$$(g, Ct) = (w, gCth) = 1, \quad N(gw) \leq b\ell, \quad N(h) \leq \frac{C_{\Phi} b^2 \ell^2 N(e)}{Y N(g)^2},$$

conditions which make tg/e and Cew coprime and squarefree. For each such choice and each ξ' the new column scale, coefficient and coupled weight are

$$\begin{aligned} X' &= \frac{\ell}{N(gw)}, \quad a^{\sharp}(n) = a_{\xi'}(n) \mathbf{1}_{(n, tg/e)=1} \chi_n(deh) \chi_n(Cew)^4, \\ \mathcal{K}(x_1, x_2) &= U_0(x_1) \overline{U_0(x_2)} \widehat{\Phi} \left(\frac{Y N(h) N(g)^2}{N(e) \ell^2 x_1 x_2} \right). \end{aligned} \tag{B.60}$$

Define also $Z = Y \widehat{\Phi}(0) \sum_{(g, CtS)=1}^* |U(N(g)/\ell)|^2 \prod_{p|g} (1 - N(p)^{-1})$.

Lemma B.19 (second Poisson identity). *With these index ranges, and a star on n_1, n_2 restricting each index separately to squarefree primary elements,*

$$\mathcal{M} = Z + \sum_{g,e,w,h,\xi'} \frac{Y\mu(e)\mu(w)c_{\xi'}N(g)}{N(e)\ell} \sum_{(n_1n_2,S)=1}^* a^\#(n_1)\overline{a^\#(n_2)} \mathcal{K}\left(\frac{N(n_1)}{X'}, \frac{N(n_2)}{X'}\right), \quad (\text{B.61})$$

$$|Z| \ll_{I_*, \Phi} Y\ell \|U\|_\infty^2.$$

Proof. Expand the square, put $g = (n_1, n_2)$ and $n_i = gz_i$. The nonzero terms satisfy $(z_1, z_2) = (z_1z_2, gCt) = (g, Ct) = 1$ and $\chi_{n_1}(y)\chi_{n_2}(y) = \mathbf{1}_{(y,g)=1}\chi_{z_1}(y)\chi_{z_2}(y)$. Apply Lemma B.4 to this character with exclusion g and divisor $e \mid g$. Its zero frequency occurs exactly when $z_1 = z_2 = 1$ and gives Z , whose bound follows by counting g . For the nonzero frequencies, (B.11), (B.13) and the Chinese remainder theorem give

$$\mu(z_1)\mu(z_2)\xi_1(z_1)\overline{\xi_1(z_2)}\gamma(\overline{\chi_{z_1}\chi_{z_2}}) = \sum_{\xi'} c_{\xi'} a_{\xi'}(z_1)\overline{a_{\xi'}(z_2)},$$

and moreover $\chi_z(dh)\overline{\chi_z(e)}\chi_z(C)^4 = \chi_z(deh)\chi_z(Ce)^4$ and

$$\frac{U(N(gz_1)/\ell)\overline{U(N(gz_2)/\ell)}}{\sqrt{N(z_1)N(z_2)}} = \frac{N(g)}{\ell} U_0\left(\frac{N(gz_1)}{\ell}\right) \overline{U_0\left(\frac{N(gz_2)}{\ell}\right)}.$$

Hence

$$\begin{aligned} \mathcal{M} - Z &= \sum_{\xi'} c_{\xi'} \sum_{(g,Ct)=1}^* \sum_{e \mid g} \frac{Y\mu(e)N(g)}{N(e)\ell} \sum_{\substack{h \neq 0_{z_1, z_2}, \\ (z_1, z_2)=1 \\ (z_1z_2, gCt)=1}} \sum_{(z_1, z_2)=1}^* a_{\xi'}(z_1)\overline{a_{\xi'}(z_2)} \\ &\quad \times \chi_{z_1}(deh)\overline{\chi_{z_2}(deh)}\chi_{z_1}(Ce)^4\overline{\chi_{z_2}(Ce)^4} \\ &\quad \times U_0\left(\frac{N(gz_1)}{\ell}\right)\overline{U_0\left(\frac{N(gz_2)}{\ell}\right)} \widehat{\Phi}\left(\frac{YN(h)}{N(e)N(z_1)N(z_2)}\right). \end{aligned}$$

Insert $\mathbf{1}_{(z_1, z_2)=1} = \sum_{w \mid z_1, w \mid z_2} \mu(w)$ and put $z_i = wn_i$, with $(w, gCt) = 1$. By (B.12) and the zero-extended characters,

$$\begin{aligned} a_{\xi'}(wn) &= a_{\xi'}(w)a_{\xi'}(n)\chi_n(w)^4 \quad ((w, n) = 1), \quad |a_{\xi'}(w)\chi_w(deh)\chi_w(Ce)^4|^2 = \mathbf{1}_{(w, h)=1}, \\ \mathbf{1}_{(n, gCtw)=1}\chi_n(deh)\chi_n(Cew)^4 &= \mathbf{1}_{(n, tg/e)=1}\chi_n(deh)\chi_n(Cew)^4, \end{aligned}$$

so that

$$\begin{aligned} \mathcal{M} - Z &= \sum_{\xi'} c_{\xi'} \sum_{\substack{g, w \\ (g, Ct)=(w, gCt)=1}}^* \sum_{e \mid g} \frac{Y\mu(e)\mu(w)N(g)}{N(e)\ell} \sum_{\substack{h \neq 0 \\ (h, w)=1}} \sum_{(n_1n_2, S)=1}^* a^\#(n_1)\overline{a^\#(n_2)} \\ &\quad \times U_0\left(\frac{N(gwn_1)}{\ell}\right)\overline{U_0\left(\frac{N(gwn_2)}{\ell}\right)} \widehat{\Phi}\left(\frac{YN(h)}{N(e)N(w)^2N(n_1)N(n_2)}\right), \end{aligned}$$

with no remaining condition $(n_1, n_2) = 1$. The column support gives $N(gw) \leq b\ell$ and $N(z_1z_2) \leq b^2\ell^2/N(g)^2$, so the support of $\widehat{\Phi}$ imposes the stated bound on $N(h)$; substituting (B.60) gives (B.61), with finitely many terms. $\square$

Lemma B.20. *Let $I = [u, v]$, $I_* = [u/2, 2v]$, $m \geq 0$ an integer and S_m the supremum in (B.35), over test functions supported in $I' = [u/16, 4v]$ with $C^m(I')$ norm at most one, an empty supremum being zero. Then for every $\varepsilon_0 > 0$ and $U \in C_c^\infty(I_*)$,*

$$\mathcal{Q}_{\xi_1}(U) \ll D^{\varepsilon_0} \Sigma(1 + S_m) \|U\|_{C^{2m+4}(I_*)}^2.$$

Proof. Apply Lemma B.19 with input length ℓ , row scale $Y_{C,d}$, the same C, d, t and the character ξ_1 ; h is now the Fourier variable of the sum over y . The nonzero output has squarefree $g, w, e \mid g$, $(g, Ct) = (w, gCth) = 1$, and with

$$r = tg/e, \quad f' = Cew, \quad k' = deh$$

the pair r, f' is coprime and squarefree and $a^\sharp(n) = a_{\xi'}(n) \mathbf{1}_{(n,r)=1} \chi_n(k') \chi_n(f')^4$. Writing X'_0 for the individual column scale X' of (B.60), the length, the coefficient and the Fourier parameter simplify to

$$X'_0 = \frac{\ell}{N(g)N(w)} = \frac{L}{N(r)N(f')}, \quad w_{C,d} \frac{Y_{C,d}N(g)}{N(e)\ell} = \frac{c_I \Sigma N(r)}{L^2},$$

$$\frac{Y_{C,d}N(h)N(g)^2}{N(e)\ell^2} = \frac{c_I \Sigma F N(k') N(r)^2}{\mathcal{H}L}.$$

The coupled kernel is therefore

$$\mathcal{K}_{r,f',k'}(x_1, x_2) = U_0(x_1) \overline{U_0(x_2)} \widehat{\Phi} \left(\frac{c_I \Sigma F N(k') N(r)^2}{\mathcal{H}L x_1 x_2} \right), \quad x_j = N(n_j)/X'_0,$$

and its nonzero support requires $N(r)N(f') \leq 2vL$ and

$$0 < N(k') \leq \frac{4C_\Phi v^2}{c_I} \frac{\mathcal{H}L}{\Sigma F N(r)^2} \leq \frac{\mathcal{H}L}{\Sigma F N(r)^2},$$

since $C_\Phi(2v)^2/c_I \leq 1$.

For fixed r, f', k' all preimages are obtained by choosing $t \mid r$, $f' = Cew$ with $e \mid k'$, $(w, k') = 1$ and $d \mid (C, k')$, and setting $g = e(r/t)$, $h = k'/(de)$; the factorisation $f' = Cew$ is disjoint and squarefree. All preimages have the same kernel, since $Ct gw = rf'$ and the Fourier parameter depends only on r, k' , so the support restrictions discard only zero kernels. The choices $t \mid r$ contribute $\tau_{\text{div}}(r)$. At a prime $p \mid f'$ the choices $p \mid C$, $p \mid e$, $p \mid w$ contribute $(1 + \mathbf{1}_{p \mid k'}) - \mathbf{1}_{p \mid k'} - \mathbf{1}_{p \nmid k'} = \mathbf{1}_{p \mid k'}$ to the sum of $\mu(e)\mu(w)$ over the preimages, the two choices inside the first term being $p \nmid d$ and $p \mid d$.²⁰ Let $\mathcal{Z}$ be the total zero-frequency contribution, the sum of the terms Z of Lemma B.19 with the weights $w_{C,d}$ and the ranges of (B.57). Regrouping before taking absolute values,

$$\mathcal{Q}_{\xi_1}(U) - \mathcal{Z} = \frac{c_I \Sigma}{L^2} \sum_{\xi'} c_{\xi'} \sum_{\substack{r, f' \text{ squarefree} \\ (r, f')=1}} N(r) \tau_{\text{div}}(r) \sum_{\substack{k' \neq 0 \\ f' \mid k'}} \times \sum_{n_1, n_2}^* a^\sharp(n_1) \overline{a^\sharp(n_2)} \mathcal{K}_{r,f',k'} \left(\frac{N(n_1)}{X'_0}, \frac{N(n_2)}{X'_0} \right).$$

Fix dyadic ranges $R \leq N(r) < 2R$ and $F' \leq N(f') < 2F'$ and put

$$X' = \frac{L}{RF'}, \quad \Sigma' = X'F' = \frac{L}{R}, \quad \mathcal{H}' = \frac{\mathcal{H}L}{\Sigma FR^2}.$$

Retain only the r for which a nonzero term occurs in these ranges. Since $f' \mid k'$, each such r satisfies $F' \leq N(f') \leq N(k') \leq \mathcal{H}L/(\Sigma F N(r)^2)$; in particular $\mathcal{H}' \geq 1$, and for every $j \mid r$, using $\mathcal{H} \leq \Sigma$,

$$\frac{X'}{N(j)} \geq \frac{\Sigma F N(r)^2}{\mathcal{H} R N(j)} \geq \frac{F \Sigma}{\mathcal{H}} \geq 1. \quad (\text{B.62})$$

Moreover $1 \leq X'/X'_0 < 4$. The rescaled kernels $\mathcal{K}_{r,f',k'}(x_1, x_2) = \mathcal{K}_{r,f',k'}(\rho x_1, \rho x_2)$, $\rho = X'/X'_0$, have support in $[u/8, 2v]^2 \subset \text{int}(I'^2)$ and satisfy, over the retained indices of the block,

$$\sup \|\mathcal{K}_{r,f',k'}\|_{C^{2m+4}(I'^2)} \ll_{I^*, m, \Phi} \|U\|_{C^{2m+4}(I^*)}^2,$$

by the Schwartz bounds for $\widehat{\Phi}$. The common parameters satisfy exactly

$$\frac{\mathcal{H}'}{\Sigma'} = \frac{\mathcal{H}}{\Sigma FR} \leq \frac{\mathcal{H}}{\Sigma}, \quad \Sigma' \leq L, \quad \mathcal{H}' \leq \frac{\mathcal{H}L}{\Sigma F}.$$

²⁰So the sum over preimages vanishes unless $f' \mid k'$, and then it equals $\tau_{\text{div}}(r)$: the signs of the two Poisson summations cancel the Möbius factors introduced by the coprimality conditions. This is why the transfer loses only a divisor function.

Fix $\delta > 0$. For each retained r enlarge the f', k' ranges by positivity, keeping the set of r fixed. Lemma B.6 then gives, for $V \in C_c^\infty(I')$,

$$\sum_{F' \leq N(f') < 2F'}^* \sum_{0 < N(k') \leq \mathcal{H}'} \left| \sum_n^* a^\sharp(n) V(N(n)/X') \right|^2 = \Sigma' \mathcal{E}_{(r)}(\mathcal{H}', X', F'; \xi', V) \\ \ll_\delta D^\delta (\Sigma')^2 S_m \|V\|_{C^m(I')}^2 :$$

the removal of the exclusion replaces (X', F') by $(X'/N(j), F'N(j))$, $j \mid r$, preserving $\mathcal{H}'$ and Σ' , and (B.62) gives $X'/N(j) \geq 1$ and $F'N(j) \geq 1$, so every resulting mean square lies in the defining supremum. There are $O(R)$ ideals r in the dyadic range. Lemma B.22 with row index (r, f', k') , the preceding mean-square bound and the uniform kernel bound show that the block contributes at most

$$D^{2\delta} \frac{\Sigma R}{L^2} R (\Sigma')^2 S_m \|U\|_{C^{2m+4}(I_*)}^2 = D^{2\delta} \Sigma S_m \|U\|_{C^{2m+4}(I_*)}^2,$$

because $(\Sigma R/L^2)R(L/R)^2 = \Sigma$. Finally, the zero-frequency bound of Lemma B.19 gives

$$|\mathcal{Z}| \ll_{I_*, \Phi} \|U\|_\infty^2 \sum_{N(Ct) \leq 2vL} \sum_{d \mid C} w_{C,d} Y_{C,d} \ell \\ \ll_{I_*, \delta} \Sigma \|U\|_\infty^2 \sum_C \frac{\tau_{\text{div}}(C)}{N(C)^2} \sum_{N(t) \leq 2vL} \frac{1}{N(t)} \ll_{I_*, \delta} D^\delta \Sigma \|U\|_\infty^2,$$

which uses only $Y_{C,d} > 0$. Sum the $O_{I_*, C_0}(\log^2(2D))$ dyadic blocks and the finite set of characters, and take $\delta = \varepsilon_0/4$. $\square$

Proof of Proposition B.11. Lemma B.20, with the loss $D^{\varepsilon/4}$, supplies the hypothesis of Lemma B.18 with $j = 2m + 4$ and with $M \ll D^{\varepsilon/4} \Sigma(1 + S_m)$; that lemma with the same loss gives $\mathcal{A}(W) \ll D^{\varepsilon/2} \Sigma(1 + S_m) \|W\|_{C^{4m+12}(I)}^2$, which is (B.35). $\square$

B.14. The theta transformation with fixed ray class twists. We prove Proposition B.13, Lemma B.14 and Lemma B.15. The argument is the theta-transformation method of Dunn and Radziwiłł [8, §5 and Appendix A], which extends Patterson [26] and Yoshimoto [28]; what is added is the treatment of the fixed ray class twists, in which the residue h_0 below need not be prime to the modulus, and the uniformity statements. We use the setting preceding the proposition: Ψ_0 and S are fixed, $\mathcal{P}$ is the varying finite set of primes outside S and j_p are the exponents of their character factors in Ψ . Character powers follow the zero-extension convention of the footnote in Section B.1, so that $\chi_p^0(x) = \mathbf{1}_{p \nmid x}$, also when we write $\chi_p(x)^j$. Write $w = (z, v)$ for a point of upper half-space, θ for the function (B.37), and $\check{e}(z) = \exp(2\pi i(z + \bar{z}))$ for the additive character of the Fourier expansions, so that $e(z) = \check{e}(z/\lambda)$.

Proof of Proposition B.13 and Lemmas B.14 and B.15. Finite Fourier expansion. Define $\phi(n) = \chi_n(\lambda)^2 \Psi_0(n)$ for $n \equiv 1 \pmod{3}$ prime to S and $\phi(n) = 0$ otherwise, and for $\Psi(n) = \Psi_0(n) \prod_{p \in \mathcal{P}} \chi_p^{j_p}(n)$ put

$$\Theta_\Psi(z, v) = \sum_{\substack{\ell \in \lambda^{-3}\mathcal{O} \\ \ell \neq 0}} \overline{\tau(-\ell)} \phi(\lambda^3 \ell) \left(\prod_{p \in \mathcal{P}} \chi_p^{j_p}(\lambda^3 \ell) \right) v K_{1/3}(4\pi|\ell|v) \check{e}(\ell z), \quad (\text{B.63})$$

which recovers Θ_k of (B.39) when $\Psi = \Psi_k$. Choose once and for all a nonzero $L \in \mathcal{O}$ with prime divisors in S , divisible by the conductor of Ψ_0 , by every prime of S and by sufficiently high powers of the primes above 2 and 3. The supplementary law of cubic reciprocity for λ [8, (1.5)], which evaluates $\chi_n(\lambda)^2 = (\lambda/n)_3$, makes ϕ periodic modulo L .

Put $\widehat{\phi}(h_0) = N(L)^{-1} \sum_{x \bmod L} \phi(x) e(-h_0 x/L)$ for $h_0 \in \mathcal{O}/(L)$, so $|\widehat{\phi}(h_0)| \leq 1$. For $p \in \mathcal{P}$ and $0 \leq j \leq 5$ expand χ_p^j on $\mathcal{O}/(p)$ in additive characters,

$$C_{p,j}(h_p) = \frac{1}{N(p)} \sum_{y \bmod p} \chi_p^j(y) e(-h_p y/p), \quad \chi_p^j(x) = \sum_{h_p \in \mathcal{O}/(p)} C_{p,j}(h_p) e(h_p x/p). \quad (\text{B.64})$$

Multiplying over $p \in \mathcal{P}$ and evaluating at $x = \lambda^3 \ell$ turns each additive character into a shift of $\bar{\theta}$: writing $\mathbf{h} = (h_0, (h_p)_{p \in \mathcal{P}})$,

$$\Theta_\Psi(z, v) = \sum_{\mathbf{h}} c_F(\mathbf{h}) \overline{\theta(z + z_{\mathbf{h}}, v)}, \quad z_{\mathbf{h}} = \lambda^2 \left(\frac{h_0}{L} + \sum_{p \in \mathcal{P}} \frac{h_p}{p} \right), \quad c_F(\mathbf{h}) = \widehat{\phi}(h_0) \prod_{p \in \mathcal{P}} C_{p,j_p}(h_p). \quad (\text{B.65})$$

Indeed, expanding the fixed factor ϕ as well, the finite Fourier expansions identify all nonzero Fourier coefficients, and the constant terms of the translates cancel because $\sum_{\mathbf{h}} c_F(\mathbf{h}) = \phi(0) \prod_p \chi_p^{j_p}(0) = 0$. For $j \not\equiv 0 \pmod{6}$ the change $y \mapsto -y$ gives $N(p)^{-1/2} \sum_{y \bmod p} \chi_p(y)^j e(-y/p) = \chi_p(-1)^j \gamma_j(p)$ with $|\gamma_j(p)| = 1$; substituting $y = h_p^{-1} u$ for $h_p \neq 0$, using orthogonality for $h_p = 0$, $j \neq 0$, and summing the additive character over the nonzero residues for $j = 0$,

$$C_{p,j}(h_p) = \begin{cases} N(p)^{-1/2} \chi_p(-1)^j \gamma_j(p) \chi_p(h_p)^{-j}, & j \neq 0, h_p \neq 0, \\ 0, & j \neq 0, h_p = 0, \\ -N(p)^{-1}, & j = 0, h_p \neq 0, \\ 1 - N(p)^{-1}, & j = 0, h_p = 0. \end{cases} \quad (\text{B.66})$$

Call $\mathcal{A}(\mathbf{h}) = \{p \in \mathcal{P} : h_p \neq 0\}$ the set of active primes of $\mathbf{h}$. By (B.66), $c_F(\mathbf{h}) \neq 0$ forces $\{p : j_p \neq 0\} \subseteq \mathcal{A}(\mathbf{h})$, so only primes with $j_p = 0$ can be inactive, as the proposition asserts. Put $r = \prod_{p \in \mathcal{A}(\mathbf{h})} p$ and let c_0 be the reduced denominator of $\lambda^2 h_0/L$. At each active prime $v_p(\lambda^2 h_p/p) = -1$ while $v_p(z_{\mathbf{h}} - \lambda^2 h_p/p) \geq 0$, so p cannot cancel from the denominator, and at the primes dividing L all the terms $\lambda^2 h_p/p$ are integral. Hence the reduced denominator of $z_{\mathbf{h}}$ is $c = c_0 r$ up to a unit, also when $(h_0, L) \neq 1$. The expansion (B.65) thus writes Θ_Ψ as $O_L(2^{|\mathcal{P}|})$ groups of translates, indexed by $h_0 \in \mathcal{O}/(L)$ and the active set $\mathcal{A}$, each group with a common reduced denominator $c = c_0 \prod_{p \in \mathcal{A}} p$ up to a unit.

Reduced denominators and cusp coefficients. Each translate $\overline{\theta(z + z_{\mathbf{h}}, v)}$ is now expressed through one of the three cusp expansions (B.44). The representatives $\gamma_0, \gamma_+, \gamma_-$ of (B.43) are $\gamma_1, \gamma_{10}, \gamma_{19}$ in the numbering of [8, §5.1], which follows [26, Table II]. Set $M = \lambda^{12} L^4$. Since $(p, M) = 1$ for active p , the Chinese remainder theorem lets us choose representatives $h_p \in \mathcal{O}$ with $h_p \equiv 0 \pmod{M^2}$; changing a representative modulo p changes $z_{\mathbf{h}}$ by an element of $\lambda^2 \mathcal{O} = 3\mathcal{O}$, under which θ is periodic. To track the dependence on the primes of r , restrict r to a residue class modulo M^2 . For fixed h_0 , active set and residue class write $z_{\mathbf{h}} = a/c$ in lowest terms, with $a \equiv 1 \pmod{3}$ if $\lambda \mid c$ and $c \equiv 1 \pmod{3}$ otherwise. Since $a = \lambda^2 c(h_0/L + \sum_{p|r} h_p/p)$, these choices fix the residue of a modulo $M c_0$. As $(a, c) = 1$ and $(r, M) = 1$, choose $\delta' \in \mathcal{O}$ with $a\delta' \equiv 1 \pmod{q^{v_q(M c_0)}}$ for $q \mid c_0$, $\delta' \equiv 0 \pmod{q^{v_q(M)}}$ for $q \mid M$, $q \nmid c_0$, and $a\delta' \equiv 1 \pmod{r}$, and put

$$b_g = \frac{a\delta' - 1}{c}, \quad g = \begin{pmatrix} a & b_g \\ c & \delta' \end{pmatrix} \in \text{SL}_2(\mathcal{O}),$$

so that $b_g \equiv 0 \pmod{q^{v_q(M)}}$ for $q \mid c_0$ and $b_g \equiv -c^{-1} \pmod{q^{v_q(M)}}$ for $q \mid M$, $q \nmid c_0$. To choose the cusp expansion at a/c put

$$H = \begin{cases} I, & 3 \mid c, \\ \begin{pmatrix} 1 & 0 \\ u_0 & 1 \end{pmatrix}, & v_\lambda(c) = 1, u_0 \in \{\lambda, -\lambda\}, u_0 \equiv c \pmod{3}, \\ \begin{pmatrix} u_0 & -1 \\ 1 & 0 \end{pmatrix}, & (c, \lambda) = 1, u_0 \equiv a \pmod{3}, \end{cases}$$

with u_0 in the last case taken from a fixed set of representatives modulo 3, and put $g_1 = gH^{-1}$, so $g = g_1H$ and $g_1 \equiv I \pmod{3}$, as in [8, §5.1]. The invariances $\theta(\gamma w) = \theta(w)$ for $\gamma \in \text{SL}_2(\mathbb{Z})$ and $\theta(z + t, v) = \theta(z, v)$ for $t \in \mathbb{Z} + 3\mathcal{O}$ reduce $\bar{\theta}(Hw)$ to one of the three functions $\bar{\theta}(\gamma_\sigma w)$, $\sigma \in \{0, -, +\}$: translating by ω or $-\omega$ and then inverting gives the representatives labelled $-$ and $+$ [8, (5.9)–(5.15), Appendix A]. Write $t_\sigma(\ell)$ for the coefficient of $vK_{1/3}(4\pi|\ell|v)\check{e}(\ell z)$ in $\bar{\theta}(\gamma_\sigma w)$ and let $\tau_1, \tau_2 : \lambda^{-4}\mathcal{O} \setminus \{0\} \rightarrow \mathbb{C}$ be the sequences of [8, (5.13), (5.14)]. Patterson's cusp calculation [26, Theorem 8.1 and Table III] gives

$$t_0(\ell) = \tau(\ell), \quad t_-(\ell) = \omega^2 \tau_1(\omega^2 \ell) \check{e}(\ell), \quad t_+(\ell) = \omega \tau_2(\omega \ell) \check{e}(\ell), \quad (\text{B.67})$$

and complex conjugation changes the frequency ℓ to $-\ell$, so in the normalisation (B.44)

$$d_\sigma(\ell) = \overline{t_\sigma(-\ell)}, \quad \sigma \in \{0, +, -\}. \quad (\text{B.68})$$

For indices $\ell = u\lambda^m n b^3$ as in (B.46) the explicit formulas give $|t_0(\ell)| \leq 3^{k/2+2}|b|$ for $m = 3k-4$, $k \geq 1$, and $|t_0(\ell)| \leq 3^{k/2+5/2}|b|$ for $m = 3k-3$, $k \geq 0$, while the $m = -4$ coefficients of $t_\pm$ have magnitude at most $9|b|$; in each case the bound is at most $27 \cdot 3^{m/6}|b|$, and the coefficients vanish outside the indices (B.46).²¹ This proves the coefficient assertions of Lemma B.15. For the chosen H , with σ the corresponding index,

$$\overline{\theta(H(z, v))} = \frac{3^{5/2}}{2} \mathbf{1}_{\sigma=0} v^{2/3} + \sum_{0 \neq \ell \in \lambda^{-4}\mathcal{O}} d_\sigma(\ell) v K_{1/3}(4\pi|\ell|v) \check{e}(\ell z).$$

The local character transformation. The factorisation $g = g_1H$ lets us apply the automorphy law $\theta(g_1 w) = \kappa(g_1)\theta(w)$, $\kappa(g_1) = (c_1/a_1)_3$ for $g_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \equiv I \pmod{3}$, where κ is Kubota's cubic character [8, (5.4), (5.6)]. For the translated theta function,

$$\begin{aligned} \overline{\theta(z + a/c, v)} &= \overline{\kappa(g_1)} \overline{\theta(Hg^{-1}(z + a/c, v))}, \\ g^{-1}(z + a/c, v) &= \left(-\frac{\delta'}{c} - \frac{\bar{z}}{c^2(v^2 + |z|^2)}, \frac{v}{N(c)(v^2 + |z|^2)} \right), \end{aligned} \quad (\text{B.69})$$

so at $z = 0$ the term of index ℓ in the expansion of $\bar{\theta}(Hw)$ acquires the phase $\check{e}(-\delta'\ell/c)$. We claim

$$\kappa(g_1) = \begin{cases} (c_0/a)_3 (a/r)_3, & 3 \mid c, \\ (-u_0/(a - u_0 b_g))_3 ((c_0/u_0)/a)_3 (a/r)_3, & v_\lambda(c) = 1, \\ (a/c_0)_3 (a/r)_3, & (c, \lambda) = 1. \end{cases} \quad (\text{B.70})$$

This follows from the determinant equation and cubic reciprocity: for $v_\lambda(c) = 1$ the needed congruences are $a(c - u_0 \delta') \equiv -u_0 \pmod{a - u_0 b_g}$ and $b_g c \equiv -1 \pmod{a}$; for $(c, \lambda) = 1$ the congruence $-b_g c = 1 - a\delta' \equiv 1 \pmod{9}$ removes the supplementary factors and reciprocity at the remaining primes gives $(\delta'/(-b_g c))_3 = 1$. The factors other than $(a/r)_3$ depend only on the fixed residues at the primes of S , so $\kappa(g_1) = \kappa_0(a/r)_3 = \kappa_0 \prod_{p|r} \chi_p(a)^2$ with $|\kappa_0| = 1$ fixed once h_0 , the active set and $r \bmod M^2$ are fixed.

Fix h_0 and the active set, so that $c = c_0 r$ is fixed while the nonzero residues h_p vary, and put $D_0 = \lambda^3 c_0$. For each active p define in $\mathcal{O}/(p)$

$$\sigma_p = \lambda^2 c/p, \quad \epsilon_p = -((\lambda^3 c/p)\sigma_p)^{-1},$$

the inverse existing because r is squarefree and $p \nmid \lambda c_0$; the expression for a gives $a \equiv \sigma_p h_p \pmod{p}$. For a dual index $\ell \in \lambda^{-4}\mathcal{O}$ the additive form of the Chinese remainder theorem yields

$$\check{e}(-\delta'\ell/c) = \psi(\lambda^4 \ell) \prod_{p|r} e(\epsilon_p h_p^{-1} \lambda^4 \ell/p), \quad \psi(\lambda^4 \ell) = e(-\delta'_0 r^{-1} \lambda^4 \ell/D_0), \quad \delta'_0 = \delta' \bmod D_0, \quad (\text{B.71})$$

²¹At the index $\lambda^{-3} n b^3$ with n squarefree the coefficient is $3^{5/2}|b|$ times a Gauss sum of modulus one, by (B.38), so the bound $27 \cdot 3^{m/6}|b|$ is attained up to the constant at $m = -3$. The growth $|b|$ in the cube part is the reason for the factor $N(b)^{-1}$ in the completed sums (B.28), which makes the cube variable harmless.

where δ'_0 and $r^{-1} \bmod D_0$ are fixed by the choices above, so ψ ranges over a fixed finite family of additive characters of $\mathcal{O}$. We claim the local transformation identity

$$\sum_{h_p \neq 0} C_{p,j}(h_p) \chi_p(a)^{-2} e(\epsilon_p h_p^{-1} \lambda^4 \ell / p) = \chi_p(\sigma_p)^{-2} \omega_{p,j} B_{p,j}(\lambda^4 \ell), \quad (B.72)$$

$$\omega_{p,j} = \begin{cases} \chi_p(-1)^j \gamma_j(p) \gamma_{j+2}(p) \chi_p(\epsilon_p)^{-j-2}, & j \neq 0, 4, \\ \gamma_4(p), & j = 4, \\ -\gamma_2(p) \chi_p(\epsilon_p)^{-2}, & j = 0 \text{ active,} \end{cases}$$

with $B_{p,j}$ as in (B.41), the indices of $\gamma_j(p)$ read modulo six, and $|\omega_{p,j}| = 1$ in each case. To prove it combine (B.66), the conjugate of (B.70) and (B.71); the substitution $y = h_p^{-1}$ changes the character exponent by $\chi_p(h_p)^{-j} \chi_p(a)^{-2} = \chi_p(\sigma_p)^{-2} \chi_p(y)^{j+2}$. For $j \neq 0, 4$ the character χ_p^{j+2} is nontrivial and its Gauss sum gives the first case of $B_{p,j}$. For $j = 4$ it is trivial on the nonzero residues and, after rescaling by $\epsilon_p \neq 0$, $\sum_{y \neq 0} e(\lambda^4 \ell y / p) = -1 + N(p) \mathbf{1}_{p|\lambda^4 \ell}$. For active $j = 0$ the coefficient $-N(p)^{-1}$ combines with a cubic Gauss sum to give $B_{p,0}(\lambda^4 \ell) = N(p)^{-1/2} \chi_p(\lambda^4 \ell)^{-2}$, the unit factors going into $\omega_{p,0}$. An inactive prime has $j = 0$ and contributes only the scalar $1 - N(p)^{-1}$. Consequently, for fixed h_0 and $\mathcal{A}$, combining the coefficients $d_\sigma(\ell)$ of $\theta(H(z, v))$ with the sums over $h_p \neq 0$ gives $d_\sigma(\ell) \psi(\lambda^4 \ell) \prod_{p \in \mathcal{A}} B_{p,j_p}(\lambda^4 \ell)$ up to a scalar independent of ℓ : these are the arithmetic factors of (B.45), and $B_{p,1}(\lambda^4 \ell) = \chi_p(\lambda^4 \ell)^3$ is the quadratic factor of (B.53).²²

The archimedean transform. For $\operatorname{Re} s > 1$ introduce the Dirichlet series attached to (B.27),

$$\mathcal{T}(s, \Psi) = \left(\sum_{\substack{n \equiv 1 \pmod{3} \\ (n, S) = 1}}^* \overline{\alpha(n)} \gamma_2(n) \Psi(n) N(n)^{-s} \right) \left(\sum_{\substack{b \equiv 1 \pmod{3} \\ (b, S) = 1}} \overline{\alpha(b)}^3 \Psi(b)^3 N(b)^{-3s+1/2} \right),$$

so that Mellin inversion gives

$$T(X; \Psi) = \frac{1}{2\pi i} \int_{(\sigma)} \widehat{V}_*(s - \tfrac{1}{2}) \mathcal{T}(s, \Psi) X^{s-1/2} ds \quad (\sigma > 1), \quad (B.73)$$

and (B.40), with Ψ in place of Ψ_k , identifies these coefficients with those of Θ_Ψ after insertion of $\alpha(nb^3)$. For $z = x + iy$ use $\partial_{\bar{z}} = (\partial_x + i\partial_y)/2$. Differentiation in $\bar{z}$ supplies a factor $\bar{\ell}$ in each Fourier mode, which with the factor $|\ell|^{-1}$ from the Bessel integral produces the angular factor $\overline{\alpha(\bar{\ell})}$. Define, for $\operatorname{Re} s > 1$,

$$\mathcal{J}(s) = \int_0^\infty \partial_{\bar{z}} \Theta_\Psi(z, v) \Big|_{z=0} v^{2s-1} dv.$$

For each nonzero mode, $\partial_{\bar{z}} \check{\ell}(lz) = 2\pi i \bar{\ell} \check{\ell}(lz)$ and, by the Mellin integral of $K_{1/3}$ [7, (10.43.19)],

$$\int_0^\infty v^{2s} K_{1/3}(4\pi|\ell|v) dv = \frac{2^{2s-1} \Gamma(s+1/3) \Gamma(s+2/3)}{(4\pi|\ell|)^{2s+1}} \quad (\operatorname{Re} s > 0). \quad (B.74)$$

For $\operatorname{Re} s > 1$ the coefficient bounds and (B.74) show that the sum of the integrals of the absolute values is finite, so the differentiated series may be integrated term by term. For $\ell = \lambda^{-3} n b^3$ the phase and scale simplify to $i \overline{\alpha(\bar{\ell})} |\ell|^{-2s} = 27^s \overline{\alpha(n)} \overline{\alpha(b)}^3 N(n)^{-s} N(b)^{-3s}$, and with (B.38)

$$\mathcal{J}(s) = \frac{3^{5/2}}{4} \left(\frac{27}{(2\pi)^2} \right)^s \Gamma(s+1/3) \Gamma(s+2/3) \mathcal{T}(s, \Psi). \quad (B.75)$$

We show that $\mathcal{J}$ is entire and use (B.75) to continue $\mathcal{T}(s, \Psi)$. Write $(z', v') = g^{-1}(z + a/c, v)$; differentiating (B.69) at $z = 0$ gives $\partial z' / \partial \bar{z} = -(cv)^{-2}$ and $\partial \bar{z}' / \partial \bar{z} = \partial v' / \partial \bar{z} = 0$, so the derivative in cusp coordinates is $-(cv)^{-2} \partial_{z'}$, with no height derivative. The defining

²²This is the arithmetic heart of the method: a row twist χ_p of order six at an active prime is turned by the Fourier expansion and the Kubota multiplier into the twist $\chi_p^{-3} = \chi_p^3$ of order two on the dual side, to which the quadratic large sieve applies. The cubic part of the symbol is absorbed by the automorphy of θ .

series (B.63) controls $v \rightarrow \infty$; for $v \rightarrow 0$ use (B.69) and the cusp expansions (B.44), in which the horizontal derivative removes the constant mode, the remaining modes decaying exponentially in the transformed height $1/(N(c)v) \rightarrow \infty$. Thus $\mathcal{J}$ is entire, and so is $\mathcal{T}(s, \Psi)$ after division by the gamma factors.

For the term indexed by $\mathbf{h}$ write $c_{\mathbf{h}}, \delta'_{\mathbf{h}}, g_{1,\mathbf{h}}, H_{\mathbf{h}}$ for the choices above, $\sigma_{\mathbf{h}}$ for the cusp index of $H_{\mathbf{h}}$, and define the cusp Mellin transform

$$\mathcal{J}_{\mathbf{h}}^{\vee}(s) = \int_0^{\infty} \partial_z \{ \overline{\theta(H_{\mathbf{h}}(z, v))} \} \Big|_{z=-\delta'_{\mathbf{h}}/c_{\mathbf{h}}} v^{2s-1} dv.$$

Substituting $v \mapsto (N(c_{\mathbf{h}})v)^{-1}$ in each translated term of $\mathcal{J}(s)$ and using (B.69) gives the functional equation

$$\mathcal{J}(s) = - \sum_{\mathbf{h}} c_{\mathbf{F}}(\mathbf{h}) \overline{\kappa(g_{1,\mathbf{h}})} \overline{\alpha(c_{\mathbf{h}})}^2 N(c_{\mathbf{h}})^{1-2s} \mathcal{J}_{\mathbf{h}}^{\vee}(1-s), \quad (\text{B.76})$$

and the cusp coefficients with (B.74) give, for $\text{Re } s > 1$,

$$\mathcal{J}_{\mathbf{h}}^{\vee}(s) = \frac{i \Gamma(s+1/3) \Gamma(s+2/3)}{4(2\pi)^{2s}} \sum_{0 \neq \ell \in \lambda^{-4}\mathcal{O}} \frac{d_{\sigma_{\mathbf{h}}}(\ell) \alpha(\ell)}{N(\ell)^s} \check{e}(-\delta'_{\mathbf{h}} \ell / c_{\mathbf{h}}). \quad (\text{B.77})$$

The series $\mathcal{T}(s, \Psi)$ converges absolutely for $\text{Re } s > 1$, and the series of (B.77) at $1-s$ converges absolutely for $\text{Re } s < 0$; the functional equation and Stirling's formula [17, §5.A.4] give polynomial bounds in $|\text{Im } s|$ for $\mathcal{T}(s, \Psi)$ on both sides of the strip $0 \leq \text{Re } s \leq 1$, splitting the integral of $\mathcal{J}$ at $v = 1$ gives finite order, and Phragmén–Lindelöf [17, Theorem 5.53] gives the same bound inside the strip (compare [8, Propositions 5.1–5.2]). With the rapid decay of $\widehat{V}_*$ on vertical lines this justifies moving the contour of (B.73) to $\text{Re } s < 0$, no pole being crossed; setting $t = \frac{1}{2} - s$ gives a line $\text{Re } t > 1/2$ on which the dual series converges absolutely. For fixed h_0 and active set, insert (B.77) into (B.76) and use (B.72) to sum over the nonzero h_p . After the normalisation (B.75), the scalar of (B.45) for this group is

$$C = -\frac{i}{81} \overline{\alpha(c)}^2 \widehat{\phi}(h_0) \overline{\kappa_0} \prod_{p \text{ inactive}} (1 - N(p)^{-1}) \prod_{p \text{ active}} \chi_p(\sigma_p)^{-2} \omega_{p,j_p}, \quad |C| \leq \frac{1}{81}.$$

Dividing (B.75) by its gamma factors and setting $t = \frac{1}{2} - s$ produces the gamma quotient of (B.42); its numerator has its first pole at $t = -5/6$ and its reciprocal denominator is entire, so no pole lies between $\text{Re } t > 1/2$ and $\text{Re } t = 0$, and the decay of $\widehat{V}_*$ with Stirling's formula allows the kernel contour to be shifted to $\text{Re } t = 0$. This gives $V_*^{\sharp}$ evaluated at $N(\ell)X/N(c)^2$ and, with (B.72), the dual sum (B.45).

Uniformity. For the family Ψ_{k_0} of Lemma B.14, fix h_0 , the factors of Ψ other than χ_{k_0} and the active or inactive choices at their primes. Every prime dividing k_0 has exponent 1 and is active, so $r = k_0 \prod_{p \in \mathcal{A}_{\text{fix}}} p$: r/k_0 is fixed, and fixing $k_0 \bmod M^2$ fixes $r \bmod M^2$. The choices of H , c_0 and the residues in (B.71) are then fixed in each such class, so $d = d_{\sigma}$, ψ and c_0 depend on k_0 exactly as asserted.

The weight. The first numerator pole of the gamma quotient in (B.42) is at $t = -5/6$. Shift the kernel contour to $\text{Re } t = -1/4$ for $0 < x \leq 1$, which gives the factor $x^{1/4}$, and to $\text{Re } t = A$ for $x \geq 1$, which gives x^{-A} ; each application of $x \partial_x$ introduces a factor $-t$, and Stirling's formula on $-1/4 \leq \text{Re } t \leq A$ gives, for $A > 0$, $j \geq 0$,

$$|(x \partial_x)^j V_*^{\sharp}(x)| \ll_{A,j} \min\{x^{1/4}, (1+x)^{-A}\} \sup_{-A \leq \eta \leq 1/4} \int_{\mathbb{R}} (1+|u|)^{[4A]+j+2} |\widehat{V}_*(\eta + iu)| du.$$

When V_* is supported in a fixed compact $I \subset (0, \infty)$, repeated integration by parts in the Mellin transform gives rapid decay in $|u|$ uniformly in $-A \leq \eta \leq 1/4$, so for $J = J(A, j)$ large enough the supremum is $\ll_{A,j,I} \|V_*\|_{C^J(I)}$. This is (B.47), and it completes the proof of the three statements: $T(X; \Psi)$ is a sum of $O_{\Psi_0, S}(2^{|\mathcal{P}|})$ dual sums with $|C| \leq 1/81$, the asserted dependence of (d, ψ, c_0) , and the support, coefficient and weight bounds just established. $\square$

B.15. Separating variables in smooth weights. The last ingredients are three lemmas on smooth weights: the separation of variables by Mellin inversion, the passage from mean squares with a common test function to kernels depending on the row, and the recombination of separated sums.

Lemma B.21. *Fix a box $\mathcal{I} = I_1 \times \cdots \times I_d$ of compact intervals in $(0, \infty)$. Every $\mathcal{K} \in C_c^\infty((0, \infty)^d)$ supported in $\mathcal{I}$ has a representation*

$$\mathcal{K}(\mathbf{x}) = \int_{\mathbb{R}^d} b(\mathbf{t}) \prod_{j=1}^d x_j^{it_j} d\mathbf{t} \quad (x_j > 0), \quad (\text{B.78})$$

and for $J \geq 0$ and every even integer $q > J + d$, $\int_{\mathbb{R}^d} |b(\mathbf{t})| (1 + |\mathbf{t}|)^J d\mathbf{t} \ll_{\mathcal{I}, J, q, d} \|\mathcal{K}\|_{C^q(\mathcal{I})}$.

Proof. Take the Fourier transform in logarithmic coordinates,

$$b(\mathbf{t}) = (2\pi)^{-d} \int_{\mathbb{R}^d} \mathcal{K}(e^{y_1}, \dots, e^{y_d}) e^{-it \cdot \mathbf{y}} d\mathbf{y};$$

Fourier inversion gives (B.78). Since the intervals are fixed and bounded away from zero, the C^q norm in logarithmic coordinates is bounded by a constant times $\|\mathcal{K}\|_{C^q(\mathcal{I})}$, and applying $(1 - \Delta_{\mathbf{y}})^{q/2}$ under the integral gives $|b(\mathbf{t})| \ll_{\mathcal{I}, q, d} (1 + |\mathbf{t}|)^{-q} \|\mathcal{K}\|_{C^q(\mathcal{I})}$ for even $q \geq 0$. Multiply by $(1 + |\mathbf{t}|)^J$ and integrate. $\square$

In the applications the weight has the form $\mathcal{K}_R(\mathbf{x}) = \prod_{j=1}^d W_j(x_j) F(R \prod_j x_j^{a_j})$ with $R > 0$, fixed real exponents a_j , $W_j \in C_c^\infty((0, \infty))$ supported in I_j , and $F \in C^\infty((0, \infty))$ with

$$\|F\|_{A, q} = \max_{0 \leq m \leq q} \sup_{u > 0} (1 + u)^A |(u \partial_u)^m F(u)| < \infty \quad (A \geq 0, q \in \mathbb{Z}_{\geq 0}).$$

On the fixed box the argument of F is comparable to R , so the chain rule and the pointwise bound of the proof give coefficients b_R with, for $A \geq 0$ and even $q \geq 0$,

$$|b_R(\mathbf{t})| \ll (1 + R)^{-A} (1 + |\mathbf{t}|)^{-q} \|F\|_{A, q} \prod_{j=1}^d \|W_j\|_{C^q(I_j)}, \quad (\text{B.79})$$

and hence, for $J \geq 0$ and even $q > J + d$,

$$\int_{\mathbb{R}^d} |b_R(\mathbf{t})| (1 + |\mathbf{t}|)^J d\mathbf{t} \ll (1 + R)^{-A} \|F\|_{A, q} \prod_{j=1}^d \|W_j\|_{C^q(I_j)}, \quad (\text{B.80})$$

the constants depending only on A, J, q, d , the intervals and the exponents: separation preserves the decay in R at the cost of finitely many derivatives.

Lemma B.22. *Fix compact intervals $I \subset \text{int } I_*$ in $(0, \infty)$ and an integer $m \geq 0$. Let $S_{j,r}(U) = \sum_n a_{j,r}(n) U(x_{j,r,n})$, $j = 1, 2$, $x_{j,r,n} > 0$, be two families of finite sums over a finite set of rows r with nonnegative weights w_r , and suppose $\sum_r w_r |S_{j,r}(U)|^2 \leq M_j \|U\|_{C^m(I_*)}^2$ for every $U \in C_c^\infty(I_*)$. Then for $|c_r| \leq w_r$ and smooth kernels $\mathcal{K}_r$ supported in I^2 ,*

$$\left| \sum_r c_r \sum_{n_1, n_2} a_{1,r}(n_1) \overline{a_{2,r}(n_2)} \mathcal{K}_r(x_{1,r,n_1}, x_{2,r,n_2}) \right| \ll_{I, I_*, m} \sqrt{M_1 M_2} \sup_r \|\mathcal{K}_r\|_{C^{2m+4}(I^2)}. \quad (\text{B.81})$$

Proof. Choose a real $V \in C_c^\infty(I_*)$ equal to one on I and put $U_t(x) = V(x)x^{it}$. Apply (B.78) to each $\mathcal{K}_r$ with $d = 2$, replace the second Mellin variable by its negative and multiply by $V(x)V(y)$, which is one on the support: $\mathcal{K}_r(x, y) = \int_{\mathbb{R}^2} b_r(s, t) U_s(x) \overline{U_t(y)} ds dt$ with $\sup_r |b_r(s, t)| \ll_{I, q} \sup_r \|\mathcal{K}_r\|_{C^q(I^2)} (1 + |s| + |t|)^{-q}$ for even q . The hypothesis applies to each U_t with $\|U_t\|_{C^m(I_*)} \ll_{I, I_*, m} (1 + |t|)^m$. Taking the common bound for b_r before integrating and applying Cauchy–Schwarz in r bounds the left side of (B.81) by a constant times $\sqrt{M_1 M_2} \sup_r \|\mathcal{K}_r\|_{C^q(I^2)} \int_{\mathbb{R}^2} (1 + |s|)^m (1 + |t|)^m (1 + |s| + |t|)^{-q} ds dt$, and $q = 2m + 4$ makes the integral converge. $\square$

Lemma B.23. *Let r, ι range over finite sets, $d \geq 1$, $w_r \geq 0$, and let measurable coefficients satisfy $|c_{\iota,r}(\mathbf{t})| \leq b_\iota(\mathbf{t})$ with $M = \sum_\iota \int_{\mathbb{R}^d} b_\iota(\mathbf{t}) d\mathbf{t} < \infty$. Then for measurable $F_{\iota,\mathbf{t}}(r)$, whenever the right side is finite,*

$$\sum_r w_r \left| \sum_\iota \int c_{\iota,r}(\mathbf{t}) F_{\iota,\mathbf{t}}(r) d\mathbf{t} \right|^2 \leq M \sum_\iota \int b_\iota(\mathbf{t}) \sum_r w_r |F_{\iota,\mathbf{t}}(r)|^2 d\mathbf{t}, \quad (\text{B.82})$$

and the same holds for finite sums in place of the integrals.

Proof. For fixed r , weighted Cauchy–Schwarz gives

$$\left| \sum_\iota \int c_{\iota,r} F_{\iota,\mathbf{t}}(r) d\mathbf{t} \right|^2 \leq M \sum_\iota \int b_\iota |F_{\iota,\mathbf{t}}(r)|^2 d\mathbf{t};$$

multiply by w_r and sum. $\square$

B.16. Conclusion. The chain is now closed. Lemma B.17 and Proposition B.9 rest on the theta transformation of Section B.14, the quadratic large sieve of Section B.11 and the weight lemmas of Section B.15; Proposition B.11 rests on the two Poisson summations of Section B.13; Proposition B.8 follows from these by the induction of Section B.8; Proposition B.2 follows in Section B.9; and Section B.2 deduces Theorem B.1 from Proposition B.2. Every input not proved here is a published theorem cited at the place where it is used: Patterson’s computation of the Fourier coefficients of the cubic theta function [26], the transformation formulas of Dunn and Radziwiłł [8], the quadratic large sieve of Goldmakher and Louvel [9], Landau’s prime ideal theorem [20], Hecke’s analytic continuation of the L -functions of the field [15], and the standard facts on Gauss sums, Stirling’s formula and Phragmén–Lindelöf from [17]. Applied with $\theta = 11/12$ to the Dirichlet L -functions $L(s, \chi\psi)$ of Corollary 1.3, Theorem B.1 gives Theorem 1.5.

DECLARATIONS

Funding. No funding was received for this work.

Conflict of interest. The authors declare no conflict of interest.

Data availability. The proofs in this paper are complete by hand, the zero-density theorem of Appendix A and the zero-free half-plane of Appendix B included. Release v3.0.0 of the repository named in the note under the abstract (10.5281/zenodo.23672710) contains the source of the paper; programs that check the numerical inequalities between the constants of Sections 3 to 7 and of Appendix A, and that compute $M(\chi)/(\sqrt{q} \log \log q)$ and $|L'/L(1, \chi)|/\log \log q$ for the characters to small prime moduli; and, as an independent check of Theorem 1.5, the Lean 4 sources of the proof of the half-plane $\text{Re } s > 7/8$ of [23], which we compiled with Lean 4.34.1 and the Mathlib library, with the report of the Lean kernel that the theorem depends on no axiom beyond the three standard ones `propext`, `Classical.choice` and `Quot.sound`, and a Lean 4 proof, checked in the same build, of the two bounds of Theorem 1.5 from that theorem. No statement of the paper depends on these programs.

REFERENCES

- [1] D. Bhattacharjee, Establishing equivariant class $[O]$ for hyperbolic groups, *Asian Res. J. Math.* 18 (2022), no. 11, 362–369. doi:10.9734/arjom/2022/v18i11615.
- [2] D. Bhattacharjee, Non-commutativity over canonical suspension η for genus $g \geq 1$ in hypercomplex structures for potential $\rho\phi$, *Asian Res. J. Math.* 18 (2022), no. 11, 332–341. doi:10.9734/arjom/2022/v18i11607.
- [3] D. Bhattacharjee, U. Bhattacharya, Iterated nerve complexes for k -fold sensor coverage, *Preprints.org* (2026). doi:10.20944/preprints202606.1351.v2.
- [4] J. Bober, L. Goldmakher, A. Granville, D. Koukoulopoulos, The frequency and the structure of large character sums, *J. Eur. Math. Soc.* 20 (2018), no. 7, 1759–1818. doi:10.4171/JEMS/799.
- [5] V. Blomer, L. Goldmakher, B. Louvel, L -functions with n -th-order twists, *Int. Math. Res. Not. IMRN* 2014, no. 7, 1925–1955. doi:10.1093/imrn/rns257.
- [6] H. Davenport, *Multiplicative Number Theory*, 2nd ed., revised by H. L. Montgomery, Graduate Texts in Mathematics 74, Springer, New York, 1980. doi:10.1007/978-1-4757-5927-3.

- [7] NIST Digital Library of Mathematical Functions, §10.43, equation (10.43.19), <https://dlmf.nist.gov/10.43.E19>.
- [8] A. Dunn, M. Radziwiłł, Bias in cubic Gauss sums: Patterson’s conjecture, *Ann. of Math. (2)* 200 (2024), no. 3, 967–1057. doi:10.4007/annals.2024.200.3.3.
- [9] L. Goldmakher, B. Louvel, A quadratic large sieve inequality over number fields, *Math. Proc. Cambridge Philos. Soc.* 154 (2013), no. 2, 193–212. doi:10.1017/S0305004112000370.
- [10] A. Granville, K. Soundararajan, Large character sums: pretentious characters and the Pólya–Vinogradov theorem, *J. Amer. Math. Soc.* 20 (2007), 357–384. doi:10.1090/S0894-0347-06-00536-4.
- [11] H. Halberstam, H.-E. Richert, *Sieve Methods*, London Mathematical Society Monographs 4, Academic Press, London, 1974.
- [12] D. R. Heath-Brown, A mean value estimate for real character sums, *Acta Arith.* 72 (1995), no. 3, 235–275. doi:10.4064/aa-72-3-235-275.
- [13] D. R. Heath-Brown, Kummer’s conjecture for cubic Gauss sums, *Israel J. Math.* 120 (2000), 97–124. doi:10.1007/s11856-000-1273-y.
- [14] D. R. Heath-Brown, S. J. Patterson, The distribution of Kummer sums at prime arguments, *J. Reine Angew. Math.* 310 (1979), 111–130. doi:10.1515/crll.1979.310.111.
- [15] E. Hecke, Eine neue Art von Zetafunktionen und ihre Beziehungen zur Verteilung der Primzahlen, I, II, *Math. Z.* 1 (1918), 357–376; 6 (1920), 11–51. doi:10.1007/BF01465095; doi:10.1007/BF01202991.
- [16] A. E. Ingham, On the estimation of $N(\sigma, T)$, *Quart. J. Math. Oxford Ser.* 11 (1940), 201–202. doi:10.1093/qmath/os-11.1.201.
- [17] H. Iwaniec, E. Kowalski, *Analytic Number Theory*, American Mathematical Society Colloquium Publications 53, American Mathematical Society, Providence, RI, 2004. doi:10.1090/coll/053.
- [18] Y. Ihara, V. K. Murty, M. Shimura, On the logarithmic derivatives of Dirichlet L -functions at $s = 1$, *Acta Arith.* 137 (2009), 253–276. doi:10.4064/aa137-3-6.
- [19] T. Kubota, On Automorphic Functions and the Reciprocity Law in a Number Field, *Lectures in Mathematics*, Department of Mathematics, Kyoto University 2, Kinokuniya, Tokyo, 1969.
- [20] E. Landau, Neuer Beweis des Primzahlsatzes und Beweis des Primidealsatzes, *Math. Ann.* 56 (1903), 645–670. doi:10.1007/BF01444310.
- [21] H. L. Montgomery, *Topics in Multiplicative Number Theory*, Lecture Notes in Mathematics 227, Springer, Berlin, 1971. doi:10.1007/BFb0060851.
- [22] H. L. Montgomery, R. C. Vaughan, Exponential sums with multiplicative coefficients, *Invent. Math.* 43 (1977), 69–82. doi:10.1007/BF01390204.
- [23] OpenAI, The quasi-Riemann hypothesis: a zero-free half-plane $\text{Re } s > 7/8$, OpenAI Math Release preprint with Lean 4 sources, 30 September 2026, <https://github.com/openai/math>, directory **The-Quasi-Riemann-Hypothesis-September-30-2026**.
- [24] OpenAI, The quasi-Riemann hypothesis: a cubic theta argument with a fixed ray class character, OpenAI Math Release preprint, 5 October 2026, <https://github.com/openai/math>, directory **The-Quasi-Riemann-Hypothesis-October-5-2026**.
- [25] R. E. A. C. Paley, A theorem on characters, *J. London Math. Soc.* s1-7 (1932), 28–32. doi:10.1112/jlms/s1-7.1.28.
- [26] S. J. Patterson, A cubic analogue of the theta series, *J. Reine Angew. Math.* 296 (1977), 125–161. doi:10.1515/crll.1977.296.125.
- [27] W. Rudin, *Real and Complex Analysis*, 3rd ed., McGraw-Hill, New York, 1987.
- [28] A. Yoshimoto, On the cubic theta function, *Nagoya Math. J.* 105 (1987), 153–167. doi:10.1017/S0027763000000817.

FORMERLY, ELECTRO-GRAVITATIONAL SPACE PROPULSION LABORATORY (EGSPL), BHUBANESWAR, ODISHA 751030, INDIA

FORMERLY, UNIVERSITY OF CALCUTTA, SENATE HOUSE, 87/1 COLLEGE STREET, KOLKATA 700073, WEST BENGAL, INDIA

Email address: d.bhattacharjee@erl-forschung.de

Email address: itsdeep@live.com

Email address: egspl@proton.me

FORMERLY, NATIONAL TAIWAN UNIVERSITY, TAIPEI 106319, TAIWAN

Email address: rai.bhattacharya.in@gmail.com