DIRICHLET L-FUNCTIONS AND THE CRITICAL LINE

ABSTRACT. We give an expanded mathematical account of the proof of the critical-line theorem for Dirichlet L-functions in the accompanying Lean development. The analytic reduction detects poles from one-sided bounds on arbitrarily distant logarithmic windows. The arithmetic estimate combines Fourier filtering in abelian CM fields, local divisibility in theta towers, and a transfer of mean slopes through approximate tight frames. We prove the analytic, finite-dimensional, and numerical mechanisms in detail and state the deeper geometric and cohomological inputs with their hypotheses and source locations. The exposition separates established results from construction-specific arguments and provides a concordance with the principal Lean theorems.

1. THE THEOREM AND THE ARGUMENT

For a Dirichlet character χ modulo $N \geq 1$, let $L(s,\chi)$ denote the meromorphic continuation of $\sum_{n\geq 1}\chi(n)n^{-s}$, initially defined for $\operatorname{Re} s > 1$. The character is extended by zero on integers not relatively prime to its modulus. The character of conductor 1 gives the Riemann zeta function. The main assertions of the development are the following.

Theorem 1.1 (`LFunction_ne_zero_of_half_lt_re`)**.** *For every Dirichlet character χ, one has $L(s,\chi) \neq 0$ when $\operatorname{Re} s > 1/2$, with the pole at $s = 1$ excluded when χ is principal.*

Theorem 1.2 (`LFunction_eq_zero_re`)**.** *If $0 < \operatorname{Re}\rho < 1$ and $L(\rho,\chi) = 0$, then $\operatorname{Re}\rho = 1/2$. This holds for every modulus and every Dirichlet character, including imprimitive characters.*

Theorem 1.3 (`riemannHypothesis`)**.** *Every zero of ζ other than the zeros $-2,-4,-6,\ldots$ has real part $1/2$.*

These statements concern Dirichlet L-functions. No corresponding assertion for arbitrary automorphic L-functions is being asserted here. Automorphic forms enter the proof as auxiliary objects that furnish an arithmetic inequality. The public declarations are in `Main.lean`; the predicates they prove are written independently of the argument in `Statements.lean`. Their precise relation to the usual meromorphic convention at $s = 1$ is recorded in Section A.

The purpose of this paper is to make the path from that arithmetic inequality to the zero statement readable, while also explaining how the inequality is constructed. Two different kinds of completeness need to be kept separate. The analytic pole argument, the frame inequality, the cancellation of normalizing terms, and the final choice of parameters are proved below. The substantially longer representation-theoretic, geometric, and cohomological constructions are presented through explicit intermediate theorems, with the hypotheses that are needed here and links to their proofs in the source. Thus the paper is a detailed guide to the full development; its treatment of those deep inputs is by precise citation, not an independent reconstruction of every imported theorem. In particular, a statement labeled as a source input is not supplied with a short argument that would purport to replace its formal proof.

Here is the central analytic quantity. For a nonprincipal character put $\ell_\chi = L'(1,\chi)/L(1,\chi)$, and for $y > 0$ set

$$S_\chi(y) = \sum_{n\leq y}\Lambda(n)\chi(n)(y/n-1), \qquad G_\chi(y) = S_\chi(y) + y\ell_\chi. \tag{1.1}$$

The finite sum is empty when $0 < y < 1$. For the character of conductor 1 the regularization is instead $G_1(y) = S_1(y) - y\log y + (1+\gamma_E)y$, where γ_E is Euler's constant. The constants in this definition are forced by the Laurent expansion at $s = 1$; changing the sign of the ℓ_χ term would leave the wrong pole in the Mellin transform. If ψ is primitive of order o, let $\psi^{[j]}$ denote the

primitive character inducing ψ^j, and define

$$H_\psi(y) = G_1(y) + \frac{1}{3o+1}\sum_{j=1}^{o-1} \operatorname{Re} G_{\psi^{[j]}}(y). \tag{1.2}$$

The coefficient of the principal character is 1, and every retained nonprincipal character has coefficient $1/(3o+1)$. In particular, the chosen character occurs with a strictly positive coefficient; if $o = 1$, the sum is empty and $H_\psi = G_1$. This positivity prevents cancellation of a zero in the associated sum of logarithmic derivatives. These are the coefficients produced by the source's cyclic level and by the Fourier selector below.

The arithmetic output has a particularly simple form. For every sufficiently small $\varepsilon > 0$ and every prescribed lower bound, there is a larger Y such that

$$H_\psi(y) \le y^{1/2+9\varepsilon} \qquad (Y \le y \le Y^{3/2}). \tag{1.3}$$

There is no absolute value on the left, and the intervals need not cover a final half-line. Neither omission is a defect: the analytic argument is designed for exactly this amount of information. On writing $y = e^u$, the interval becomes $[U, 3U/2]$. A high derivative of the Laplace transform has the positive kernel $u^k e^{-\sigma u}/k!$. Its mass can be concentrated in this interval while its tails are made exponentially small. Positivity then compares the derivative at a real point with the modulus of a derivative at a suitable nonreal point. A single closest pole contradicts the resulting upper bound. The separate case of a closest real pole is disposed of by the positive sign of its residue. Section 2 contains the full argument.

The arithmetic construction is organized around a comparison of two contributions to a hermitian lattice. Frobenius at suitable finite primes forces additional divisibility of theta coefficients. At the infinite places, explicit norms contribute periods and elementary logarithmic terms. Fourier inversion in an abelian CM host field identifies their difference with a smoothed prime sum. The host and its archimedean weights are chosen so that the part of this sum which survives is the positive combination H_ψ in (1.2).

A direct comparison of dimensions would lose too much information. There are many small source spaces and one much larger target space; individual evaluation maps need not be close to surjective. The useful invariant is the average slope, and the useful analytic hypothesis is an approximate tight-frame estimate for the entire family of maps. It bounds the codimension of the span of high-slope source subspaces. Integrating those codimension bounds through the Harder–Narasimhan filtration transfers the average slope with a logarithmic loss, rather than a loss proportional to the full range of slopes. This is the role of Theorem 5.3.

The remaining scale parameter is the length m of a theta tower. Its order is close to $x^{1/2}$, where x is the argument of the final prime sum. The finite-place gain is repeated along the tower, whereas the principal slope and cohomological costs have a different dependence on m. After division by the total tower weight, the admissible inequalities for m, the ranks, and the level give an error of size $x^{1/2+O(\varepsilon)}$. The square-root exponent is consequently a quantitative balance in the construction; it is not an additional analytic hypothesis.

The distinction between classical ingredients and the contribution of this argument is most useful at the level of mechanisms. The logarithmic derivative of an Euler product, its Mellin transform, and Landau's use of positivity are classical [9, 19]. The formulation on a selected sequence of multiplicative windows and its compatibility with a one-sided arithmetic estimate are the features used here. Behrend's construction of large sets without short additive configurations is classical [3]; the simultaneous parity condition, density bound, and family of thick translates required by the CM host are specific to this construction. Harder–Narasimhan filtrations and arithmetic slope inequalities are established tools [13, 6]; the pooled frame estimate, its layer-cake integration, and the exact cancellation with the theta norm formula form the interface specific to this construction.

Likewise, the local theta correspondence, the Rallis inner-product formula, Lubin–Tate theory, arithmetic volumes, and L^2 extension are not new theories here. What requires proof is that their integral models, measures, weights, and error estimates can be chosen together in the stated range. The same distinction applies to the Galois-cohomological part. Vanishing of

an Iwasawa μ-invariant does not by itself provide the uniform bounds on characteristic roots that the argument needs. Those bounds, the passage from automorphic congruences to local conditions, and the construction of a common window are separate propositions. Sections 6–9 state them at the appropriate level of precision.

We use "construction-specific" to identify a result whose exact formulation is supplied by this development and is not an immediate invocation of a cited classical theorem. This is a statement about the organization and mathematical burden of the present proof, not a claim that no analogous idea occurs elsewhere. Historical priority would require a separate literature comparison. The especially useful features to watch for are the one-sided window criterion, the positive Fourier selector, the pooling of slopes through a frame, and the cancellation of local, archimedean, and level normalizations before estimation. None of these features can be assessed solely from a list of theorem names.

All unspecified constants are allowed to depend on the fixed primitive character ψ and on a fixed positive ε, unless uniformity is explicitly stated. Such constants never depend on the variable scale x or on a later choice of an auxiliary prime. We write $a \ll_{\mathcal{D}} b$ when the implicit constant depends only on the displayed data $\mathcal{D}$, and $a \asymp_{\mathcal{D}} b$ when both inequalities hold. Complex embeddings are counted with their multiplicities: a conjugate pair contributes twice in an arithmetic degree and once, with the corresponding factor 2, when we sum over complex places. These conventions become essential in the slope calculation.

The argument may be read in two passes. Sections 2, 5, 10, and 11 give the analytic, linear-algebraic, and quantitative deductions. The intervening sections explain the data that make those deductions applicable and distinguish their classical inputs from the specialized construction theorems. Only principal propositions retain the source's declaration names in their headings. Other correspondence with Lean is placed where it helps locate a proof, and collected in the final concordance.

2. One-sided windows and poles

We begin with the analytic reduction because it explains precisely what the arithmetic construction must deliver. The only classical facts about Dirichlet L-functions used in this section are their Euler products in $\operatorname{Re} s > 1$, meromorphic continuation, nonvanishing at $s = 1$ for nonprincipal characters, and their functional equations. These are understood with the standard primitive normalization [9]. All transforms, residues, and localization estimates needed for the reduction are derived below.

For a primitive nonprincipal character χ, retain S_χ, G_χ, and ℓ_χ from (1.1). We use $\chi = 1$ only for the primitive character of conductor 1, and write

$$G_1(y) = S_1(y) - y\log y + (1+\gamma_E)y, \qquad G_\chi(y) = S_\chi(y) + y\ell_\chi \quad (\chi \neq 1). \tag{2.1}$$

Both functions are continuous and piecewise smooth on $[1,\infty)$. No estimate for primes is needed for the preliminary growth bound. Since $0 \le \Lambda(n) \le \log n$, one has $|S_\chi(y)| \le y\sum_{n\le y}(\log n)/n \ll y(1+\log y)^2$. For each fixed character, therefore, $|G_\chi(y)| \ll_\chi y^2$. This intentionally weak bound is sufficient to control the tails of a differentiated Laplace transform.

Proposition 2.1 (`poleData`). *For $\operatorname{Re} s > 1$ the Mellin transform $F_\chi(s) = \int_1^\infty G_\chi(y) y^{-s-1}\,dy$ is*

$$F_\chi(s) = -\frac{L'(s,\chi)/L(s,\chi) - s\ell_\chi}{s(s-1)} \qquad (\chi \neq 1), \tag{2.2}$$

$$F_1(s) = -\frac{\zeta'(s)/\zeta(s) + (s-1)^{-1} + 1 - (1+\gamma_E)s}{s(s-1)}. \tag{2.3}$$

These expressions continue meromorphically to $\operatorname{Re} s > 0$ and are regular at $s = 1$. If ρ is a zero of $L(s,\chi)$ in that half-plane, of multiplicity $m_\chi(\rho)$, its contribution to the residue is $-m_\chi(\rho)/(\rho(\rho-1))$.

Proof. Write $a = \operatorname{Re} s > 1$. The interchange of the sum and integral is justified by the convergent nonnegative sum

$$\sum_{n\ge1} \Lambda(n) \int_n^\infty (y/n - 1) y^{-a-1}\, dy = \frac{1}{a(a-1)} \sum_{n\ge1} \frac{\Lambda(n)}{n^a} < \infty.$$

Direct integration gives $\int_n^\infty (y/n-1)y^{-s-1}\,dy = n^{-s}/[s(s-1)]$. The logarithmic derivative of the absolutely convergent Euler product is $-L'(s,\chi)/L(s,\chi) = \sum_n \Lambda(n)\chi(n)n^{-s}$. Adding the Mellin transform $\ell_\chi/(s-1)$ of $y\ell_\chi$ proves (2.2).

For the principal character, the two elementary transforms are $\int_1^\infty y \log y\, y^{-s-1}\,dy = (s-1)^{-2}$ and $\int_1^\infty y\, y^{-s-1}\,dy = (s-1)^{-1}$. Thus $F_1 = -\zeta'/[\zeta s(s-1)] - (s-1)^{-2} + (1+\gamma_E)/(s-1)$, which is (2.3). The terms 1 and $(1+\gamma_E)s$ in its numerator must both be retained.

For $\chi \ne 1$, the numerator in (2.2) vanishes at 1. For $\chi = 1$, the Laurent expansion $\zeta'/\zeta(s) = -(s-1)^{-1} + \gamma_E + O(s-1)$ gives the same conclusion for the regularized numerator in (2.3). These are removable singularities. At a zero $\rho \ne 0, 1$, write $L(s,\chi) = (s-\rho)^m u(s)$ with u holomorphic and $u(\rho) \ne 0$. Then $L'/L = m/(s-\rho) + u'/u$, giving the asserted residue. The pole is simple even when the zero has higher multiplicity. □

For a finite family of primitive characters with positive weights a_χ, put $G(y) = \sum_\chi a_\chi \operatorname{Re} G_\chi(y)$, $g(u) = G(e^u)$, and

$$\Phi(s) = \int_0^\infty g(u) e^{-su}\, du = \frac{1}{2} \sum_\chi a_\chi (F_\chi(s) + F_{\bar\chi}(s)). \tag{2.4}$$

The equality initially holds for $\operatorname{Re} s > 1$. The function g is real-valued, and $|g(u)| \le C_0 e^{2u}$ for a fixed C_0. The meromorphic continuation in (2.4) has residue

$$\operatorname{Res}_{s=\rho} \Phi(s) = -\frac{b_\rho}{\rho(\rho-1)}, \qquad b_\rho = \frac{1}{2} \sum_\chi a_\chi (m_\chi(\rho) + m_{\bar\chi}(\rho)) > 0 \tag{2.5}$$

at every zero represented by the family. The positivity of b_ρ means that a shared zero cannot disappear in the sum. At a real zero $\beta \in (0,1)$ the residue is $b_\beta/[\beta(1-\beta)] > 0$. These are two distinct uses of positivity: the weights prevent cancellation of poles, while the real residue has the sign needed in the real-pole case below.

We isolate the localization argument as a general statement. A meromorphic function has a *locally finite* set of poles if every compact subset of its domain meets only finitely many poles; this holds automatically for a meromorphic function that is not identically infinite.

Theorem 2.2 (`landau_window`)**.** *Let $g : [0,\infty) \to \mathbb{R}$ be measurable, with $|g(u)| \le C_0 e^{2u}$. Suppose its Laplace transform, initially defined in $\operatorname{Re} s > 2$, extends meromorphically to $\operatorname{Re} s > 0$, with simple poles, no poles in $\operatorname{Re} s > 1$, and no pole at $s = 1$. Suppose further that every real pole in $(0,1)$ has positive residue. Fix $0 < \theta < 1$ and $\delta > 0$.*

Assume that for every $\eta > 0$ there are arbitrarily large U and nonnegative measurable functions R_U on $[U, (1+\delta)U]$ such that $g \le R_U$ on that interval and

$$\int_U^{(1+\delta)U} R_U(u) e^{-\theta u}\, du \le e^{\eta U}. \tag{2.6}$$

Then the meromorphic continuation has no pole in $\operatorname{Re} s > \theta$.

The theorem only requires an upper bound for g. Large negative values are permitted everywhere. It also makes no assumption on the gaps between the admissible windows. Landau's positivity principle is classical [19]; the point of the formulation is to match an arithmetic construction that is available only on a selected sequence of windows. We now give its proof, including the concentration and pole-selection steps.

Lemma 2.3. *For every $\delta > 0$ there are $\sigma > 2$ and $C_\delta > 0$ such that, for each $U \ge 1$, an integer $k \ge U$ satisfies*

$$\int_{(0,\infty)\setminus[U,(1+\delta)U]} \frac{u^k e^{-(\sigma-2)u}}{k!}\, du \le \frac{C_\delta}{\sigma^k}. \tag{2.7}$$

Proof. For $s > \ell > 0$, the gamma integral $\int_0^\infty u^k e^{-au}\,du/k! = a^{-k-1}$ gives the elementary exponential-tail estimate

$$\int_{(0,\infty)\setminus[U,V]} \frac{u^k e^{-su}}{k!}\,du \le \frac{e^{-\ell V}}{(s-\ell)^{k+1}} + \frac{e^{\ell U}}{(s+\ell)^{k+1}}. \tag{2.8}$$

Indeed, for $u > V$ multiply the kernel with parameter $s-\ell$ by $e^{-\ell V}$, and for $u < U$ multiply the kernel with parameter $s+\ell$ by $e^{\ell U}$. In either range the resulting function dominates the original kernel, so its integral over the whole positive half-line is an upper bound.

Choose $\ell = 2 + 6/\delta$, $M = \ell(1+\delta)$, $\sigma = M(\ell+2)$, and $b = 1+\delta/2$, and take $k = \lceil b\sigma U\rceil$. Then $k \ge U$ and $s = \sigma - 2 > \ell$. Write $a_- = \sigma - 2 - \ell$ and $a_+ = \sigma - 2 + \ell$. The identities $b(\ell+2) = M-1$ and $b(\ell-2) = \ell+1$ are the reason for this choice.

For every $a > 0$, $\log(\sigma/a) \le (\sigma-a)/a$. Hence $e^x a^{-k-1} \le e^c/(a\sigma^k)$ whenever $x + k(\sigma-a)/a \le c$. For the first term of (2.8), with $V = (1+\delta)U$, one has $(\sigma - a_-)/a_- = 1/(M-1)$ and $b\sigma = M(M-1)$. Since $k \le b\sigma U + 1$, it follows that $-\ell V + k/(M-1) \le 1$. For the second term, $(\sigma - a_+)/a_+ = -(\ell-2)/a_+$. The inequality $\ell a_+ \le (\ell+1)\sigma$, together with $k \ge b\sigma U$, gives $\ell U - k(\ell-2)/a_+ \le 0$. Applying these two bounds to (2.8) proves (2.7) with $C_\delta = e/a_- + 1/a_+$. □

For $k \ge 0$ and $\operatorname{Re} z > 2$, differentiation under the integral in (2.4) is justified by the integrable majorant $C_0 u^k e^{-(\operatorname{Re} z - 2)u}$. Set

$$A_k(z) = \frac{(-1)^k}{k!}\Phi^{(k)}(z) = \int_0^\infty g(u)\frac{u^k e^{-zu}}{k!}\,du. \tag{2.9}$$

Since g is real, $A_k(\sigma)$ is real. The triangle inequality and $|g| + g = 2\max(g,0)$ give

$$|A_k(\sigma+it)| + \operatorname{Re} A_k(\sigma) \le 2\int_0^\infty \max(g(u),0)\frac{u^k e^{-\sigma u}}{k!}\,du. \tag{2.10}$$

This is the step that replaces an unavailable two-sided bound. A negative contribution to $A_k(\sigma)$ helps the inequality rather than having to be estimated in absolute value.

The elementary bound $v^k e^{-v}/k! \le 1$ for $v \ge 0$ follows by retaining the kth term in the exponential series. On an admissible window, therefore,

$$\frac{u^k e^{-\sigma u}}{k!} \le (\sigma-\theta)^{-k} e^{-\theta u}.$$

Off that window we use $\max(g,0) \le C_0 e^{2u}$ and Lemma 2.3. Thus, for the integer k supplied by that lemma and every $t \in \mathbb{R}$,

$$|A_k(\sigma+it)| + \operatorname{Re} A_k(\sigma) \le \frac{2e^{\eta U}}{(\sigma-\theta)^k} + \frac{2C_0 C_\delta}{\sigma^k}. \tag{2.11}$$

The constants σ and C_δ do not depend on the chosen window. The derivative order k increases with U, which will make a nearest pole visible.

Lemma 2.4. *Let Z be a locally finite set in $\operatorname{Re} s > 0$, with every element satisfying $\operatorname{Re} s \le 1$. Let $\sigma > 2$ and $0 < \theta < 1$. If Z has an element with real part greater than θ, one of the following holds:*

(i) there is $0 < D \le \sigma - \theta$ such that the open disk $B(\sigma, D)$ contains no element of Z, and some $\rho \in Z$ satisfies $\sigma - \operatorname{Re}\rho < D$;

(ii) there is a real $\beta > \theta$ in Z and a radius $r > \sigma - \beta$ such that $r \le \sigma - \theta$ and β is the only element of Z in $B(\sigma, r)$.

Proof. Put $E = \sigma - \theta$. The closed disk $\overline{B}(\sigma, E)$ is a compact subset of $\operatorname{Re} s > 0$, so it meets Z in a finite set. If that set is empty, take $D = E$ and use any element of Z lying to the right of θ. Otherwise let D be the smallest distance to σ. If some ρ satisfies $\sigma - \operatorname{Re}\rho < D$, we are in case (i). If none does, every nearest element satisfies $D = |\sigma - \rho| \ge \sigma - \operatorname{Re}\rho \ge D$. Equality forces ρ to be the real point $\beta = \sigma - D$. The assumed element to the right of θ has horizontal distance less than E, so $D < E$ and $\beta > \theta$. There is only one point at distance D, and local finiteness permits a slightly larger disk, still of radius at most E, containing no other point of Z. This is case (ii). □

Proof of Theorem 2.2. Suppose a pole exists to the right of θ, and apply Lemma 2.4 to the pole set. Put $E = \sigma - \theta$, with σ chosen in Lemma 2.3.

In case (i), let ρ_1 have horizontal distance $\sigma - \operatorname{Re}\rho_1 < D$ and put $t_0 = \operatorname{Im}\rho_1$. We can choose a real t close to t_0 so that $z = \sigma + it$ has a unique closest pole, at distance $d < D$. Here is a justification that also addresses potentially equal distances. Take $a > 0$ so small that $D + a < \sigma$ and $\sigma - \operatorname{Re}\rho_1 + a < D$. All poles which could lie within distance D of $\sigma + it$ for $|t - t_0| < a$ lie in the fixed compact disk $\overline{B}(\sigma + it_0, D + a)$, hence form a finite set. For two distinct such poles, equality of their squared distances from $\sigma + it$ is a linear equation in t. It has at most one solution unless their imaginary parts agree. In the latter case equal distances would require equal horizontal distances; since both real parts are less than σ, that would make the poles equal. Avoiding finitely many values of t therefore gives a unique nearest pole, while ρ_1 guarantees that its distance is less than D.

Call that pole ρ, with nonzero residue c. Choose r with $d < r < D$ such that ρ is the only pole in $\overline{B}(z, r)$. Then $\Phi(w) - c/(w - \rho)$ is holomorphic near that closed disk. Cauchy's derivative estimate gives

$$\left| A_k(z) - \frac{c}{(z-\rho)^{k+1}} \right| \le M r^{-k}.$$

The disk $B(\sigma, D)$ is pole-free, so, after possibly decreasing r while keeping $r > d$, the same estimate gives $|A_k(\sigma)| \le M' r^{-k}$. Consequently the left side of (2.11) is at least $|c| d^{-k-1} - (M + M') r^{-k}$.

Choose $\eta > 0$ with $e^\eta d/E < 1$. Multiply (2.11) by d^{k+1} and let admissible U tend to infinity. Since $k \ge U$, its right side tends to zero, as do the terms involving $(d/r)^k$. The remaining lower bound tends to $|c| > 0$, a contradiction.

In case (ii), put $d = \sigma - \beta$. The residue c_β is positive, and subtraction of its principal part gives, for some r' with $d < r' < r$, $\operatorname{Re} A_k(\sigma) \ge c_\beta d^{-k-1} - M (r')^{-k}$. The left side of (2.11) is at least this quantity. The same multiplication by d^{k+1} and choice of η again give a contradiction. Both cases are impossible, proving the theorem. □

Corollary 2.5. *Under the analytic hypotheses of Theorem 2.2, if $g(u) \le e^{\theta u}$ on arbitrarily distant intervals $[U, (1+\delta)U]$, then Φ has no pole in $\operatorname{Re} s > \theta$.*

Proof. Take $R_U(u) = e^{\theta u}$. The integral in (2.6) is δU, and $\delta U \le e^{\eta U}$ for all sufficiently large U, for each fixed $\eta > 0$. □

Apply the corollary to $g(u) = H_\psi(e^u)$. All coefficients in (1.2) are positive, so (2.5) applies. If (1.3) holds for every sufficiently small $\varepsilon > 0$, the associated transform has no pole with real part greater than $1/2 + 9\varepsilon$. A zero ρ of $L(s, \psi)$ with $\operatorname{Re}\rho > 1/2$ would be a pole for every such transform. Choosing ε with $1/2 + 9\varepsilon < \min(\operatorname{Re}\rho, 1)$ gives a contradiction. The nonvanishing assertion therefore follows for every primitive character.

For completeness, we record the classical final reductions. If χ modulo N is induced by a primitive character χ^* of conductor f, Euler-factor stripping gives

$$L(s, \chi) = L(s, \chi^*) \prod_{\substack{p \mid N \\ p \nmid f}} (1 - \chi^*(p) p^{-s}). \tag{2.12}$$

Every factor on the right is nonzero in $\operatorname{Re} s > 0$, since $|\chi^*(p) p^{-s}| = p^{-\operatorname{Re} s} < 1$. Thus nonvanishing to the right of $1/2$ extends to imprimitive characters without any assumption concerning their additional zeros on $\operatorname{Re} s = 0$.

Let $a \in \{0, 1\}$ be determined by $\chi(-1) = (-1)^a$. For primitive nonprincipal χ, the completed function $\Lambda(s, \chi) = (f/\pi)^{(s+a)/2} \Gamma((s+a)/2) L(s, \chi)$ satisfies $\Lambda(s, \chi) = W(\chi) \Lambda(1 - s, \bar{\chi})$, with $|W(\chi)| = 1$. In $0 < \operatorname{Re} s < 1$, the gamma and power factors are finite and nonzero, so a zero on the left of the critical line gives one on the right for $\bar{\chi}$. Equation (2.12) then proves Theorem 1.2. For the character of conductor 1, the same conclusion follows from the functional equation of ζ.

Finally, the asymmetric functional equation $\zeta(s) = 2^s \pi^{s-1} \sin(\pi s/2) \Gamma(1 - s) \zeta(1 - s)$ shows that in $\operatorname{Re} s \le 0$, away from $s = 0$, the only zeros are the negative even integers: the gamma

factor is finite and nonzero there, $\zeta(1-s)$ is nonzero by Theorem 1.1, and the sine factor vanishes exactly at the even integers. At $s=0$, the continuation has value $-1/2$. Together with the strip theorem and the nonvanishing in $\operatorname{Re} s > 1/2$, this proves Theorem 1.3 once the arithmetic window estimate is established.

The source proof of this section is concentrated in `poleData.lean`, `landauTails.lean`, and `landauWindow.lean`. The passage from right-half-plane nonvanishing to the public statements is in `CritLine.lean` and `RhOfRight.lean`. These references locate the principal steps without interrupting the proof with the names of their elementary supporting lemmas.

3. Auxiliary primes and uniform character estimates

The arithmetic construction requires two kinds of estimates for Dirichlet characters. A fixed, finite family is allowed to remain in the final inequality. Characters introduced by two auxiliary quadratic extensions must instead have a square-root bound on a prescribed interval. All characters of the resulting abelian field must also have a much smaller relative error at the exponentially larger arguments arising from the thin weights. These are different requirements, and their proofs use different features of the prime selection.

The analytic ingredients are classical: a zero-free region with its possible exceptional real zero, the Landau–Page uniqueness principle, a weak zero-density estimate, and primes in arithmetic progressions with the exceptional term retained. Standard references for the first three parts of this framework include Davenport [9] and Huxley [16]. The construction-specific point is the simultaneous choice of the two primes, with prescribed splitting, a controlled family of quadratic twists, and enough uniformity for every argument in a fixed power interval. We distinguish below the analytic estimates supplied by the source from the finite and elementary reductions that combine them.

Fix an integer $M \ge 1$, an exponent $J \ge 1$, and $0 < \varepsilon < 1/52$. All thresholds in this section may depend on these fixed data. Put $X = Y^J$, $z = Y^{\varepsilon/5}$, and $z' = z(\log Y)^2$. For an odd prime $p \equiv 3 \pmod 4$, write χ_{-p} for the primitive quadratic character of discriminant $-p$. Thus $\chi_{-p}(a) = (a/p)$. Consider the two finite sets

$$\mathcal{P}_1 = \{q : e^{z/2} < q \le e^z,\ q \equiv 3 \pmod 8\}, \qquad R_q = q \prod_{3 \le p \le z} p,$$

$$\mathcal{P}_2(q) = \{d : e^{z'/2} < d \le e^{z'},\ d \equiv 7 \pmod{16},\ R_q \mid d+1\}.$$

where the letters p, q, d in these sets denote primes. For sufficiently large Y, the prime q is larger than z, so this product is squarefree. The Chinese remainder theorem writes the second set as primes in one reduced residue class modulo $16R_q$. The congruence $d \equiv -1$ modulo each odd prime dividing R_q is compatible with $d \equiv 7 \pmod{16}$, and the resulting class is coprime to $16R_q$.

The usual Chebyshev bound gives $\log(16R_q) \le 4z$ after increasing the threshold. Consequently $\frac{1}{4}(\log Y)^2 \log(16R_q) \le z'$. This is the modulus condition used below. It does not imply $\log(16R_q) = o(\sqrt{z'})$; a fixed-modulus prime number theorem is therefore insufficient for counting this family. The required progression estimate has to be uniform in this growing modulus.

Proposition 3.1 (`pairEx`). *For all sufficiently large Y, there are $q \in \mathcal{P}_1$ and $d \in \mathcal{P}_2(q)$, both coprime to M, with the following properties. If ξ is any Dirichlet character modulo M, if $(i,j) \in \{0,1\}^2 \setminus \{(0,0)\}$, and if ϑ is the primitive character inducing $\xi\chi_{-q}^{\,i}\chi_{-d}^{\,j}$, then*

$$|G_\vartheta(y)| \le y^{1/2+\varepsilon/2}, \qquad |\ell_\vartheta| \le 4\log y \qquad (Y \le y \le X).$$

Moreover, if ρ is any real primitive character of conductor dividing $8M$, the primitive character inducing $\rho\chi_{-q}^{\,i}\chi_{-d}^{\,j}$ has no zero in

$$\operatorname{Re} s > 1 - \frac{1}{400}, \qquad |\operatorname{Im} s| \le Q_j, \qquad Q_j = 8Me^z e^{jz'}.$$

The statement includes all characters modulo M, not just the powers of the particular character eventually retained. This distinction supplies the real quadratic characters needed for the thin estimate. The source proves the stronger assertion that, for each fixed $0 < \eta < 1$, a

proportion at least $1-\eta$ of $\mathcal{P}_1$ is suitable, and for each such q a proportion at least $1-\eta$ of $\mathcal{P}_2(q)$ is suitable. It also allows z' to vary between $z(\log Y)^2$ and $J^2 z(\log Y)^2$. The exact large-subset statement and its existence consequence are in `PairShare` and `pairEx`. The following argument explains their analytic inputs and the uniform selection mechanism.

For positive U and coprime integers a_0, m_0, with $8 \mid m_0$, put $\mathcal{P} = \{p : \sqrt{U} < p \le U,\ p \equiv a_0 \pmod{m_0}\}$. The range used in the source is

$$\frac{1}{4}(\log Y)^2 \log m_0 \le \log U, \qquad Y^{\varepsilon/6} \le \log U \le Y^{\varepsilon/5}(\log Y)^2. \tag{3.1}$$

The following three uniform inputs are sufficient. First, throughout (3.1), $|\mathcal{P}| \ge U^{9/10}$ for sufficiently large Y. Second, for any fixed $A, B > 0$, there is at most one squarefree odd integer $s_* > 1$, coprime to m_0, that must be excluded from the estimate

$$\left|\sum_{p\in\mathcal{P}} \left(\frac{p}{s}\right)\right| \le Y^{-B}|\mathcal{P}| \quad (1 < s \le Y^A,\ s \text{ odd and squarefree},\ (s, m_0) = 1,\ s \ne s_*). \tag{3.2}$$

Third, the number of primitive characters of conductor at most Q having a zero in $\operatorname{Re} s > 1 - 1/400$, $|\operatorname{Im} s| \le Q$ is $O(Q^{1/4})$. Only this weak consequence of zero density is required. For example, Huxley's hybrid estimate $(QT)^{40(1-\sigma)/9+\delta}$, with a sufficiently small fixed $\delta > 0$, gives a substantially smaller exponent when $\sigma = 399/400$ and $T = Q$ [16].

The precise progression estimate behind the first two inputs keeps the possible exceptional contribution in its main term and bounds the remainder relative to that main term. This avoids multiplying an error by the large factor $\varphi(m_0)$. The source derivations are `MSqCount`, `twistK`, and `CountBad`. These uniform analytic statements are inputs to the present exposition; they are not consequences of elementary orthogonality alone. The exceptional value s_* is the part of an exceptional conductor outside m_0. An exceptional character already induced from m_0 does not create further exceptional values of $s > 1$, because the characters occurring in the s-twisted progression sum have outside conductor exactly s.

It is useful to make the progression count more explicit. The source uses a relative estimate of the form

$$\left|\pi(U; m, a) - \frac{\operatorname{Li}(U) - \chi_*(a)\operatorname{Li}(U^{\beta_*})}{\varphi(m)}\right| \le C\left(e^{-c\log U/\log m} + e^{-\sqrt{c\log U}}\right)\frac{\operatorname{Li}(U) - \chi_*(a)\operatorname{Li}(U^{\beta_*})}{\varphi(m)},$$

once $\log U \ge C' \log m \log(\log m + 2)$. The exceptional term is omitted when there is no exceptional character. The constants are uniform; the exceptional character is real and its zero satisfies $0 < \beta_* < 1$. Siegel's theorem, used with a fixed exponent, gives $1 - \beta_* \ge c_*/m$, after allowing the constant to be ineffective. We may decrease c_* so that $0 < c_* \le 1$. In the worst sign $\chi_*(a) = 1$, and for $\log U \ge 1$,

$$U - U^{\beta_*} \ge U(1 - e^{-c_*/m}) \ge \frac{c_* U}{2m}, \qquad \operatorname{Li}(U) - \operatorname{Li}(U^{\beta_*}) \ge \frac{c_* U}{2m \log U}.$$

The second inequality follows by integrating $1/\log t$ on $[U^{\beta_*}, U]$; the exceptional zero is near 1, so the lower endpoint exceeds 2 in the range at issue. For the other sign, or with no exceptional term, the same lower bound holds for large U. Since $\varphi(m) \le m$, the progression main term is at least $c_* U/(2m^2 \log U)$.

Under (3.1), $\log m/\log U \le 4/(\log Y)^2$, and $\log\log U/\log U \to 0$ uniformly. Consequently $m^2 \log U = U^{o(1)}$, while the relative error tends uniformly to zero. The hypothesis on $\log U$ also follows from (3.1): its available ratio is of order $(\log Y)^2$, whereas $\log(\log m + 2) = O(\log Y)$. Thus $\pi(U; m, a) \ge 2U^{9/10}$ after increasing Y. Removing the primes at most $\sqrt{U}$ proves $|\mathcal{P}| \ge U^{9/10}$. This argument uses the correct modulus range without appealing to the inapplicable comparison with $\sqrt{\log U}$.

Here is the finite second-moment calculation. Its usefulness is that it requires cancellation for every outside squarefree part except one; it does not require the exclusion of an exceptional real

character by an ineffective choice of signs. For $y \geq 3$, $Z \geq y$, and $|\kappa(n)| \leq 1$, set

$$c_y(n) = \Lambda(n)\min(1, y/n), \qquad S_p(y) = \sum_{n \leq Z} c_y(n)\kappa(n)\left(\frac{n}{p}\right).$$

Assume $p > Z^2$ for every $p \in \mathcal{P}$. If $n, n' \leq Z$, let $s(n, n')$ be the product of those primes outside m_0 that occur to an odd power in nn'. Quadratic reciprocity and the fixed class of p modulo m_0 give

$$\left(\frac{n}{p}\right)\left(\frac{n'}{p}\right) = \epsilon_{n,n'}\left(\frac{p}{s(n,n')}\right), \qquad \epsilon_{n,n'} \in \{-1, 1\},$$

with $\epsilon_{n,n'}$ independent of p. The assumption $8 \mid m_0$ fixes the signs from 2 and quadratic reciprocity; primes dividing m_0 also contribute fixed signs. The inequality $p > Z^2$ excludes zero Jacobi symbols.

Since c_y is supported on prime powers, its outside squarefree part is either 1 or a single prime. Let

$$R = \sum_{\substack{n \leq Z \\ \text{outside squarefree part of } n=1}} c_y(n), \qquad A_\ell = \sum_{\substack{k \geq 1 \text{ odd} \\ \ell^k \leq Z}} c_y(\ell^k) \quad (\ell \nmid m_0), \qquad B_0 = \sup_{\ell \mid m_0} A_\ell.$$

The total coefficient of pairs with $s(n, n') = 1$ is at most $R^2 + \sum_\ell A_\ell^2$. For a fixed $s > 1$, it is at most $2B_0R + 2B_0^2$: if s is prime, one factor has outside part 1 and the other has outside part s; if s is the product of two distinct primes, these are the two outside parts; otherwise there are no contributing pairs. Expanding the square, taking absolute values of the factors $\kappa(n)\overline{\kappa(n')}$, and using (3.2) therefore yields

$$\frac{1}{|\mathcal{P}|}\sum_{p \in \mathcal{P}} |S_p(y)|^2 \leq R^2 + \sum_\ell A_\ell^2 + 2B_0R + 2B_0^2 + Y^{-B}\left(\sum_{n \leq Z} c_y(n)\right)^2. \tag{3.3}$$

This argument accounts for the mixed terms between prime powers supported inside and outside m_0. They cannot be omitted from the exceptional class.

For completeness, all coefficient bounds needed in (3.3) follow from an elementary Chebyshev estimate $\sum_{n \leq u} \Lambda(n) \ll u$. For a fixed prime ℓ, the powers $\ell^k \leq y$ contribute at most $\log y$, and the geometric tail contributes $O(\log \ell)$ when $\ell \leq y$. When $\ell > y$, the entire sum is $O(y \log \ell/\ell)$. Consequently $B_0 \ll \log(2y)$. The total contribution of all even powers is

$$\sum_{\ell, k \geq 1} \log \ell \, \min(1, y/\ell^{2k}) = 2y\int_{\sqrt{y}}^{\infty} \frac{\sum_{n \leq t} \Lambda(n)}{t^3}\, dt \ll \sqrt{y}.$$

The equality follows by writing the minimum as the integral of its decreasing kernel, so it remains valid without any differentiability assumption on the Chebyshev function. Suppose now that every prime dividing m_0 is at most $z = Y^{\varepsilon/5}$ or greater than Z. The additional contribution to R is at most $O(z \log(2y))$. Uniformly for $Y \leq y \leq Y^J$, this is $o(\sqrt{y})$, and hence $R \ll \sqrt{y}$.

To bound $\sum_\ell A_\ell^2$, first take $\ell \leq y$. The sum of all powers below y is $O(y)$. For each such prime, the powers above y contribute at most $O(\log \ell)$, since the first omitted power is already larger than y. Summing gives $\sum_{\ell \leq y} A_\ell \ll y$, and therefore $\sum_{\ell \leq y} A_\ell^2 \ll y \log(2y)$. For $\ell > y$, partial summation with $\sum_{\ell \leq t} \log \ell \ll t$ gives

$$\sum_{\ell > y} A_\ell^2 \ll y^2 \sum_{\ell > y} \frac{(\log \ell)^2}{\ell^2} \ll y \log(2y).$$

Finally, partial summation gives $\sum_{n \leq Z} c_y(n) \ll y(1 + \log Z)$. With $Z = X^{400}$, choose $A = 800J$ and $B = 2J + 2$ in (3.2). The last term of (3.3) is then $O(1)$, uniformly for $Y \leq y \leq X$. We have proved

$$\sum_{p \in \mathcal{P}} |S_p(y)|^2 \ll y \log(2y)\, |\mathcal{P}|, \qquad Y \leq y \leq X, \tag{3.4}$$

uniformly for every sequence $|\kappa(n)| \le 1$. This is the finite content of `MeanSqAbs` and its uniform form in `meanSqK`.

A bound at each fixed y is not yet a bound simultaneously for every y in the interval. The following maximal argument supplies exactly that step. Each S_p is absolutely continuous on compact intervals. Away from integers, $S_p'(t) = \sum_{t<n\le Z} \Lambda(n)\kappa(n)(n/p)/n$. The coefficient of $tS_p'(t)$ is at most $c_t(n)$, so (3.4) applies to $tS_p'(t)$ as well. For $Y \le T \le X$, put $V_T = \min(2T, X)$. For $T \le v \le V_T$, the fundamental theorem of calculus and Cauchy–Schwarz give

$$|S_p(v)|^2 \le 2|S_p(T)|^2 + 2T \int_T^{V_T} |S_p'(t)|^2\, dt.$$

Summing over p and applying (3.4) shows that $\sum_p \sup_{T\le v\le V_T} |S_p(v)|^2 \ll T\log(2T)|\mathcal{P}|$. The fraction of primes for which $|S_p(v)| > \frac{1}{2}v^{1/2+\tau}$ somewhere in this interval is therefore $O(T^{-2\tau}\log(2T))$. A geometric sum over the dyadic intervals covering $[Y, X] = [Y, Y^J]$, with the last interval truncated at X, tends to zero for every fixed $\tau > 0$. This proves the large-subset statement uniformly in y; a finite union handles all the boundedly many base characters. No passage from an uncountable union of fixed-y exceptional sets is being made.

Two remaining analytic operations convert this estimate into Proposition 3.1. If $\vartheta \ne 1$ is primitive, $Z \ge y$, and

$$T_\vartheta(Z) = -\ell_\vartheta - \sum_{n\le Z} \frac{\Lambda(n)\vartheta(n)}{n},$$

then a direct rearrangement of the finite sums gives

$$G_\vartheta(y) = -\sum_{n\le Z} \Lambda(n)\vartheta(n)\min(1, y/n) - yT_\vartheta(Z). \tag{3.5}$$

In particular, this formula has a minus sign in front of the truncated sum. The contour input used in the source says that, if the conductor is at most Q, the stated zero-free box holds, and $Y^{\varepsilon/6} \le \log Q \le Y^{\varepsilon/4}$, then

$$X\,|T_\vartheta(X^{400})| \le Y^{\varepsilon}. \tag{3.6}$$

The contour shift uses a positive slack $\varepsilon/(1600J)$. Its dependence on J is essential: it converts powers of $X = Y^J$ back into the required power of Y. The complete quantified estimate is in `TailKExp`; its use in (3.5) is in `BoxOfS`. Since $y \le X$ and $\varepsilon < 1/4$, $y|T_\vartheta(X^{400})| \le Y^{\varepsilon} \le \frac{1}{2}y^{1/2+\tau}$ for sufficiently large Y. Together with the maximal estimate this gives the asserted bound for G_ϑ.

The second operation removes primes for which the hard zero-free box fails. If a primitive base character has conductor at most $F \le \sqrt{U}$, then twisting it by χ_{-p}, $p \in \mathcal{P}$, introduces the new prime p into its conductor. A resulting primitive character can come from at most one such p: its conductor has the form fp with $f \le F < p$. The zero-density input thus discards only $O((FU)^{1/4}) = O(U^{3/8})$ primes. Dividing by $|\mathcal{P}| \ge U^{9/10}$ gives a quantity tending to zero. Here a primitive inducing character is used throughout; treating an imprimitive character as if its extra Euler factors were absent would invalidate the conductor argument.

Apply this procedure first to $U = e^z$, $m_0 = 8$, and all primitive characters inducing characters modulo $8M$. Apply it a second time, for each retained q, to $U = e^{z'}$, $m_0 = 16R_q$, and the primitive characters obtained by also allowing the twist χ_{-q}. Use the common conductor ceiling $F = 8Me^z \ge 8Mq$, which is less than $\sqrt{U}$ for large Y. Then the common conductor ceiling is $FU = 8Me^{z+z'}$, uniformly for every retained q. The prime divisors of m_0 are either at most z or equal to $q > Z$, as required for (3.4). The lower bound on $\log U$ also gives $p > Z^2$. The estimates on $\log Q$ used in (3.6) follow from $\log Q \le \log(8M) + z + z'$, because $(\log Y)^2 = o(Y^{\varepsilon/20})$. For the wider permitted range of z', the same proof is applied with the corresponding fixed rescaling of the power interval.

Lastly, the same zero-free box gives $|\ell_\vartheta| \le 4\log Y$ throughout the chosen family. This is the logarithmic-derivative input in `Kappa4Loglog`; the dependence on the conductor is controlled through $\log\log Q$, not by an uncontrolled multiple of $\log Q$. More precisely, the contour estimate at the truncation $Z_0 = (\log Q)^{1000}$, together with $\sum_{n\le Z_0} \Lambda(n)/n \le \log Z_0 + O(1)$,

gives $|\ell_\vartheta| \le 1000 \log\log Q + C$, with a constant independent of the selected character. Since $\log Q \le C_{M,J} Y^{\varepsilon/5} (\log Y)^2$, the right side is at most $200\varepsilon \log Y + 2000 \log\log Y + O_{M,J}(1)$. The strict inequality $200\varepsilon < 4$, which follows from $\varepsilon < 1/52$, absorbs the lower-order terms. This explains both the stated constant and the numerical restriction on ε at this point of the argument. Taking $\tau = \varepsilon/2$, and then taking the union of the finitely many exceptional sets, proves Proposition 3.1 from the explicitly stated analytic inputs. Siegel's theorem enters the uniform progression count. Its possible ineffectivity affects the threshold for Y, not the logical form of the existence statement.

The congruence conditions have a separate geometric purpose. The prime 2 is inert in $\mathbb{Q}(\sqrt{-q})$ and split in $\mathbb{Q}(\sqrt{-d})$. For every odd $p \le z$, and also for $p = q$, the congruence $d \equiv -1 \pmod p$ gives $(-d/p) = 1$. Thus all these primes split in $\mathbb{Q}(\sqrt{-d})$. This is why the second prime is chosen after the first, and why a selection of arbitrary large primes would not suffice.

Proposition 3.2 (`thinH`). *Fix $B, D > 0$ and $0 < \varepsilon < 1/52$. Let Q_0 be an integer whose odd part is at most B, and let q, d be distinct odd primes coprime to Q_0. Consider a family of primitive characters whose conductors divide $Q_0 qd$, and suppose $\log(Q_0 qd) \le D x^{\varepsilon/5} (\log x)^2$. Assume that every real nonprincipal member whose conductor is divisible by q or d has no real zero greater than $1 - 1/400$. Then, for sufficiently large x, uniformly for every character ϑ in this family, including the primitive principal character, and every $y \ge \frac{1}{3} e^{x^{\varepsilon/3}}$,*

$$|G_\vartheta(y)| \le y \exp\Bigl(-\tfrac{1}{2} x^{\varepsilon/8}\Bigr).$$

This is the uniform thin-scale consequence of the source theorem; its interface uses $D = 2$ and the characters of the host field. We give the reduction, including the scale comparison. A real primitive nonprincipal character has fundamental discriminant conductor. Its odd part is squarefree and its 2-part divides 8. If neither q nor d divides its conductor, that conductor is at most $8B$. There are only finitely many such characters, and $L(1, \vartheta) \ne 0$ gives a common zero-free real interval to the left of 1. The hypothesis supplies such an interval for the remaining real characters. Combining the two intervals with the classical zero-free region gives a constant $c > 0$, independent of x and the character, such that every zero $\beta + i\gamma$ with $\beta < 1$ satisfies

$$\beta \le 1 - \frac{c}{\lambda + \log(|\gamma| + 3)}, \qquad \lambda = \max(1, \log \operatorname{cond} \vartheta).$$

The functional equation supplies the corresponding control near the other edge of the critical strip when bounding the explicit formula.

Write $\psi_0(\vartheta, u) = \sum_{n \le u} \Lambda(n)\vartheta(n) - \mathbf{1}_{\vartheta=1} u$. The uniform explicit-formula consequence used here is

$$|\psi_0(\vartheta, u)| \le C u (\log u + \lambda)^2 \exp\Bigl(-\frac{c \log u}{\lambda + \sqrt{\log u}}\Bigr), \qquad u \ge e. \tag{3.7}$$

This is proved in `PsiThin`. It includes the principal character with its pole term subtracted. Partial summation, followed by the value of the regularized logarithmic derivative at 1, gives the exact identity

$$G_\vartheta(y) = -y \int_y^\infty \frac{\psi_0(\vartheta, u)}{u^2}\, du. \tag{3.8}$$

For a nonprincipal character, one can check this directly by first integrating to R, writing $\sum_{n \le R} \Lambda(n)\vartheta(n)/n$ by partial summation, and letting $R \to \infty$ using (3.7). The limiting sum is $-L'/L(1, \vartheta)$. In the principal case, $\sum_{n \le R} \Lambda(n)/n - \log R \to -\gamma_E$; substitution yields precisely the normalization $G_1 = S_1 - y \log y + (1 + \gamma_E) y$. Thus no expression $L'/L(1, 1)$ is introduced.

Put $v_0 = \log y$. For large x, $v_0 \ge \frac{1}{2} x^{\varepsilon/3}$ and $\lambda \le D x^{\varepsilon/5} (\log x)^2$. Because $61/300 > 1/5$, it follows uniformly for $v \ge v_0$ that $\lambda \le v^{61/100}$. Also $\sqrt{v} \le v^{61/100}$, and $(v + \lambda)^2 \le 4v^2$. Equations (3.7) and (3.8) give

$$\frac{|G_\vartheta(y)|}{y} \le 4C \int_{v_0}^\infty v^2 e^{-(c/2) v^{39/100}}\, dv \le C' e^{-(c/4) v_0^{39/100}}.$$

For the last inequality, substitute $t = v^{39/100}$ and absorb the resulting fixed power of t into half of the exponential. Finally, $39/300 = 13/100 > 1/8$, so the final bound is at most $\exp(-\frac{1}{2}x^{\varepsilon/8})$ after increasing the threshold. This proves the proposition from (3.7).

For the host constructed next, the required real-character hypothesis is supplied by Proposition 3.1. Indeed, a real primitive character of conductor dividing Q_0qd splits as a real primitive character of conductor dividing 8 $\mathrm{odd}(Q_0)$, times $\chi_{-q}^{i}\chi_{-d}^{j}$. When q or d occurs in the conductor, $(i,j) \neq (0,0)$. The bounded factor is included in the family of real characters modulo $8M$. This conductor argument, recorded in `RealBox`, is the necessary link between the finite twist selection and the thin estimate for every host character. The final source theorem is in `GaussSumFunctionEq`.

4. The host field and character extraction

We now construct the abelian field in which the retained characters and the auxiliary quadratic characters occur together. The construction has three features that will be used separately. A cyclic subgroup permits a sparse set of thick archimedean places with controlled Frobenius orbits. A biased distribution of top parameters produces positive coefficients for the retained characters. Finally, finite Fourier inversion turns a divisibility functional into a linear combination of the regularized prime sums. The group theory, the sphere construction of Behrend, and character orthogonality are classical. The particular combination of these choices, and its compatibility with the geometric inequality proved later, are the features specific to this argument. A claim of historical priority for each of these elementary components is neither needed nor intended.

Let ψ be primitive of conductor M and order o. Write its finite image as an abstract cyclic group B of order o, using additive notation, and let $\bar{\psi} : (\mathbb{Z}/M\mathbb{Z})^\times \to B$ be the associated surjection. For e sufficiently large put

$$Q_e = \mathrm{lcm}(2^{e+2}, M), \qquad H_e = (\mathbb{Z}/2^{e+2}\mathbb{Z})^\times/\{\pm 1\},$$
$$A = \mathrm{im}(\pi : (\mathbb{Z}/Q_e\mathbb{Z})^\times \longrightarrow H_e \times B), \qquad \pi(a) = (\mathrm{cl}(a), \bar{\psi}(a)).$$

Here H_e is cyclic of order 2^e. One can take the class of 5 as a generator: induction from $v_2(5^{2^r} - 1) = r + 2$ gives its order, and the units modulo 2^{e+2} are the disjoint union of its powers and their negatives. Both coordinate projections of A are surjective. Therefore, with $n = |A|$ and $A_0 = \ker(A \to B)$, one has $[A : A_0] = o$; moreover, the first projection embeds A_0 in H_e, so A_0 is cyclic of 2-power order $n_0 = n/o$.

There is a fixed integer $c \mid o$ such that $n = c2^e$ for all sufficiently large e. Indeed, the kernel of the projection $A \to H_e$ is the image under $\bar{\psi}$ of those units whose 2-adic component is 1 or -1. Once $e + 2 \geq v_2(M)$, the reduction of this condition modulo M is independent of e: the odd component is arbitrary and the 2-component is one of these two signs. Its image has a fixed order c, proving the assertion. Thus, for each sufficiently large x, we may choose e with

$$x^{2+8\varepsilon} \leq n \leq 2x^{2+8\varepsilon}. \tag{4.1}$$

Only existence of such an integer is needed; inclusive endpoints need not give a unique choice.

Take the primes q, d from Proposition 3.1, with Y large enough that they are coprime to Q_e. Define $w : \{\pm 1\} \to A$ by $w(1) = 0$ and $w(-1) = \pi(-1) = (0, \bar{\psi}(-1))$. This is a homomorphism, since $2\pi(-1) = 0$. On the units modulo Q_eqd, set

$$\mathrm{art}(a) = (\chi_{-q}(a), \chi_{-d}(a), \pi(a) + w(\chi_{-q}(a))) \in \{\pm 1\}^2 \times A. \tag{4.2}$$

The Chinese remainder theorem makes the two quadratic coordinates independent of the residue modulo Q_e. Given desired coordinates (s_q, s_d, a_0), first choose a unit modulo Q_e with π-image $a_0 - w(s_q)$, then choose residues modulo q, d with the indicated Legendre symbols. Hence art is surjective.

Let E be the subfield of $\mathbb{Q}(\zeta_{Q_eqd})$ fixed by $\ker(\mathrm{art})$. The elementary Galois correspondence identifies $\mathrm{Gal}(E/\mathbb{Q})$ with $\{\pm 1\}^2 \times A$ and gives $[E : \mathbb{Q}] = 4n$. Complex conjugation is $\mathrm{art}(-1) = (-1, -1, 0)$, because the two occurrences of $\pi(-1)$ cancel. Its fixed field F is totally real, and E/F is a totally imaginary quadratic extension. Both $\mathbb{Q}(\sqrt{-q})$ and $\mathbb{Q}(\sqrt{-d})$ lie in E; neither lies in F, so $E = F(\sqrt{-q}) = F(\sqrt{-d})$.

The parity correction in (4.2) also explains how every power of ψ, including an odd one, is retained. If $\chi = \psi^k$, let $\lambda_\chi : A \to \mathbb{C}^\times$ be the character obtained from the second projection. Define

$$\theta_\chi(s_q, s_d, a) = \lambda_\chi(a + w(s_q)).$$

Then $\theta_\chi(\mathrm{art}(v)) = \lambda_\chi(\pi(v) + 2w(\chi_{-q}(v))) = \chi(v)$. Thus the primitive inducing character of the pullback is exactly the primitive character belonging to ψ^k, with no unintended quadratic twist. The characters of the host trivial on A_0 are precisely the $4o$ pullbacks $\psi^k \chi_{-q}^{\,i} \chi_{-d}^{\,j}$, with $0 \le k < o$, $i, j \in \{0, 1\}$. This follows either by duality on the quotient, of order $4o$, or by checking that these characters are distinct: the factors at the coprime primes q, d detect i, j, and the remaining factor detects k. The three families with $(i, j) \neq (0, 0)$ are therefore exactly the characters to which the auxiliary square-root bounds apply.

The corresponding source construction is `AnyLevel`, with the field identification in `IsChiField` and the character decomposition in `TwistAny`. The preceding argument gives the group-theoretic reason for the parity correction, rather than treating it as a convention hidden in those files.

From now on interchange the two sign coordinates and write

$$G = (\mathbb{Z}/2\mathbb{Z})^2 \times A, \qquad c = (1, 1, 0), \qquad \Phi = \{(\alpha, \beta, a) : \alpha = 0\}, \qquad f = |\Phi| = 2n.$$

The first sign character $\eta(\alpha, \beta, a) = (-1)^\alpha$ is χ_{-d}, and the second is χ_{-q}. Let $\mathcal{K} = \{\theta_{\psi^k} : 0 \le k < o\}$ be the retained set and let $\Psi = \{\theta \in \widehat{G} : \theta|_{A_0} = 1\}$. Then $|\Psi| = 4o$, $\mathcal{K} \subset \Psi$, and

$$\theta(1, 0, 0) = 1 \quad (\theta \in \mathcal{K}). \tag{4.3}$$

For $t \in A$, the map $\rho_t(g) = c + g + (0, 0, t)$ is a bijection from $G \setminus \Phi$ to Φ. In particular, $\rho_t(1, 0, s + u - t) = (0, 1, s + u)$. This identity will locate the thick weights after Fourier inversion.

Proposition 4.1 (`hostThickExists`). *Here L_{orb} denotes the orbit length parameter. There is an absolute constant $C > 0$ such that, for integers $n_0, L_{\mathrm{orb}} \ge 2$, there is a set $T_2 \subset \{s \in \mathbb{N} : 4L_{\mathrm{orb}} s < n_0\}$ with*

$$|T_2| \ge \frac{n_0}{L_{\mathrm{orb}}} \exp\!\big(-C\sqrt{\log n_0 \log L_{\mathrm{orb}}}\big)$$

and the following property: whenever $s_1, s_2, s_3 \in T_2$ are pairwise distinct and $i, j \in \mathbb{Z}$ satisfy $0 < |i|, |j| \le L_{\mathrm{orb}}$, $i \neq j$, then $j(s_2 - s_1) \neq i(s_3 - s_1)$. For $n_0 = n/o$ as above and $L_{\mathrm{orb}} \asymp_\varepsilon \log x$, one can remove 0 and still arrange $\delta_2 = |T_2|/n \ge x^{-\varepsilon}$ for sufficiently large x.

Proof. This is the sphere construction of Behrend [3], with a larger radix to exclude all the indicated bounded coefficients. We include the construction because that strengthened relation is what controls the Frobenius orbits. Suppose first that $n_0 \ge (8L_{\mathrm{orb}})^{16}$, and put

$$S = \left\lfloor \sqrt{\frac{\log n_0}{\log(8L_{\mathrm{orb}})}} \right\rfloor, \qquad t = \left(\frac{n_0}{8L_{\mathrm{orb}}}\right)^{1/S}, \qquad d = \left\lfloor \frac{t}{4L_{\mathrm{orb}}} \right\rfloor, \qquad B_1 = 4L_{\mathrm{orb}} d.$$

These choices give $S \ge 4$, $d \ge 2$, and $B_1^S \le n_0/(8L_{\mathrm{orb}})$. Among the d^S vectors $v \in \{0, \dots, d-1\}^S$, the squared Euclidean norm takes at most $S(d-1)^2 + 1$ values. Some sphere therefore contains at least $d^S/(S(d-1)^2 + 1)$ such vectors. Map this sphere injectively to integers by $v \mapsto \sum_{a=0}^{S-1} v_a B_1^a$. Every image is less than B_1^S, hence satisfies $4L_{\mathrm{orb}} s < n_0$.

Suppose a forbidden relation holds between the images of three vectors $v^{(1)}, v^{(2)}, v^{(3)}$. Its a-th signed digit is $c_a = j(v_a^{(2)} - v_a^{(1)}) - i(v_a^{(3)} - v_a^{(1)})$, of absolute value less than $2L_{\mathrm{orb}} d < B_1$. Since $\sum_a c_a B_1^a = 0$, reduction modulo B_1 gives $c_0 = 0$; division by B_1 and induction give $c_a = 0$ for every a. Consequently $v^{(2)} - v^{(1)} = (i/j)(v^{(3)} - v^{(1)})$. All three vectors have the same norm. Expanding the two norm equalities gives $(i/j)(i/j - 1)\|v^{(3)} - v^{(1)}\|^2 = 0$. Since $i/j \notin \{0, 1\}$, the vectors cannot be pairwise distinct, a contradiction.

It remains to record the size, without hiding a loss of a power of n_0. We have $d \ge t/(8L_{\mathrm{orb}})$, so

$$\log |T_2| \ge \log n_0 - \log(8L_{\mathrm{orb}}) - S \log(8L_{\mathrm{orb}}) - \frac{2 \log n_0}{S} - \log(2S).$$

Here we used $S(d-1)^2 + 1 \le 2Sd^2$. Since S lies between $\frac{3}{4}\sqrt{\log n_0 / \log(8L_{\mathrm{orb}})}$ and $\sqrt{\log n_0 / \log(8L_{\mathrm{orb}})}$, the subtracted terms are $O(\sqrt{\log n_0 \log L_{\mathrm{orb}}})$, with an absolute constant.

This proves the claimed lower bound after increasing C. If $n_0 < (8L_{\mathrm{orb}})^{16}$, take $T_2 = \{0\}$. In this range $\log n_0 \le 64 \log L_{\mathrm{orb}}$, so a sufficiently large absolute C makes the claimed lower bound at most 1. The two cases prove the assertion for all $n_0, L_{\mathrm{orb}} \ge 2$.

For the application, $\log n_0 \asymp_\varepsilon \log x$ and $\log L_{\mathrm{orb}} = O_\varepsilon(\log\log x)$. Thus the logarithm of the reciprocal density is $O_{\psi,\varepsilon}(\sqrt{\log x \log\log x}) = o(\log x)$. The displayed lower bound tends to infinity, so removing the single element 0 changes it by at most a fixed factor. After dividing by $n = on_0$, that factor and the fixed o are absorbed by x^ε. □

Identify A_0 with $\mathbb{Z}/n_0\mathbb{Z}$, and view T_2 as a subset of A_0. Its consequence for orbits is particularly simple. For any $h, t \in A$ and $u \in A_0$, if $r = \mathrm{ord}(h) \le 4L_{\mathrm{orb}}$, then the orbit $t + \mathbb{Z}h$ meets $T_2 + u$ in at most one distinct point. Indeed, two intersections give $n_0 \mid r(s_2 - s_1)$, whereas $|r(s_2 - s_1)| < n_0$; hence $s_1 = s_2$. If $r > 4L_{\mathrm{orb}}$, any L_{orb} consecutive orbit positions meet $T_2 + u$ in at most two points. Otherwise three indices $i_1 < i_2 < i_3 < i_1 + L_{\mathrm{orb}}$ give three distinct labels s_1, s_2, s_3. Eliminating h gives

$$(i_3 - i_1)(s_2 - s_1) - (i_2 - i_1)(s_3 - s_1) \equiv 0 \pmod{n_0}.$$

The integer on the left has absolute value less than $n_0/2$, so the congruence is an equality, contradicting the proposition. The qualification about distinct points in a short orbit matters: a stationary orbit repeats the same point at every time. The proof does not assert that such repetitions are forbidden. The construction and orbit statements are recorded in `ThickExists` and `HostThickExists`.

The top parameters use a different finite construction. Put $\delta_1 = x^{-\varepsilon/8}$ and choose a random subset $T \subset A$, independently including a with probability δ_1 on A_0 and $3\delta_1/4$ off A_0. For a coset $C \in A/A_0$, let $N_C = |T \cap C|$ and $\mu_C = \mathbb{E}N_C$. Then $\mu_{A_0} = \delta_1 n_0$, $\mu_C = 3\delta_1 n_0/4$ otherwise, and $\mu = \mathbb{E}|T| = (3o+1)\delta_1 n/(4o)$. Independence gives $\mathbb{E}\sum_C (N_C - \mu_C)^2 = \sum_{a \in A} p(a)(1-p(a)) \le \delta_1 n$.

The following elementary selection principle explains how a geometric exceptional-set estimate can be imposed at the same time. Let $\mathcal{B}(T)$ be any nonnegative function of the random subset, and suppose $\mathbb{E}\mathcal{B}(T) \le b_0$. Markov's inequality shows that $\sum_C (N_C - \mu_C)^2 > 2\delta_1 n$ has probability at most $1/2$, whereas $\mathcal{B}(T) > 4b_0$ has probability at most $1/4$. If $b_0 = 0$, the latter function is zero almost surely. In either case there is a deterministic T such that

$$|N_C - \mu_C| \le \sqrt{2\delta_1 n} \quad \text{for every } C, \qquad \mathcal{B}(T) \le 4b_0. \tag{4.4}$$

The source applies this with $\mathcal{B}(T) = \sum_p \omega_p |\{t \in T : \mathrm{Bad}(p,t,T)\}|$, where the weights are nonnegative, the prime set is finite, and Bad is its specified failure of the orbit interval and tile conditions. The separate input $\Pr(t \in T, \mathrm{Bad}(p,t,T)) \le C_b x^{-2}\delta_1$ gives $b_0 \le C_b x^{-2}\delta_1 n \sum_p \omega_p$. The probability estimate, including its precise orbit predicate, is supplied by `topsGroup`; the finite selection principle itself is `select`, and its host application is `topsChoiceH`. The variance computation alone proves the distribution of top parameters; it does not prove the additional orbit probability estimate.

Let T_1 be a chosen subset satisfying (4.4). For fixed o, the errors there are $o(\delta_1 n)$. Thus $\frac{1}{2}\delta_1 n \le r := |T_1| \le 2\delta_1 n$ for large x. If $q_C = 4/(3o+1)$ on the zero coset and $3/(3o+1)$ on the other cosets, then

$$|N_C - q_C r| \le |N_C - \mu_C| + q_C |r - \mu| \le 3\sqrt{2\delta_1 n}.$$

Using (4.1) shows that the normalized error is at most $x^{-1+\varepsilon}$ for sufficiently large x. For a nontrivial character λ of A/A_0, the constant part of q_C sums to zero and the extra mass at the zero coset contributes $1/(3o+1)$. Hence

$$\left| \frac{1}{r} \sum_{t \in T_1} \lambda(t) - \frac{1}{3o+1} \right| \le o\, x^{-1+\varepsilon}. \tag{4.5}$$

The same statement holds with complex conjugates. The trivial character has average exactly 1. The positive coefficient in (4.5) is the reason for the ratio $1 : 3/4$ in the sampling probabilities. Uniform sampling would give zero as the limiting coefficient for every nontrivial character.

We next establish the identity connecting these finite choices to the prime sums. For each rational prime p, let $I_p \subset G$ be inertia and let $F_p \in G$ represent arithmetic Frobenius modulo I_p. For $\theta \in \widehat{G}$, put

$$v_\theta(p^a) = \frac{1}{|I_p|} \sum_{h \in I_p} \theta(aF_p + h) \qquad (a \geq 1).$$

If θ is nontrivial on inertia, this average is zero. Otherwise it is $\theta(F_p)^a$. These are precisely the prime-power coefficients of the primitive Dirichlet character attached to θ. In particular, use of the inertia average accounts for the Euler factors at ramified primes. Extend v_θ multiplicatively to the positive integers, with $v_\theta(1) = 1$. Let G_θ denote the regularized prime sum for that primitive character, with the principal normalization already fixed in (2.1), and put $\ell_\theta = L'/L(1,\theta)$ for $\theta \neq 1$, where $L(s,\theta)$ likewise denotes the attached primitive Dirichlet L-function.

For $a > 0$, $t \in A$, and a positive function $w : \Phi \to (0,\infty)$, define

$$\mathcal{A}(a,t,w) = \sum_p \sum_{j \geq 1} \frac{2\log p}{|I_p|} \sum_{g \in (jF_p + I_p)\setminus\Phi} \left(\frac{a}{p^j} - w(\rho_t(g)) \right)_+ ,$$

$$\mathcal{M}(a,w) = a \left(\log a - 1 - \gamma_E + \ell_\eta - \frac{1}{f} \sum_{\tau \in \Phi} \log w(\tau) \right).$$

Here $u_+ = \max(u,0)$. The first sum is finite, since a nonzero summand requires $p^j < a/\min_\Phi w$. This formulation also explains the explicit finite cutoff used in the source: any integer $N \geq a/\min_\Phi w$ bounds both the contributing primes and the relevant exponents. The number ℓ_η is real, since $\eta = \chi_{-d}$ is a real character and $L(1,\eta) \neq 0$. Thus $\mathcal{M}$ is real; writing $\operatorname{Re}\ell_\eta$ in its definition makes no change.

Proposition 4.2 (`ch_balance`)**.** *For every positive a, w as above,*

$$\mathcal{A}(a,t,w) - \mathcal{M}(a,w) = \frac{1}{f} \sum_{g \notin \Phi} \sum_{\theta \in \widehat{G}} \overline{\theta(g)}\, w(\rho_t(g)) G_\theta\!\left(\frac{a}{w(\rho_t(g))} \right). \tag{4.6}$$

Proof. Orthogonality on the finite abelian group, whose order is $2f$, gives

$$\sum_{\theta \in \widehat{G}} \overline{\theta(g)} v_\theta(p^j) = \frac{2f}{|I_p|} \mathbf{1}_{g \in jF_p + I_p}.$$

Multiply by $f^{-1}\log p\,(a/p^j - w(\rho_t(g)))_+$, then sum over p, j and $g \notin \Phi$. The resulting right side is $\mathcal{A}$. On the left, $(a/p^j - w)_+ = w(a/(wp^j) - 1)_+$, so the inner sum is the finite prime sum $wS_\theta(a/w)$. For $\theta \neq 1$, this equals $wG_\theta(a/w) - a\ell_\theta$. For the trivial character, it equals $wG_1(a/w) + a\log a - a\log w - (1+\gamma_E)a$.

The partner bijection gives $\sum_{g\notin\Phi} \log w(\rho_t(g)) = \sum_{\tau\in\Phi} \log w(\tau)$. Finally, $\sum_{g\notin\Phi} \overline{\theta(g)}$ is f for $\theta = 1$, is $-f$ for $\theta = \eta$, and is zero for every other character. Thus the nonprincipal linear terms contribute $a\ell_\eta$, and all the displayed elementary terms combine into $\mathcal{M}$. This proves (4.6). □

For this host, the invariant expression for $\mathcal{A}$ is exactly the more concrete divisibility sum needed later. Primes dividing $Q_e q$ are split in $\mathbb{Q}(\sqrt{-d})$, by the congruences of Section 3; their inertia and Frobenius cosets are contained in Φ, and they contribute nothing. At an unramified prime split in that quadratic field the same conclusion holds. At an unramified inert prime, $jF_p \notin \Phi$ exactly for odd j, and the coefficient is $2\log p$. At $p = d$, inertia is $\{0, (1,0,0)\}$, so every Frobenius coset contains exactly one point outside Φ. Its coefficient is $\log d$ for every integer exponent $j \geq 1$. There are no half-integer powers at the ramified prime. The source proves this identity, with the same inertia factors, in `ChBalance`.

The extraction from (4.6) uses two elementary continuity estimates. For a nonprincipal character, differentiation away from the finitely many breakpoints gives

$$\frac{d}{dw}\left[wG_\theta(a/w) \right] = -\sum_{n \leq a/w} \Lambda(n) v_\theta(n).$$

The term $w(a/w)\ell_\theta = a\ell_\theta$ is constant. For the principal character the derivative is $a/w - \sum_{n\le a/w}\Lambda(n)$. Since $\sum_{n\le u}\Lambda(n) \le 3u$, both derivatives have absolute value at most $3a/w$. Continuity at the breakpoints and integration give

$$|wG_\theta(a/w) - vG_\theta(a/v)| \le \frac{3a}{\min(v,w)}|w-v|. \tag{4.7}$$

The elementary bound with constant 3 follows from the central binomial coefficient. For an integer $N \ge 2$, the product of the primes associated with prime powers in $(\lceil N/2\rceil, N]$ divides $\binom{N}{\lfloor N/2\rfloor}$. Hence $\psi(N) - \psi(\lceil N/2\rceil) \le N\log 2$. Iterate with $N_j = \lceil N/2^j\rceil$ until $N_j = 1$. Before that time $N_j \le 2N/2^j$, so $\sum_j N_j < 4N$ and $\psi(N) \le (4\log 2)N < 3N$. Taking $N = \lfloor u\rfloor$ proves the stated bound for real u, with the case $u < 2$ immediate.

For $|y-x| \le 1$, partial summation gives $\sum_{n\le y}\Lambda(n)/n \le 3 + 3\log y$. Suppose that $|\ell_\theta| \le K_0$ on $\mathcal{K}\setminus\{1\}$ and $|\ell_\theta| \le C_0\log x$ on $\Psi\setminus\mathcal{K}$, with $K_0, C_0 \ge 0$. It follows, for $x \ge 3$, that

$$|G_\theta(y) - G_\theta(x)| \le V_x \quad (\theta\in\Psi), \qquad V_x = 4(1+\log(x+1)) + |\gamma_E| + K_0 + C_0\log x. \tag{4.8}$$

The principal character satisfies the same estimate by differentiating its explicit main terms. A crude bound needed only for the coefficient errors is

$$|G_\theta(x)| \le P_x \quad (\theta\in\mathcal{K}), \qquad P_x = x\{4(1+\log(x+1)) + 2 + |\gamma_E| + K_0\}. \tag{4.9}$$

Both estimates follow by integrating a bounded finite prime sum, so no unproved cancellation is used for the retained characters.

Proposition 4.3 (`ch_core`)**.** *Let $T_1 \subset A$ and $T_2 \subset A_0$ be nonempty, put $r = |T_1|$, $\delta_2 = |T_2|/n$, and let $m \ge 1$ be an integer. Assume $b, b', X' > 0$, $h, \vartheta \ge 0$, $x \ge 3$, and*

$$b'X' \le xb, \qquad 2f(x+1)\vartheta \le x^{-1}, \qquad 6f(x+1)h \le b/x.$$

For each $t \in T_1$, choose $a_t > 0$ with $|a_t/b - x| \le 1$. For every $u \in A_0$ and $1 \le j \le m$, let $w_{t,j,u} : \Phi \to (0,\infty)$ satisfy $b \le w_{t,j,u} \le b+h$ on $\{(0,1,s+u) : s \in T_2\}$, and $w_{t,j,u} \le 2b'$ elsewhere. Suppose $|G_\theta(y)| \le \vartheta y$ for every $\theta \in \widehat{G}$ and $y \ge X'/3$, and suppose the logarithmic-derivative bounds preceding (4.8) hold.

Write $c_0 = 1/(3o+1)$, set $q_1 = 1$ and $q_\theta = c_0$ for $\theta \in \mathcal{K}\setminus\{1\}$, and assume $|r^{-1}\sum_{t\in T_1}\theta(0,0,t) - q_\theta| \le \rho$ on $\mathcal{K}$, with $\rho \ge 0$. If, for every $u \in A_0$,

$$\sum_{t\in T_1}\sum_{j=1}^{m}\{\mathcal{A}(a_t,t,w_{t,j,u}) - \operatorname{Re}\mathcal{M}(a_t,w_{t,j,u})\} \le E,$$

then, with $\Sigma(x) = \sum_{\theta\in\Psi\setminus\mathcal{K}}|G_\theta(x)|$,

$$G_1(x) + c_0\sum_{\theta\in\mathcal{K}\setminus\{1\}}\operatorname{Re}G_\theta(x) \le \Sigma(x) + 4o(V_x + \rho P_x) + \frac{2E}{rm\delta_2 b} + \frac{4}{\delta_2 x}. \tag{4.10}$$

Proof. Let $\mathcal{D}$ denote the right side of (4.6). At every thin place the argument is at least $(x-1)b/(2b') \ge X'/3$. Hence its contribution has absolute value at most $2f\vartheta a_t \le b/x$, after summing over all characters and all thin places and dividing by f. At a thick place replace its weight by b. Equation (4.7) bounds the change in the whole sum by $6f(x+1)h \le b/x$. Thus

$$\left|\mathcal{D} - \frac{b}{f}\sum_{s\in T_2}\sum_{\theta\in\widehat{G}}\overline{\theta(1,0,s+u-t)}G_\theta(a_t/b)\right| \le 2b/x.$$

All normalizing factors are positive by hypothesis.

Average this identity over $u \in A_0$. Character orthogonality annihilates precisely the characters outside Ψ. For $\theta \in \Psi$, the factor involving $s+u$ is 1. Since $|T_2|/f = \delta_2/2$, summing over t, j and using the assumed inequality gives

$$\operatorname{Re}\left\{\frac{1}{r}\sum_{t\in T_1}\sum_{\theta\in\Psi}\overline{\theta(1,0,-t)}G_\theta(a_t/b)\right\} \le \frac{2E}{rm\delta_2 b} + \frac{4}{\delta_2 x}.$$

Replace a_t/b by x. Equation (4.8) bounds the change by $|\Psi|V_x = 4oV_x$. On $\mathcal{K}$, (4.3) removes the sign coordinate and (4.5), or the stated coefficient hypothesis, replaces the top average by q_θ. Equation (4.9) bounds the resulting error by $4o\rho P_x$. For each discarded character the top average has absolute value at most 1, so its real contribution is bounded below by $-|G_\theta(x)|$. Moving these terms to the right proves (4.10). □

This finite inequality identifies the exact role of the geometric result. It must give an upper bound for the gap at every translate $u \in A_0$, with the displayed choices of weights and top parameters. An assertion about a different average is insufficient unless its relation to this average is proved. The source extraction, including the later numerical absorption, is `ChCore`; the geometric input used for the individual translates is exposed through `ChTop`. Proposition 4.3 proves the Fourier reduction itself without presupposing that geometric bound.

The scale choices make its analytic errors small for transparent reasons. Take

$$b = e^{x^\varepsilon}, \qquad h = e^{x^{\varepsilon/4}} + 1, \qquad X' = e^{x^{\varepsilon/3}}, \qquad b' = \frac{xb}{X'}.$$

Proposition 3.2 allows $\vartheta = \exp(-\frac{1}{2}x^{\varepsilon/8})$. Since $f \asymp x^{2+8\varepsilon}$, both numerical smallness conditions in Proposition 4.3 hold for large x. If a third, still smaller weight is chosen as $b'' = b'/(xn)$, then $xb/b'' = xnX'$. This is the exact argument at the nominal top $a = xb$; an extra factor x must not be inserted. All these weights are strictly positive.

The retained logarithmic derivatives are bounded constants, because their primitive characters are fixed. Proposition 3.1 gives $\Sigma(x) \le 3o\,x^{1/2+\varepsilon/2}$ whenever $Y \le x \le Y^J$, and also gives the required discarded logarithmic-derivative bound. With $\rho = o\,x^{-1+\varepsilon}$, the term $4o(V_x + \rho P_x)$ is $O_\psi(x^\varepsilon \log x)$; the remaining elementary error is at most $4x^{-1+\varepsilon}$. The only substantial term left in (4.10) is therefore the normalized geometric gap $2E/(rm\delta_2 b)$. Sections 5 and 10 construct and estimate that gap. The arithmetic and finite Fourier arguments here isolate its required size and ensure that every retained nontrivial character has the same positive coefficient $1/(3o+1)$.

5. Mean slopes from approximate frames

The arithmetic construction produces many maps from small lattices to one large lattice. A useful estimate must retain the mean slope of the sources, rather than only their largest slope. This section proves the finite-dimensional and slope-theoretic mechanism. The underlying Harder–Narasimhan theory is classical; its use with a pooled approximate frame is the construction-specific step.

Let K be a number field of degree d. A hermitian $\mathcal{O}_K$-lattice E is a projective $\mathcal{O}_K$-module of finite rank together with compatible euclidean or hermitian norms on its archimedean scalar extensions. For a nonzero rational section v of $\det E$, its normalized arithmetic degree is

$$\widehat{\deg}_{\mathrm{Ar}}(E) = \frac{1}{d}\left(\sum_{\mathfrak{p}} \mathrm{ord}_{\mathfrak{p}}(v; \det E)\log N\mathfrak{p} - \sum_{\sigma: K\hookrightarrow\mathbb{C}} \log\|v\|_\sigma\right). \tag{5.1}$$

The local order is measured relative to a generator of $(\det E)_{\mathfrak{p}}$. It can be negative, since v need not belong to $\det E$. Multiplying v by $a \in K^\times$ changes the two sums by the same amount, by the product formula including the finite places. Hence (5.1) is independent of v. This definition avoids choosing a global integral basis, which need not exist.

For $r = \mathrm{rk}\,E > 0$, write $\mu_{\mathrm{Ar}}(E) = \widehat{\deg}_{\mathrm{Ar}}(E)/r$. Exact sequences are equipped with induced subspace and quotient norms; arithmetic degrees are then additive. Enlarging a lattice by an ideal, or changing its archimedean metrics, has a particularly transparent effect. If $I = \prod_{\mathfrak{p}} \mathfrak{p}^{-n_{\mathfrak{p}}}$ and the infinite norms are unchanged, then $\mu_{\mathrm{Ar}}(IE) = \mu_{\mathrm{Ar}}(E) + d^{-1}\sum_{\mathfrak{p}} n_{\mathfrak{p}} \log N\mathfrak{p}$. If every archimedean norm is multiplied by e^{-t}, the slope increases by t.

The source uses squared norms. Throughout this section and the quantitative assembly, we therefore use

$$\mu(E) = 2\mu_{\mathrm{Ar}}(E), \qquad h(B) = \frac{2}{d}\sum_{\sigma: K\hookrightarrow\mathbb{C}} \log\|B_\sigma\|_{\mathrm{op}} \quad (B \neq 0). \tag{5.2}$$

An operator norm is allowed to be less than 1, so $h(B)$ need not be nonnegative. If K is totally complex, with $r_2 = d/2$, put

$$\mathcal{H}(E) = 2^{r_2 r}|\Delta_K|^{-r/2}\operatorname{covol}_{\mathbb{R}}(E), \qquad \mu(E) = -\frac{2}{dr}\log \mathcal{H}(E). \tag{5.3}$$

The first equality removes the covolume of $\mathcal{O}_K$ in the Minkowski space, and the second agrees with (5.1)–(5.2). For a lattice generated by a K-basis this follows by taking the determinant of its Gram matrix at each complex place; the general case follows by multiplying by the finite lattice index. In particular, there is no additional discriminant shift once the normalized quantity $\mathcal{H}$ is used. The affine discriminant term appearing in the source's conversion from raw euclidean slopes has already been absorbed in (5.3).

We recall the Harder–Narasimhan filtration in the form needed here. The maximum slope is the largest slope of a nonzero saturated sublattice, and the minimum slope is the smallest slope of a nonzero torsion-free quotient, with quotient metrics. A lattice has a canonical filtration by saturated sublattices whose successive quotients are semistable and have strictly decreasing slopes. Repeating a quotient slope according to its rank gives real numbers $\lambda_1(E) \ge \cdots \ge \lambda_r(E)$ with $r^{-1}\sum_j \lambda_j(E) = \mu(E)$. The filtration subspace $E_{\ge t} \subset E \otimes K$ is the sum of the successive pieces with slopes at least t. Thus $\mu_{\min}(E_{\ge t}) \ge t$ whenever it is nonzero, whereas the nonzero quotient $E/E_{\ge t}$ has maximum slope strictly less than t. These statements, including attainment and the canonical filtration, belong to arithmetic reduction theory [13, 6]. We will use only these explicitly stated properties.

Lemma 5.1. *Let $B : E \to F$ be a nonzero integral map of hermitian lattices. Then $\mu_{\min}(E) \le \mu_{\max}(F) + h(B)$. More precisely,*

$$B(E_{\ge t}) \subset F_{\ge t - h(B)} \qquad (t \in \mathbb{R}). \tag{5.4}$$

Proof. Let $I = B(E)$, with the metric induced from F. The map $E/\ker B \to I$ is an isomorphism of $\mathcal{O}_K$-modules. At each embedding its determinant norm is at most $\|B_\sigma\|_{\mathrm{op}}^{\operatorname{rk} I}$, when the domain has the quotient norm. Formula (5.1) therefore gives $\mu(E/\ker B) \le \mu(I) + h(B)$. Saturating I in F can only increase its degree, so $\mu(I) \le \mu_{\max}(F)$. Since $\mu_{\min}(E) \le \mu(E/\ker B)$, the first assertion follows.

For the filtered assertion, compose the restriction of B to $E_{\ge t}$ with the quotient map $F \to F/F_{\ge t-h(B)}$. A quotient norm does not increase the operator norm, so the composite, if nonzero, has height at most $h(B)$. Its domain has minimum slope at least t and its target has maximum slope strictly less than $t - h(B)$, contradicting the first assertion. Thus the composite is zero, which is (5.4). □

The following frame estimate contains no number theory. For a linear map between finite-dimensional complex Hilbert spaces, $\|A\|_{\mathrm{HS}}^2 = \operatorname{Tr}(A^*A)$ is the squared Hilbert–Schmidt norm. All traces of positive operators below are real.

Proposition 5.2 (`frame_linear_algebra_first`)**.** *Let $B_i : V_i \to \mathcal{H}$ be finitely many linear maps, let $c_i > 0$, and put $R = \sum_i \dim V_i > 0$, $D = \dim \mathcal{H} > 0$, and $Q = \sum_i c_i B_i B_i^*$. Suppose*

$$\sum_i \|c_i B_i^* B_i - I_{V_i}\|_{\mathrm{HS}}^2 \le e_F R, \qquad \operatorname{Tr}(Q^2) \le (1+\eta_F)\frac{(\operatorname{Tr} Q)^2}{D}, \tag{5.5}$$

where $0 \le e_F \le 1/4$ and $\eta_F \ge 0$. If subspaces $N_i \subset V_i$ satisfy $\sum_i \operatorname{codim} N_i \le \varepsilon_F R$, then

$$\operatorname{codim}_{\mathcal{H}} \sum_i B_i N_i \le (\eta_F + 4\varepsilon_F + 4\sqrt{e_F})D. \tag{5.6}$$

Proof. Write $K_i = c_i B_i^* B_i - I_{V_i}$. Cauchy–Schwarz gives $|\sum_i \operatorname{Tr} K_i| \le \sqrt{R}(\sum_i \|K_i\|_{\mathrm{HS}}^2)^{1/2} \le R\sqrt{e_F}$. Consequently $T = \operatorname{Tr} Q \ge (1-\sqrt{e_F})R \ge R/2 > 0$.

Let $S = \sum_i B_i N_i$, let P be the orthogonal projection onto $S^\perp$, and let P_i project onto $N_i^\perp$. Since PB_i vanishes on N_i, it equals PB_iP_i. Positivity of $I - P$ then implies

$$\operatorname{Tr}(PQ) = \sum_i c_i \|PB_i\|_{\mathrm{HS}}^2 \le \sum_i \operatorname{Tr}(P_i(I + K_i)).$$

The sum of the identity terms is at most $\varepsilon_F R$. For the remaining terms, $|\operatorname{Tr}(P_i K_i)| \le \sqrt{\operatorname{rk} P_i}\,\|K_i\|_{\mathrm{HS}}$, and $\sum_i \operatorname{rk} P_i \le R$. A second Cauchy–Schwarz estimate gives $\operatorname{Tr}(PQ) \le (\varepsilon_F + \sqrt{e_F})R$.

Put $z = \operatorname{Tr}(PQ)/T$. Then $0 \le z \le 1$ and $z \le 2(\varepsilon_F + \sqrt{e_F})$. Applying Hilbert–Schmidt Cauchy–Schwarz to the operators $I - P$ and Q gives

$$(1-z)^2T^2 = \bigl(\operatorname{Tr}((I-P)Q)\bigr)^2 \le \dim S\,\operatorname{Tr}(Q^2) \le \dim S\,(1+\eta_F)T^2/D.$$

It follows that $\dim S/D \ge (1-z)^2/(1+\eta_F) \ge 1 - \eta_F - 2z$. The latter inequality follows on multiplying by $1+\eta_F$, since the difference is $z^2 + \eta_F^2 + 2\eta_F z \ge 0$. Substitution of the bound for z proves (5.6). □

The two assumptions in (5.5) have different roles. The first says that the individual source metrics are respected on average, after positive rescaling. The second says that the total frame operator has small variance about a scalar operator. Indeed, $\operatorname{Tr}[(Q - (\operatorname{Tr} Q/D)I)^2] = \operatorname{Tr}(Q^2) - (\operatorname{Tr} Q)^2/D$. Neither condition follows merely from the maps being defined over a number field, or from an individual scalar norm identity. In the application the maps are furnished by theta evaluation, while their collective trace estimate is a separate input.

Theorem 5.3 (`theoremFSharp`). *Let I be a finite nonempty index set, let E_i $(i \in I)$ and F be nonzero hermitian lattices over a totally complex number field, and let $B_i : E_i \to F$ be integral maps. Put $d_i = \operatorname{rk} E_i$, $R = \sum_i d_i$, and $p_i = d_i/R$. For every nonzero B_i, assume $h(B_i) \le h_i$.*

At one complex embedding, choose auxiliary Hilbert structures, a surjective linear projection $\pi : \widetilde{\mathcal{H}} \to F \otimes \mathbb{C}$, and lifts $\widetilde{B}_i : E_i \otimes \mathbb{C} \to \widetilde{\mathcal{H}}$ satisfying $\pi\widetilde{B}_i = B_i$. Assume $0 \le e_F \le 1/4$ and $\eta_F \ge 0$. Suppose these lifts, with positive coefficients c_i, satisfy (5.5), and suppose

$$\operatorname{rk} F \ge (1-\delta_F)\dim \widetilde{\mathcal{H}}, \qquad 0 \le \delta_F < 1. \tag{5.7}$$

For real s, $u \ge 0$, and $H_0 \ge 0$, assume

$$\sum_i p_i(\mu(E_i) - h_i) \ge s, \qquad \mu_{\max}(E_i) - h_i \le s + u, \qquad \mu_{\min}(F) \ge s - H_0. \tag{5.8}$$

Then, with $u\log(1 + H_0/u)$ interpreted as 0 when $u = 0$,

$$\mu(F) \ge s - \frac{4u\log(1 + H_0/u) + (\eta_F + 4\sqrt{e_F})H_0}{1-\delta_F}. \tag{5.9}$$

Proof. Pool the shifted source slopes $\lambda_j(E_i) - h_i$, giving each individual slope weight $1/R$. Their mean is at least s, and each is at most $s + u$. For $\tau > 0$, let $q(\tau)$ be the fraction of these slopes that are less than $s - \tau$. Comparing their mean with $q(\tau)(s-\tau) + (1-q(\tau))(s+u)$ gives $q(\tau) \le u/(u+\tau)$. This also holds with $q(\tau) = 0$ if $u = 0$: every shifted slope must then equal s.

Set $N_i = (E_i)_{\ge s-\tau+h_i} \otimes \mathbb{C}$. Lemma 5.1 shows that $B_i N_i$ is contained in $F_{\ge s-\tau} \otimes \mathbb{C}$; if $B_i = 0$, this inclusion is automatic. The total source codimension is $Rq(\tau)$. Proposition 5.2 bounds the codimension of $\sum_i \widetilde{B}_i N_i$ in $\widetilde{\mathcal{H}}$. Passing to the quotient $F \otimes \mathbb{C}$ cannot increase that codimension. Using (5.7), we find that the fraction of target slopes below $s - \tau$ is at most

$$\frac{1}{1-\delta_F}\left(\eta_F + 4\sqrt{e_F} + \frac{4u}{u+\tau}\right). \tag{5.10}$$

The auxiliary Hilbert structures at the chosen embedding enter only this dimension calculation. The map heights in (5.8) still involve all embeddings and the original arithmetic metrics.

Every target slope λ is at least $s - H_0$, and $\min(\lambda, s) = s - \int_0^{H_0} \mathbf{1}_{\{\lambda < s-\tau\}}\,d\tau$. Average this identity over the target slopes and use $\lambda \ge \min(\lambda, s)$. Integrating (5.10) proves (5.9) when $u > 0$. For $u = 0$ the fraction of omitted source slopes is zero for every $\tau > 0$, and direct integration gives the stated limiting formula. The case $H_0 = 0$ is immediate from the minimum-slope hypothesis. □

The weights p_i are determined by the ranks. They cannot be replaced by arbitrary probabilities while retaining the same unweighted frame hypothesis. An alternative weighting requires a corresponding replication or weighted frame statement. The number of maps also does not

appear in (5.9). A weaker estimate in the source includes a $\log|I|$ term, but the sharp theorem used here removes it. This matters when the family of evaluation maps is large.

For later use we give a convenient way to estimate the logarithmic loss without introducing spurious powers of x.

Lemma 5.4. *If $u, U, H_0, b, X \geq 0$, $u \leq U$, and $H_0 \leq Xb$, then*

$$u\log(1 + H_0/u) \leq (U + b)\log(1 + X), \tag{5.11}$$

with the same convention at $u = 0$.

Proof. For $u > 0$, concavity of log at $1+X$ gives $\log(1+H_0/u) \leq \log(1+X)+(H_0/u-X)/(1+X)$. After multiplication by u and use of $H_0 \leq Xb$, the right side is at most $u\log(1+X)+Xb/(1+X)$. Since $X/(1+X) \leq \log(1+X)$ and $u \leq U$, this proves (5.11). The limit at $u = 0$ is zero. □

This estimate is particularly useful when H_0/b is polynomial in x. Its contribution is then a logarithm of x times $U + b$, rather than the product UH_0. Similarly, if a frame error is exponentially small, it can be made smaller than every fixed negative power of x: for fixed $A, B, c > 0$, $x^A e^{-x^c} \leq x^{-B}$ for all sufficiently large x, because $x^c/\log x \to \infty$. The constants and exponents must be fixed before the threshold is enlarged. This elementary order of choices will recur in the final assembly.

The normalizations are defined in `StatementPl.lean`; the finite-dimensional statement is `FrameLinearAlgebraFirst.lean`. The sharp transfer is stated in `StatementFrameF.lean` and proved through `TheoremFSharp.lean`. These files separate the classical filtered-lattice properties from the particular frame argument, just as the proof above does.

6. Theta towers and the divisibility estimate

The arithmetic comparison uses the same theta pairing in three forms: as a composition of algebraic kernels, as an integral pairing of lattices, and as a complex period. Each form controls a different part of the argument. The algebraic kernels specify the weights and the denominators. Their expansions at finite places produce the divisibility gain. Their complex norms, together with the row spectrum, account for the cost of obtaining that gain. The essential compatibility is that all three descriptions concern the same pairing and the same lattices. A valuation estimate for one normalization cannot be combined with a norm estimate for another without recording the change of scale.

Theta correspondence, the seesaw identity, formal complex multiplication, and the Rallis inner product formula belong to the established literature [20, 12]. The construction-specific contribution is the choice of a growing tower and its weights so that the finite-place gain and the archimedean cost retain compatible leading terms. We describe this compatibility explicitly. The local coefficient estimate and the global realization of the tower are substantial inputs from the accompanying source; the elementary deductions from those inputs are proved below.

Let E/F be the CM extension already chosen, with $[F : \mathbb{Q}] = f$, and put $r = k + m - 1$ and $N_j = k + j(t + 1)$. Here k, m are even, $m \geq 2$, and t is odd, with the separation inequalities imposed by the parameter construction. It is useful to distinguish the four families of Hermitian spaces at once:

space	rank over E	archimedean condition
X_j	N_j	definite
Y_j	$r + jt$	prescribed at $j = 0$
W_j	r	$(k + j - 1, m - j)$ at every place
V_j	$r + t$	definite.

The index ranges are $0 \leq j \leq m$ when both endpoints are needed. Definiteness allows either sign; a positive sign is asserted only where the construction supplies it. The rank-r spaces W_j govern the row representations. The larger spaces Y_j occur in the column tower. In particular, the signature displayed for W_j is not a signature for Y_j. At $j = m$, W_m is positive definite. For $j < m$, its Hermitian symmetric domain has complex dimension $f(k + j - 1)(m - j)$.

There are isometric embeddings $X_j \hookrightarrow Y_j$ for $j < m$ and an isometry $Y_m \simeq X'$ onto a codimension-one subspace $X' \subset X_m$. The base space contains an orthogonal decomposition $Y_0 = X_0 \perp Ez_m \perp H^{m/2-1}$, with $\langle z_m, z_m\rangle = -T_*$ and H a hyperbolic plane. A column u uses a vector $z_m(u) = c(u)^{-1}z_m$, where $c(u)\overline{c(u)}T_u = T_*$; hence its norm is $-T_u$. Finally, every W_j is identified with Y_0 at the finite adeles in the manner required to compare the levels and the local types. Existence of spaces with these simultaneous properties is part of the tower construction. The relevant specification is `DataSpaces`. Nothing about the degree f alone implies that these unitary groups are anisotropic over F. Compactification and boundary estimates therefore remain part of the geometric argument when the quotients are not compact.

For a definite stage, let $\mathcal{C}_j$ be the finite set of lattice classes and let $\Gamma_{j,L}$ be the stabilizer of a class L. The column map is a weighted sum

$$(K_j\phi)(L') = \sum_{L\in\mathcal{C}_j} \frac{K_j(L,L')\phi(L)}{|\Gamma_{j,L}|}, \qquad K^{\mathrm{col}} = K_{m-1}\cdots K_0.$$

The kernel includes the determinant characters, the unfolded theta operator, and the prescribed powers of the determinant coordinates. Those powers are integers. In the source their two exponents differ by $(t+1)/2$, which explains the parity requirement on t. The determinant characters have complex absolute value one. This removes them from an archimedean norm bound, but it does not remove the obligation to check their integrality at a finite place.

A common positive integer D_Γ is divisible by every $|\Gamma_{j,L}|$ occurring in the tower. Thus a single kernel step with integral unweighted coefficients has coefficients in $D_\Gamma^{-1}\mathcal{O}$; a composition of s such steps has coefficients in $D_\Gamma^{-s}\mathcal{O}$. This simple observation is the reason stabilizer weights must be carried through the integral construction. It is not legitimate to replace the weighted sum by an unweighted sum merely because both define operators on the same complex vector space. The finite-sum definitions and the denominator condition appear in `DataKernels`.

The row tower runs in the opposite direction. For a tuple of local characters μ, write $\mathcal{A}_\mu$ for its top space and $\mathcal{E}$ for the retained bottom row space. Composing the downward maps gives $\iota_\mu : \mathcal{A}_\mu \to \mathcal{E}$. A self-adjoint idempotent e_μ selects a subspace V_μ of $\mathcal{A}_\mu$. The kernel that is discarded is $U_\mu = V_\mu \cap \ker \iota_\mu$, and the useful quotient is represented by $Q_\mu = V_\mu \cap U_\mu^{\perp}$. These intersections matter: the ambient kernel can contain components that never belonged to V_μ. The lattice on Q_μ is obtained by projecting the integral lattice in V_μ orthogonally, and the row lattice is its image under ι_μ. Orthogonal projection is an archimedean operation; its arithmetic denominators are controlled separately.

Lemma 6.1. *Let A and E_0 be finite-dimensional complex Hilbert spaces with $E_0 \neq 0$, let $T : A \to E_0$ be linear, and let $V \subset A$. Assume $T(V) = E_0$ and $T^*(E_0) \subset V$. Set $U = V \cap \ker T$ and $Q = V \cap U^{\perp}$. Then $Q = \operatorname{im} T^*$ and $T|_Q : Q \to E_0$ is an isomorphism. If $E_0 = \bigoplus_\pi E_\pi$ is a finite orthogonal sum of nonzero spaces and $TT^*|_{E_\pi} = \lambda_\pi \operatorname{id}$ with $\lambda_\pi > 0$, then*

$$\|(T|_Q)^{-1}\|^2 = (\min_\pi \lambda_\pi)^{-1}.$$

Consequently, a functional P on V vanishing on U has a unique factorization $P = B \circ T|_V$, and $\|B\|^2 \le \|P|_Q\|^2/(\min_\pi \lambda_\pi)$.

Proof. Since $T(V) = E_0$, the map T is onto. Thus T^* is injective and $A = \ker T \oplus \operatorname{im} T^*$ orthogonally. The second summand lies in V, so restricting this decomposition to V gives $V = U \oplus \operatorname{im} T^*$. This proves the assertion about Q and the bijectivity of $T|_Q$. For $y \in E_0$ the inverse is $T^*(TT^*)^{-1}y$. Its squared norm is $\langle (TT^*)^{-1}y, y\rangle$, so the spectral decomposition gives the claimed operator norm. Finally define $B = P|_Q \circ (T|_Q)^{-1}$. The orthogonal decomposition of V proves the factorization, surjectivity gives uniqueness, and the norm inequality follows by composition. □

In the theta application, the eigenvalues take the form $\kappa_R \prod_{j=1}^m \lambda_{j,\pi}$, after incorporating the common positive normalization factor into κ_R. Their positivity and the identification with the row operators come from the inner product calculation. The classical Rallis formula relates the Petersson inner product of theta lifts to a doubling zeta integral; its application here also

requires the convergence or regularization hypotheses, the local measures, and the chosen test vectors. The source passage `RallisDM` makes these dependencies explicit. The displayed scalar is therefore part of the normalization data, not a universal constant of theta lifting.

Let P_u denote the terminal period functional, before composition with the column kernels. The lemma explains the exact form of the height estimate. When $B_u \neq 0$, all the factors below are nonzero; at a complex embedding σ, submultiplicativity yields

$$\log \|B_u\|_\sigma^2 \le \log \|P_u\|_\sigma^2 + \sum_j \log \|K_j\|_\sigma^2 - \log \kappa_{R,\sigma} - \min_\pi \sum_j \log \lambda_{j,\pi,\sigma}. \tag{6.1}$$

All the large terms in the last two summands must be retained until the kernel estimates have been inserted. Bounding the inverse row map by an unspecified constant would lose precisely the cancellation needed later.

Proposition 6.2 (`BesselId`)**.** *For the compatible tower data, the algebraic pairing B_u is defined over the coefficient field. There are a single unitary change of coordinates V on $\mathcal{E}$, unitary changes of coordinates U_u on the column spaces, and nonzero scalars $a_u \in \mathbb{C}$ such that $B_u(\phi, F) = a_u \operatorname{Bes}_u(U_u\phi, VF)$. Here Bes_u is the Bessel period belonging to the vector $z_m(u)$ and to the specified measures and test vectors.*

This is the rationality and comparison input recorded in `GlobalTheoremInputs`. It identifies the pairing whose integrality is studied with the pairing whose spectral correlations are estimated. It does not assert that the correlation operators are scalar or nearly scalar. Those estimates are a separate input in Section 7. In particular, the scalars a_u must be included when the periods are normalized; unitary changes of coordinates preserve operator norms, whereas arbitrary scalar changes do not.

We next explain exactly how a local gain changes a global lattice. Let H be the coefficient field, $d_H = [H : \mathbb{Q}]$, and let $\ell_w = \log Nw/d_H$ for a finite place w. Write v_w for the integer valuation, normalized by $v_w(\varpi_w) = 1$, and $e_w = v_w(p)$ if $w \mid p$. The convention throughout is $\mu_2 = 2\mu_{\mathrm{Ar}}$. Thus a fractional-ideal change has twice the usual normalized logarithmic contribution to the slope.

Lemma 6.3. *Let N and R be full $\mathcal{O}_H$-lattices and let $B : N_H \times R_H \to H$ be bilinear. Suppose integers a_w, zero outside a finite set, satisfy $B(N_w, R_w) \subset \varpi_w^{a_w}\mathcal{O}_{H,w}$ for every finite place. Set $I = \prod_w \mathfrak{p}_w^{-a_w}$ and $N' = IN$. Then $B(N', R) \subset \mathcal{O}_H$, and, with the same archimedean metrics,*

$$\mu_2(N') = \mu_2(N) + 2\sum_w \ell_w a_w.$$

In particular, if $a_w = \lfloor e_w s_w \rfloor$, then

$$0 \le 2\sum_w \ell_w e_w s_w - (\mu_2(N') - \mu_2(N)) < 2\sum_{w \in S} \ell_w$$

whenever S is finite and nonempty and $s_w = 0$ for $w \notin S$.

Proof. Localization gives $I_w = \varpi_w^{-a_w}\mathcal{O}_{H,w}$, so bilinearity makes $B(N'_w, R_w)$ integral at every w. An element of H integral at all finite places belongs to $\mathcal{O}_H$, which proves the first assertion. If $n = \operatorname{rk} N$, then $\det(IN) = I^n \det N$. The normalized arithmetic degree therefore changes by $-n \log N(I)/d_H$; multiplying by two and dividing by n proves the formula for μ_2. Finally $0 \le e_w s_w - \lfloor e_w s_w \rfloor < 1$ at each place. Sum these inequalities with weights $2\ell_w$. □

The lemma also applies when the row lattice is first specified locally and then patched by intersection. What must be checked is that these local lattices are the localizations of the chosen global row lattice. This is the content of the row-lattice compatibility in the source, rather than a consequence of the complex comparison alone. For the column u, the shift used there is $s_{u,w} = \sum_{j=1}^m (g_{j,w} + F_{j,w} - E_{u,j,w})$, with the designated discarded-prime gains set to zero. The three terms respectively measure the jet gain, the row-side finite contribution, and the cost of the column-dependent change of level. The resulting integral map is $N'_u \to R^\vee$. Source definitions of the ideal, its rounding, and the local-to-global row lattice are in `DataDIVDefs`.

The normalization by d_H in ℓ_w is what keeps the rounding loss uniform when the coefficient field grows. If the places at which a nonintegral shift is rounded all lie above rational primes $p \le X$, then $\sum_{w|p} \ell_w = (\log p/d_H) \sum_{w|p} f_w \le \log p$, because $\sum_{w|p} e_w f_w = d_H$ and $e_w \ge 1$. The rounding loss is therefore at most $2\sum_{p\le X} \log p$, independently of d_H. Chebyshev's bound makes this $O(X)$. Places above distinguished larger primes must be checked separately; the estimate is not permission to discard them. This is the reason the source records the support of the shift and the comparison with the original column lattice in addition to the valuation inequality itself.

The finite-place estimate is easiest to understand on a single Frobenius orbit. Fix $p \ge 5$ and an even orbit length $d \ge 2$, with indices in $\mathbb{Z}/d\mathbb{Z}$. The formal logarithms are

$$\mathcal{L}_s(X) = \sum_{h\ge 0} p^{-h} X_{s-2h}^{p^{2h}}. \tag{6.2}$$

This is an identity of formal power series: any coefficient receives only finitely many contributions. Analytic convergence on the whole unit disc is neither asserted nor needed. The shift by two preserves parity; the two parities encode the two sides of the Hermitian pairing.

The expansion used in the source contains more information than the existence of these logarithms. After a change of coordinates by a unit, it has an integral factor times the exponential of a sum of copies of a series b. The coefficients of b lie in the completed maximal unramified coefficient field. It is invariant under $X_s \mapsto \zeta^{p^s} X_s$ for even s and $X_s \mapsto \zeta^{-p^s} X_s$ for odd s, where $\zeta^{p^d-1} = 1$. Every nonzero monomial has positive degree in both parities, and bidegree $(1,1)$ is absent. Moreover, the exponentials of the two additive coboundaries of b, after substituting the formal logarithms, are integral with constant coefficient one. The last condition records the integral biextension law. It is essential to the coefficient estimate; invariance under roots of unity alone does not give a lower bound for a coefficient's valuation.

For clarity, if $\delta_+ b = b(X^+ + Y^+, X^-) - b(X^+, X^-) - b(Y^+, X^-)$, one of these requirements is the integrality of $\exp(\delta_+ b(\mathcal{L}))$; there is an analogous requirement for addition in the negative variables. The source permits an integral power-series factor in the original expansion and does not require that factor itself to be a unit. These are the admissibility conditions in `OrbJetBase`.

Lemma 6.4. *If a formal series is invariant under the preceding root-of-unity action, then a nonzero coefficient of $\prod_s X_s^{n_s}$ satisfies*

$$\sum_{s \text{ even}} p^s n_s - \sum_{s \text{ odd}} p^s n_s \equiv 0 \pmod{p^d - 1}.$$

Proof. Choose a primitive (p^d-1)st root ζ. Comparing the indicated coefficient before and after the substitution multiplies it by ζ^a, where a is the displayed signed sum. A nonzero coefficient can be unchanged only when $\zeta^a = 1$. □

This congruence is the origin of the orbit combinatorics, but it is only one part of the argument. For example, multiplying an invariant monomial by p^{-M} preserves the congruence while making its valuation arbitrarily negative. The integral coboundary conditions prevent this freedom. Also, an odd orbit displacement is not obtained by setting $h = i/2$ in (6.2): the index h is an integer and the exponents are p^{2h}. Odd displacements arise from pairing the two parity classes and from the constrained distribution of degrees.

Here is a precise form of the input that retains those constraints. For each $s \in \mathbb{Z}/d\mathbb{Z}$ choose a Young diagram ν_s with at most N_0 rows, where $N_0 < N_1$. Put $A_s = (\nu_s)_1$, and let $\mathcal{T}$ be the distinguished set of orbit positions. Set $c_s = (\nu_s)_2$ and $D_s = |\nu_s| - A_s$ for $s \in \mathcal{T}$, and $c_s = A_s$, $D_s = |\nu_s|$ otherwise; a missing second row has length zero. For nonnegative numbers D_s, define

$$\mathcal{F}_p(D) = \frac{p}{p^2-1} \sum_s D_s - \max_x \sum_{s,i} x_{s,i}, \quad \sum_i p^i x_{s,i} + \sum_i x_{s-i,i} \le D_s, \tag{6.3}$$

where $x_{s,i} \ge 0$ and the index i ranges over odd integers $1 \le i < d$. The feasible set is nonempty and compact, since the zero array is feasible and $x_{s,i} \le D_s/p^i$. Thus the maximum exists.

Summing the constraints shows $(p+1)\sum_{s,i} x_{s,i} \le \sum_s D_s$, and hence

$$\frac{1}{p^2-1}\sum_s D_s \le \mathcal{F}_p(D) \le \frac{p}{p^2-1}\sum_s D_s. \tag{6.4}$$

If all $D_s = D$, equality holds in the lower bound: take $x_{s,1} = D/(p+1)$ and all other entries zero. Thus the optimization measures a genuinely coupled capacity constraint rather than a sum of independent one-variable estimates.

For $s \in \mathcal{T}$, let d_s be the least positive odd integer with $s + d_s \in \mathcal{T}$, or $+\infty$ if there is none. Define

$$\mathcal{T}_s = \sum_{\substack{i\ge 1 \text{ odd}\\ i<d_s}} (A_s/p^i - c_{s+i})_+ + \sum_{\substack{i\ge 1 \text{ odd}\\ i>d_s}} A_s/p^i, \qquad C_s = \max\bigl(\{c_{s-i} : i \ge 1 \text{ odd},\ p^i \le |\nu_{s-i}|\}\cup\{0\}\bigr).$$

The first sum is bounded termwise by a geometric series, and the maximum is over a finite set. The term at $i = d_s$ is deliberately excluded.

Proposition 6.5 (`Loc.ThmA`)**.** *The admissible local jet construction supplies an absolute constant C with the following property. Suppose $p \le a$, $|\nu_s| \le a$ for every s, and $2\log_p a + 2 \le N_0 - 1$. Apply the products of column minors associated with the diagrams ν_s as differential operators to the theta series, then evaluate at zero. Every resulting polynomial coefficient has p-adic valuation at least*

$$\sum_{s\in\mathcal{T}} \left(\mathcal{T}_s - \frac{p}{p^2-1}C_s\right) + \mathcal{F}_p(D) - CdN_0N_1(\log_p a + 1). \tag{6.5}$$

The valuation is normalized by $v_p(p) = 1$. The differential operators include their factorial coefficients; they are not divided-power operators unless the corresponding change of lattice is also made.

This is a cited construction input. Its exact coefficient formulation is the predicate for the orbital jet estimate in `OrbJetBase`; the matrix-entry passage is developed in `TheoremA MatrixEntries`. The distinction between an integral minor basis and an arbitrary complex highest-weight basis is important here. Integrality is proved for the former and then transported through explicitly controlled changes of lattice. Formula (6.5), including its capacity loss and its logarithmic error, is the specialized quantitative statement. The classical theory of Lubin–Tate groups explains the formal endomorphisms; it does not by itself prove this estimate for the chosen Schur modules.

The proof route through `DworkPowerSeriesWeights` first identifies the admissible biextension series with the integral exponential model, then transfers a frame-jet coefficient estimate to the product-of-minors operator. This supplies (6.5) with coefficient 1 for its final logarithmic loss. Keeping the symbol C above is convenient when this estimate is combined with subsequent changes of lattice; it does not conceal a constant depending on the orbit or on the Young diagrams.

There is also a useful stability estimate for the residual optimization. Suppose $k \ge 3$, $w_s > 0$, and $0 \le E_s \le 4w_s$, and assume $D_s = (k-1)w_s + E_s$ outside $\mathcal{T}$ and $D_s = (k-3)w_s + E_s$ on $\mathcal{T}$. The local comparison input gives

$$\mathcal{F}_p(D) \ge \mathcal{F}_p((k-1)w) - \frac{11}{p}\sum_s w_s. \tag{6.6}$$

Together with (6.5), this isolates the large base weight $(k-1)w$ while charging a separately controlled cost for the modified first rows. The more refined version replaces the $\mathcal{T}_s$ in that estimate by

$$\mathcal{B}_s = \sum_{\substack{i\ge 1 \text{ odd}\\ i\ne d_s}} \bigl((A_s - (\nu_s)_2)/p^i - c_{s+i}\bigr)_+.$$

Indeed $\mathcal{B}_s \le \mathcal{T}_s$: before d_s this follows by monotonicity in the numerator, after d_s each summand is at most A_s/p^i, and both exclude d_s. The source keeps the two constants separate, using 11 for the loss in (6.6) and 1 for the coefficient error. This companion estimate is recorded in

`LocalBoundElevenComparison`. It is part of the specialized local input, rather than a general property of arbitrary perturbations of a linear program. Its hypotheses explain exactly which perturbations of the diagram are permitted by the global weight construction.

The positive parts that survive this calculation have a simple arithmetic interpretation. For $\alpha \ge 0$ and numbers $w_i \ge 0$, the geometric identity $\sum_{i\ge 1,\ i \text{ odd}} p^{-i} = p/(p^2-1)$ gives

$$\frac{p}{p^2-1}\alpha - \sum_{i\ge 1,\ i \text{ odd}} \min(\alpha/p^i, w_i) = \sum_{i\ge 1,\ i \text{ odd}} (\alpha/p^i - w_i)_+. \tag{6.7}$$

Indeed $z - \min(z,w) = (z-w)_+$ term by term, and both series converge by comparison with $\alpha\sum p^{-i}$. If $w_i \ge 1$, the right side vanishes whenever $p^i \ge \alpha$. More generally, truncating after odd indices $i \le J$ costs at most $\alpha p^{-(J+1)}/(1-p^{-2})$ when $J+1$ is the next odd index. Multiplication by $\log p$ makes these terms compatible with the prime powers in a von Mangoldt sum. This is the elementary reason that local divisibility eventually meets the analytic ramp sum of the earlier sections.

The local lattices must also absorb the coupling between adjacent weights. Their construction uses intersections of integral dual Weyl lattices over suitable discrete valuation rings. The auxiliary ring conditions include completeness, a prescribed Frobenius, a square root of p, and the required residue field and square-root conditions; they are recorded in `UnramifiedQuadraticSetting`. Passing to the lattice that satisfies the jet constraints has a cost $\tau_w + \iota_w$. More precisely, away from the distinguished column prime, the new lattice is locally contained in $\mathfrak{p}_w^{\lfloor e_w \tau_w \rfloor_+}$ times the original ramified column lattice, and its normalized local colength is at most $d_u(\tau_w + \iota_w)$. If no orbit is served at that place, the source sets all three of τ_w, ι_w, and the colength equal to zero.

At the global interface this produces, for an inert prime in the specified range, numbers $g_{j,w}$ satisfying $0 \le g_{j,w} \le 4|\mathcal{T}|a/p$ and

$$\sum_{j=1}^{m} \sum_{t\in\mathcal{T}_w} (\mathcal{V}_{j,t,p} - L_p) - mF_p \le \sum_{j=1}^{m} g_{j,w} - \tau_w. \tag{6.8}$$

Here $\mathcal{V}_{j,t,p}$ is the finite odd-power positive-part sum of (6.7) for the corresponding weight, and L_p, F_p are the specified local error functions. In particular, $F_p = CfN_m^2 \log_p a + Cf(\log_p a)^2$. The more detailed error L_p has different small-, medium-, and large-prime pieces; the medium-prime piece is allowed a residual $b^{\delta(x)}$ with $\delta(x) \to 0$. Replacing this residual by a constant is not justified. The point of (6.8) is that the cost τ_w has already been subtracted from the claimed gain. Source interface `LinkG` in `GlobalTheoremInputs` also records containment of the served orbits and the mean auxiliary bound needed to sum these inequalities.

At infinity, the normalization is governed by the degree of an integral CM eigenline. Write $\mathrm{rd}_E = |\Delta_E|^{1/[E:\mathbb{Q}]}$ for the root discriminant and put $v = \mathrm{rd}_E/2$. The normalized period used in the tower is

$$\log \Omega = \gamma_E + \log(\pi/v) - \operatorname{Re} \frac{L'(1,\chi_{-d})}{L(1,\chi_{-d})}. \tag{6.9}$$

Here is the passage from the elliptic period to this eigenline degree. Let L_τ be the integral eigenline carrying the CM character selected by the elliptic action. Over a coefficient field K splitting E, let $n_\sigma > 0$ be the squared norms of the distinguished integral Lie vector of the auxiliary CM elliptic curve, and put $P_{\mathrm{ell}} = -[K:\mathbb{Q}]^{-1}\sum_\sigma \log n_\sigma$. The elliptic period theorem and the compatible rank-one tensor construction give, respectively,

$$\begin{aligned} P_{\mathrm{ell}} &= \gamma_E + \log(2\pi/\sqrt{d}) - \operatorname{Re}\ell_{\chi_{-d}}, \\ 2\widehat{\deg}_{\mathrm{Ar}}(L_\tau) &= P_{\mathrm{ell}} + \tfrac{1}{2}\log d - \log \mathrm{rd}_E = \log\Omega. \end{aligned} \tag{6.10}$$

The tensor identity uses a coefficient field splitting E, a positive rank-one CM datum, and the specified tensor identification and polarization. These are part of the compatible construction data. The vector whose norm defines P_{ell} is a Lie vector; passing to a dual differential would change the norm convention. In particular, $P_{\mathrm{ell}} = \log\Omega + \log(\mathrm{rd}_E/\sqrt{d})$, so the elliptic Lie period and the eigenline degree must be distinguished.

The source defines these quantities separately in `EllipticFibrePrincipalData.lean`. The elliptic period is constructed in `ExistsCmEllipticCurveIntegersWithPeriod.lean`; the rank-one degree identity is proved in `TwoMulDegreeEigenLatticeTensorRankOne.lean`. The base-change reduction and cancellation in (6.10) are assembled in `CmEigenLineDegreeFormula.lean`. Classical CM period and Kronecker limit formulas underlie the first identity [25]; the integral tensor construction supplies the correction in the second. That correction is already absorbed in Ω and is separate from the common row-twisting contribution used in Section 10.

One elementary estimate explains the shape of the archimedean answer. Write $\Phi(y) = y\log y - y$ for $y > 0$ and $\Phi(0) = 0$.

Lemma 6.6. *Let G be entire on $\mathbb{C}^s$, and suppose $|G(z)| \le S\exp(\sum_i r_i^2|z_i|^2/2)$ with $r_i > 0$ and $S \ge 0$. For a multi-index $\beta \in \mathbb{N}^s$,*

$$|\partial^\beta G(0)| \le S \prod_{\beta_i>0} \beta_i!(er_i^2/\beta_i)^{\beta_i/2}.$$

If complex scalars c_i satisfy $|c_i r_i|^2 = \Omega_i^{-1}$, then

$$|c^\beta\partial^\beta G(0)|^2 \le S^2 \prod_i \Omega_i^{-\beta_i} \prod_{\beta_i>0} e^2\beta_i \exp\bigl(\Phi(\beta_i)\bigr).$$

Proof. Cauchy's formula on the product of circles of radii R_i bounds the derivative by $S\prod_i \beta_i! R_i^{-\beta_i}\exp(r_i^2R_i^2/2)$. For $\beta_i > 0$, minimize at $R_i^2 = \beta_i/r_i^2$. If $\beta_i = 0$, let $R_i \downarrow 0$. This gives the first inequality. The elementary upper Stirling bound $n! \le en^{n+1/2}e^{-n}$ for $n \ge 1$ implies $(n!)^2(e/n)^n \le e^2 n\exp(\Phi(n))$. Square the first inequality and use $|c_i r_i|^{2\beta_i} = \Omega_i^{-\beta_i}$. □

The factors $\Phi(\beta_i)$ and $-\beta_i\log\Omega_i$ are therefore forced by Gaussian growth and differentiation. A scalar Cauchy estimate does not, however, automatically give the norm of a map between Schur modules. One must control the invariant Hermitian forms, the determinant twists, the padding of the weight, and the differential operator that realizes the representation. The block estimate used in the source has the following precise form.

For each embedding τ in a finite index set, let λ_τ, ν_τ be dominant weights of length $a \ge 1$ and λ'_τ a dominant weight of length $b \ge a$. Assume ν_τ is nonnegative, differs from λ_τ by a constant determinant shift, and that a determinant shift of λ'_τ is ν_τ padded by zeros. If $a = b$, require the last entry of ν_τ to be zero. Let $r_\tau, \Omega_\tau > 0$, let $c_\tau \in \mathbb{C}$ satisfy $|c_\tau r_\tau|^2 = \Omega_\tau^{-1}$, and let G be entire on $\prod_\tau \mathrm{Mat}_{b\times a}(\mathbb{C})$ with

$$|G(z)| \le S\exp\left(\tfrac12\sum_\tau r_\tau^2\|z_\tau\|_{\mathrm{HS}}^2\right), \qquad S \ge 0.$$

Write $J = J_{\lambda,\lambda',\nu}(G;c)$ for the block jet map of the source: it applies the prescribed highest-weight construction to the Taylor derivatives of G at zero. Its derivative functional sends α to $(\prod_\tau c_\tau^{|\alpha_\tau|})\partial^\alpha G(0)$, without division by $\alpha!$, and the last entry of ν_τ is its determinant-shift parameter. Both its source and target carry the invariant Hermitian forms fixed in that construction. For this map and these norms, the block estimate is

$$\|J\|^2 \le S^2 \prod_\tau \left[\Omega_\tau^{-|\nu_\tau|}e(2|\nu_\tau| + 2)^{2ab}\exp\left(\sum_i \Phi(\nu_{\tau,i})\right)\right]. \tag{6.11}$$

This is the cited block input in `PaddedLastEntry`. The map is the one specified there by the scaled Taylor functional and its Schur-module construction. The Gaussian hypothesis ties its size to S; the weight conditions alone would give no bound on a jet of an arbitrary entire function.

The row eigenvalue estimate has a complementary gamma-factor expression. For a decreasing sequence $a_0, \dots, a_{n-1}$ in its admissible range, the source uses

$$\mathcal{H}_n(a) = \sum_{i=0}^{n-1} \log\Gamma(a_i + n - i) - \sum_{0\le i<l<n} \log(a_i - a_l + l - i).$$

The second sum is a Weyl denominator. The admissible ordering makes its arguments positive. The estimate used in (6.1) first takes the smallest spectral value over the row representations

at each embedding, and then averages its logarithm over the embeddings. Thus the relevant quantity is $[K:\mathbb{Q}]^{-1}\sum_\sigma \min_\pi \sum_j \log \lambda_{j,\pi,\sigma}$, with the factor $\kappa_{R,\sigma}$ treated separately. An average over π would not control the norm of the inverse row map. The lower bound for this embedding-average involves the displayed gamma expression, with an error polynomial in the ranks and logarithmic in the size and conductor parameters; see `KernelPlaceMeanBound`. More precisely, a stagewise lower bound $L_{6,j}$ holds for $\log \lambda_{j,\pi,\sigma}$ at every selected nonzero row component and every embedding. Summing in j, taking the minimum, and averaging therefore preserves this lower bound. The uniform stagewise assertion is `LocalQuadraticParts.lean`. The correct Stirling expansion is $\log\Gamma(y) = (y-\frac{1}{2})\log y - y + \frac{1}{2}\log(2\pi) + O(y^{-1})$. For $\log\Gamma(y+1)$ the coefficient of $\log y$ is instead $y+\frac{1}{2}$. Confusing these two formulas changes the accumulated logarithmic error.

For integral nonnegative weights there is an elementary alternative that makes the size of the error explicit.

Lemma 6.7. *Let $a_0 \ge \cdots \ge a_{n-1} \ge 0$ be integers, with $n \ge 1$, and put $A = a_0$. Then*

$$\left| \mathcal{H}_n(a) - \sum_i \Phi(a_i) \right| \le n + \frac{n(n+1)}{2}\log(A+n).$$

Proof. The gamma factors can be written exactly as $\log\Gamma(a_i+n-i) = \log(a_i!) + \sum_{h=1}^{n-i-1}\log(a_i+h)$. There are $n(n-1)/2$ terms in this last double sum, and the same number in the Weyl denominator. Each is between zero and $\log(A+n)$: dominance gives $1 \le a_i - a_l + l - i \le A+n$ for $i<l$. The difference of these two nonnegative sums therefore has absolute value at most $n(n-1)\log(A+n)/2$. For an integer $a \ge 1$, monotonicity of $\log t$ gives $\Phi(a)+1 \le \log(a!) \le \Phi(a)+1+\log a$ by comparison with $\int_1^a \log t\,dt$. For $a=0$, both $\log(a!)$ and $\Phi(a)$ vanish. Summing these factorial errors gives at most $n + n\log(A+n)$, which proves the bound. □

Thus, once the weights are identified, the hook expression has leading term $\sum_i \Phi(a_i)$ and an explicit error of order $n^2\log(A+n)$. The corresponding assertion for nonintegral shifts uses the gamma function and its usual uniform remainder estimate. The exact matching of weights remains necessary in either approach: an error of one in a row entry costs only a logarithm, whereas replacing an entire unmatched row by a crude bound can cost its full $a\log a$ contribution.

The stability of the leading term under bounded shifts can be checked without any asymptotic shorthand. If $y \ge 1$ and $|h| \le y/2$, the mean-value theorem gives $|\Phi(y+h)-\Phi(y)| \le |h|\log(2y)$, because $\Phi'(s) = \log s$ and $s \in [y/2, 3y/2]$. For $y \ge 2$ this bound is immediate; for $1 \le y \le 2$ the same estimate follows from $|\log s| \le \log(2y)$ on that interval. Thus, after the exact weight matching has been established, determinant shifts of polynomial size cost only their size times a logarithm. They do not create another term of order $a\log a$.

The crucial calculation is the matching of the kernel rows with the gamma-factor rows. The source records the kernel row as $\nu_{j,\tau,i} = s_{u,\tau,i} + w_{j,\tau} - (N_{j-1}+N_j-1-2i)/2$, where $s_{u,\tau,i}$ is the specified source parameter and $w_{j,\tau}$ is the weight of the step. Its contribution is $\sum_i \Phi(\nu_{j,\tau,i}) - (\sum_i \nu_{j,\tau,i})\log\Omega$. These explicit formulas are what must be compared with the row spectrum; a reference to Stirling's formula alone does not perform the comparison. After cancellation, the retained main term is

$$M_u = \sum_{t\in\mathcal{T}} \left[m\{\Phi(a_t) - a_t\log\Omega\} - a_t\log T_u \right]. \tag{6.12}$$

The $a_t\log T_u$ term comes from evaluation at the final vector. The period term is the same Ω as in (6.9), and the finite contribution $F_{j,w}$ is retained until the lattice comparison.

Proposition 6.8 (`LinkDIV`)**.** *For the compatible families supplied by the construction, the local jet, row-spectrum, period, and lattice-comparison inputs give a constant C and a common threshold such that, above that threshold, every column has an integral map $B_u : N_u' \to R^\vee$. Its squared-norm height is at most h_u when $B_u \ne 0$, and*

$$\mu_2(R^\vee) + h_u - F_0 \le M_u + C|\mathcal{T}|a\mathcal{L}_a + Cm(B_0+N_c)x^{C_{\mathrm{abs}}\varepsilon}, \tag{6.13}$$

where $F_0 = \sum_{j,w} 2\ell_w e_w F_{j,w}$ *and* $\mathcal{L}_a = \log(a \operatorname{rd}_E)$. *The same construction gives the bound for the aggregate column height and the maximal row slope required by the slope argument. It also gives*

$$\left| h_u + \Big(\sum_t a_t\Big) \log T_u - h_{u'} - \Big(\sum_t a_t\Big) \log T_{u'} \right| \le C(|\mathcal{T}| a \mathcal{L}_a + m B_0).$$

This is the global construction interface, with its constants uniform over the stated family. Its full list of compatibility and finite-level clauses is recorded in `GlobalTheoremInputs`. The preceding calculations explain each leading term and each change of normalization, but do not replace the cited local jet or geometric construction by an elementary proof. The distinctive feature of this interface is the simultaneous preservation of the local gain, the period, and the row eigenvalues. Once those data are supplied, the prime-sum and slope arguments can use (6.8) and (6.13) without referring again to the internal theta models.

7. Geometric estimates and spectral concentration

The theta calculation leaves three quantitative tasks. The row lattices need a lower bound for their mean slope, together with an upper bound for the maximal slope. Changes of type and level must preserve those bounds with a controlled loss. Finally, a sufficiently large family of Bessel periods must cover the row space in the quantitative sense used by the frame argument. These tasks interact, but their estimates have different mathematical sources. Arithmetic volume supplies information about sections; integral descent transports that information to the row lattices; and the spectral calculation controls correlations between different columns.

All statements in this section are uniform over the family fixed by the parameter construction. A phrase such as "for sufficiently large x" means that one threshold works for every index in that family, after the displayed constants have been fixed. It does not permit a threshold to depend on a column, on a component of a cover, or on a prime that is still being summed. This order of quantifiers is one of the main requirements of the construction.

Arithmetic Hilbert–Samuel theory and arithmetic bigness are classical inputs [21]. Their role should be stated precisely. An asymptotic count of small sections is not, by itself, a finite-level bound for the mean slope of a particular isotypic lattice. Even for Euclidean lattices, the distinction is elementary: the lattice $\mathbb{Z}^2$ with squared norm $e^{-2N}x^2 + e^{2N}y^2$ has exponentially many vectors in its unit ball, but all such vectors lie on one line when $N > 0$. A small-section count must therefore be combined with rank and spanning information. In the present construction those additional facts are provided by covers, multiplication maps, and integral descent.

Fix a row stage j. Denote its main period contribution by M_j and its error scale by E_0. There are a core lattice L_{core}, a type lattice $L \subset M$, and a self-adjoint idempotent e on their ambient Hermitian space. The retained lattice is $L_g = L \cap \operatorname{im} e$; the discarded lattice is $L_b = L \cap \ker e$. The geometric construction is organized to prove

$$\mu_2(L_{\mathrm{core}}) \ge -M_j - C_2 E_0, \qquad \mu_2(L) \ge \mu_2(L_{\mathrm{core}}) - C_1 E_0. \tag{7.1}$$

The first bound is a lower bound, and its sign is indispensable. An upper bound for $\mu_{2,\max}(M)$ cannot substitute for it. The covolume recursion then gives $\operatorname{lcv}(M_j^{\mathrm{row}})/D \le -\mu_2(L_g)$, with the row rank normalization D specified by the construction.

For the nondefinite stages, the core bound is obtained on a sufficiently deep cover of the unitary moduli space. On that cover one compares the section lattices Λ_- and Λ_+ obtained from two nearby determinant shifts of the automorphic weight, and the lattice Λ of the desired type. The source supplies finitely many integral maps $\phi_h : \Lambda_- \to \Lambda$, with trivial common kernel, such that

$$\sum_h \|\phi_h(s)\|_\sigma^2 \le e^{2b} \|s\|_\sigma^2, \qquad (\#\{h\}) \operatorname{rk} \Lambda \le \operatorname{rk} \Lambda_+, \qquad \operatorname{rk} \Lambda_+ \le \left(1 + \frac{b}{P\kappa}\right) \operatorname{rk} \Lambda_-.$$

Here $P\kappa$ is the scale of the auxiliary weight. A set of integral vectors of norm at most $e^{P\kappa}$ spans Λ over the coefficient field. Descent to the core costs at most $2b$ in squared-norm slope. The combination of the rank ratio and the bound on spanning vectors controls the loss caused by a

small complement of the image of the multiplication maps. Existence of the maps, their integral nature, and their norm estimate are separate requirements; a complex spanning statement does not imply any of them.

The arithmetic estimate on each component of the cover is also more specific than positivity of arithmetic volume. To describe its form, write q for the dimension of the relevant flag variety, T for the auxiliary tensor power, d_Z for the geometric degree on a component Z, and h_Z for its normalized height. With $b_T = \binom{T-1}{q}$, the comparison gives rank shares $b_T d_Z \le r_Z \le b_T d_Z/(1-z)$ with $0 \le z < 1$, and a degree bound of the shape

$$d_Z^{\mathrm{ar}} \ge b_T d_Z \left[T h_Z - 3(q+1)|h_Z| - A\log(2+|h_Z|) - R\right]. \tag{7.2}$$

The constants satisfy $A > 0$, $R \ge 0$, and $A\log(\max(A,1)) + R \le CE_0$. The construction additionally controls the variation of d_Z between components and provides a lower bound for an appropriate average of the heights. These hypotheses allow the component degrees to be summed with their actual ranks. Using a pointwise lower bound for h_Z that is absent from the input would incorrectly discard this part of the argument.

More explicitly, the source introduces auxiliary component heights s_Z and nonnegative numbers c_s, c_c, X with $h_Z + c_s \ge 0$, $(1-\rho)\max(0, s_Z + c_s) \le h_Z + c_s + c_c$, and $\sum_Z s_Z/\#\{Z\} \ge -X$. It then bounds TX, Tc_c, and $c_s[6q + 10 + T(z+2\rho)]$ by CE_0. The degree ratio lies between d_0 and $d_0/(1-\rho)$ for some $d_0 > 0$. These explicit inequalities are the mechanism for transferring a mean height estimate to the rank-weighted expression in (7.2). They also explain why changes of component, degree, and tensor power must be tracked together.

Lemma 7.1. *Let $\mathcal{Z}$ be a finite nonempty set. Suppose $h_Z, s_Z \in \mathbb{R}$, $c_s, c_c, X \ge 0$, and $0 \le \rho, z < 1$ satisfy $h_Z + c_s \ge 0$, $(1-\rho)\max(0, s_Z + c_s) \le h_Z + c_s + c_c$, and $(\#\mathcal{Z})^{-1}\sum_Z s_Z \ge -X$. Suppose also $d_0 \le d_Z \le d_0/(1-\rho)$ for some $d_0 > 0$, and $b_T d_Z \le r_Z \le b_T d_Z/(1-z)$ with $b_T > 0$. Then*

$$\frac{\sum_Z r_Z h_Z}{\sum_Z r_Z} \ge -X - (z+2\rho)c_s - c_c.$$

Proof. Put $\theta = (1-z)(1-\rho)$. The rank bounds imply $r_Z/\sum_W r_W \ge \theta/\#\mathcal{Z}$. Since $h_Z + c_s \ge 0$, the weighted average of these nonnegative numbers is at least θ times their unweighted average. The second hypothesis gives $(\#\mathcal{Z})^{-1}\sum_Z(h_Z + c_s) \ge (1-\rho)(c_s - X) - c_c$, because $\max(0, s_Z + c_s) \ge s_Z + c_s$. Subtracting c_s shows that the desired weighted average is at least $-\theta(1-\rho)X - [1 - \theta(1-\rho)]c_s - \theta c_c$. Finally $0 < \theta \le 1$ and $1 - (1-z)(1-\rho)^2 \le z + 2\rho$ prove the assertion. □

The lemma explains the occurrence of $T(z+2\rho)c_s$ in the height budget. Multiplication by T preserves an $O(E_0)$ loss because that combination, TX, and Tc_c have already been bounded. The absolute height term in (7.2) is controlled using $|h_Z| \le h_Z + 2c_s$, while the logarithmic term is handled by the specified bound on $A\log(\max(A,1)) + R$ and the remaining tensor-power inequalities. The full component argument needs these latter inequalities as well; the lemma isolates the averaging step so that no passage from an unweighted height mean to a rank-weighted slope is left implicit.

The resulting mean slope of the flatly normalized lattice is at least $-CE_0$. Comparing the flat normalization with the shifted automorphic weight contributes the term $-M_j$, where M_j uses the same period Ω as Section 6. The multiplication maps and descent then give the first inequality in (7.1); rationality and comparison of the Petersson and algebraic norms give the second. The final definite stage is treated directly rather than by invoking a nonexistent indefinite symmetric domain.

These are cited geometric construction inputs. The exact component, cover, and descent conditions are recorded in `CovolXOfCovol`. Their assembly is given in `TopCovolG`. The more consolidated entry point `TopCovolSlim2Use` lists its resolution, principalization, Poincaré–Lelong, arithmetic Hilbert–Samuel, arithmetic bigness, Hodge–Tate, polarization, and toroidal-model inputs explicitly. The construction uses those results with its particular bundles and metrics. General resolution of singularities or general arithmetic bigness does not assert the entire package

of uniform inequalities above. In particular, logarithmic boundary metrics require their own comparison with the metrics to which the cited arithmetic intersection results apply.

The passage from the type lattice to the retained lattice can be proved in a short and useful form. It accounts simultaneously for the discarded rank and the denominator of the projector.

Lemma 7.2. *Let $L \subset M$ be nonzero Hermitian $\mathcal{O}_H$-lattices with the metric on L induced from M, and let e be an H-linear idempotent on L_H, orthogonal at every complex embedding. Suppose an integer $g \geq 1$ satisfies $ge(L) \subset L$. Put $L_g = L \cap \operatorname{im} e$ and $L_b = L \cap \ker e$. If $0 \leq \delta \leq 1/12$ and $Q \geq 0$ satisfy*

$$\operatorname{rk} L_b \leq 3\delta \operatorname{rk} L, \qquad \mu_{2,\max}(M) \leq Q, \qquad \mu_2(L) \geq -Q, \qquad 2\log g \leq Q,$$

then $L_g \neq 0$ and $\mu_2(L_g) \geq \mu_2(L) - 12\delta Q$.

Proof. Write $r_g = \operatorname{rk} L_g$, $r_b = \operatorname{rk} L_b$, and $r = r_g + r_b = \operatorname{rk} L$. The rank identity follows from the direct-sum decomposition over H. Since $r_b \leq r/4$, we have $r_g > 0$ and $r_b/r_g \leq 3\delta/(1-3\delta) \leq 4\delta$. Projection to $\ker e$ identifies $L/(L_g \oplus L_b)$ with $(1-e)L/L_b$. Moreover $g(1-e)L \subset L_b$, so the latter quotient is contained in $g^{-1}L_b/L_b$. Its cardinality is at most $g^{[H:\mathbb{Q}]r_b}$. Orthogonality at infinity and the determinant-index formula therefore give

$$r\mu_2(L) \leq r_g\mu_2(L_g) + r_b\mu_2(L_b) + 2r_b \log g.$$

If $r_b = 0$, this already proves the result. Otherwise $\mu_2(L_b) \leq \mu_{2,\max}(M) \leq Q$. Move the term $r_b\mu_2(L)$ to the right, use $-\mu_2(L) \leq Q$ and $2\log g \leq Q$, and divide by r_g. The loss is at most $3Q(r_b/r_g) \leq 12\delta Q$. □

The denominator hypothesis cannot be omitted. An orthogonal projector need not preserve an integral lattice. Nor does a small discarded rank alone control the mean-slope loss if the discarded part has arbitrarily large slope or the original mean slope is arbitrarily negative. The four hypotheses in the lemma address these distinct possibilities. They are the content of the source truncation argument, including its factor 12, in `MajAtAddRight`.

Proposition 7.3 (`lcv_le`)**.** *Assume (7.1), with $C_1, C_2 \geq 0$, and suppose $x \geq 1$, $A, B \geq 0$, and $1 \leq b \leq E_0$ satisfy $\mu_{2,\max}(M) \leq x^A b$ and $|M_j| \leq x^B b$. Assume the projector denominator satisfies $2\log g \leq b$ and the discarded rank is at most $3\delta(x)\operatorname{rk} L$, where $\delta(x) = \exp(-x^{e_z/7})$ with $e_z > 0$. If*

$$12\delta(x)(x^A + x^B + C_1 + C_2 + 1) \leq 1,$$

then any row lattice satisfying $\operatorname{lcv}(M_j^{\mathrm{row}})/D \leq -\mu_2(L_g)$ obeys

$$\frac{\operatorname{lcv}(M_j^{\mathrm{row}})}{D} \leq M_j + (C_1 + C_2 + 1)E_0. \tag{7.3}$$

For fixed A, B, C_1, C_2, e_z, the displayed smallness condition holds for all sufficiently large x.

Proof. Put $Q = (x^A + x^B + C_1 + C_2 + 1)E_0$. The core inequalities imply $\mu_2(L) \geq -M_j - (C_1 + C_2)E_0 \geq -Q$. The other bounds give $\mu_{2,\max}(M) \leq Q$ and $2\log g \leq Q$. The smallness hypothesis implies $\delta(x) \leq 1/12$. Lemma 7.2 consequently gives $\mu_2(L_g) \geq -M_j - (C_1 + C_2)E_0 - 12\delta(x)Q$. Its last term is at most E_0, proving (7.3). Finally, for any fixed real a and $c > 0$, $x^a e^{-x^c} \to 0$: taking logarithms reduces the assertion to $a\log x - x^c \to -\infty$. Apply this with $c = e_z/7$ to the finitely many terms in the smallness condition. □

This proposition is the numerical conclusion of the geometric construction. Its proof makes clear what has to be supplied before the exponentially small bad rank is useful. The core lower bound, the maximal-slope bound, and the integral projector are all part of the input. The source proof of the family version is `CovolumeFromGramAndBounds`. It combines finitely many uniform thresholds by taking their maximum; the exponential absorption is performed after their polynomial exponents have been fixed.

An analytic ingredient in the change of row lattice is interpolation with control of the determinant of the evaluation map. We state its normalization carefully, since both the global weight condition and the Jacobian are essential.

Theorem 7.4 (`ot0`). *Let M be an n-dimensional complex manifold equipped with a locally biholomorphic map $z : M \to \mathbb{C}^n$ and a smooth plurisubharmonic exhaustion; assume $n \geq 1$. Use the volume form pulled back from Euclidean volume in these local coordinates. Let $g = (g_1, \dots, g_n)$ be holomorphic, let φ, φ_2 be smooth plurisubharmonic functions, let $r > 0$, and let $j \geq 2$ be an integer. Suppose*

$$\|g(x)\|^2 e^{-\varphi(x)} \leq r^2 \quad (x \in M), \qquad Z = \{g_1 = \cdots = g_n = 0\} \text{ is finite},$$
$$J_g(p) = \det(dg/dz)(p) \neq 0 \quad (p \in Z).$$

The Jacobian interpolation input supplies, for arbitrary numbers c_p, a holomorphic function F with $F(p) = c_p J_g(p)$ and

$$\int_M |F|^2 e^{-\varphi_2 - (n+j)\varphi}\, dV_z \leq \frac{280\pi^n r^{2n}}{(n-1)!} \sum_{p \in Z} |c_p|^2 e^{-\varphi_2(p) - j\varphi(p)}. \tag{7.4}$$

This is the precise analytic input used here. Its statement is `JacobianInterpolationStatement`; `Ot0` derives it from a fixed-parameter extension estimate, a local shell estimate, disjoint local neighborhoods, and a smooth cutoff. The analytic method belongs to the L^2 extension theory of Ohsawa–Takegoshi [22]. The constant 280 is the explicit constant of this implementation. No assertion of optimality is needed or made.

We give the normalization calculation because it explains the exact form of the theorem. Set $\rho = \|g\|^2 e^{-\varphi}/r^2$, so $0 \leq \rho \leq 1$. Near a simple zero p, changing variables from z to g cancels $|J_g|^2$ against the real Jacobian of the volume form. The polar-coordinate identity

$$\int_{a \leq \|w\|^2 \leq b} \|w\|^{-2n}\, dV_w = \frac{\pi^n}{(n-1)!} \log(b/a)$$

then shows that the shrinking shell $\varepsilon^2/2 \leq \rho \leq \varepsilon^2$ has limiting weighted mass

$$\frac{\pi^n \log 2}{(n-1)!}\, r^{2n} e^{-\varphi_2(p) - j\varphi(p)}$$

when its integrand is $|J_g|^2 e^{-\varphi_2 - (n+j)\varphi} \rho^{-n}$. The factor r^{2n} and the reduction of the exponent from $n + j$ to j on the point values are both consequences of this calculation. The polar identity follows by integrating the sphere area $2\pi^n/(n-1)!$ against dR/R between $\sqrt{a}$ and $\sqrt{b}$.

The source uses positive auxiliary functions $X_\varepsilon = 1 + \log(1 + \varepsilon^2) - \log(\rho + \varepsilon^2)$ and $\eta_\varepsilon = X_\varepsilon + \log X_\varepsilon$. The global bound $\rho \leq 1$ gives $X_\varepsilon \geq 1$ and $\eta_\varepsilon \geq 1$. This is the useful sign condition in the twisted estimate. A choice such as $1 + \log \rho$ would not have that property near Z. With a cutoff having derivative bounded by 3, the fixed-parameter estimate bounds the final norm by $(\sqrt{A_\varepsilon} + \sqrt{B_\varepsilon})^2$. Writing $Q_n = \pi^n/(n-1)!$ and $S = \sum_p |c_p|^2 e^{-\varphi_2(p) - j\varphi(p)}$, the shell and local estimates give, for sufficiently small $\varepsilon > 0$,

$$A_\varepsilon \leq 2e(1 + \varepsilon^2)^n\, 72(1 + 1/200) Q_n r^{2n} (\log 2) S, \qquad B_\varepsilon \leq 4 Q_n r^{2n} (\log 2) \varepsilon^2 S.$$

Choose $\varepsilon < 1/100$ and $(1 + \varepsilon^2)^n < 1 + 1/200$. The inequality $(a + b)^2 \leq (101/100)a^2 + 101 b^2$ reduces the remaining constant to $\frac{101}{100} 2e\, 72 (201/200)^2 \log 2 + \frac{404}{10000} \log 2$, which is less than 280. This explains the numerical constant without identifying it with a sharp extension constant. The existence of the fixed-parameter holomorphic extension remains the cited L^2 input.

For interpolation of actual values $F(p) = a_p$, one must set $c_p = a_p / J_g(p)$. Thus the right side of (7.4) contains $|a_p|^2/|J_g(p)|^2$. Rescaling a local coordinate cannot remove this factor from a global norm estimate unless the metric and measure change are also recorded. Likewise, the theorem is not an unweighted interpolation theorem on an arbitrary Stein manifold: the condition $\|g\|^2 e^{-\varphi} \leq r^2$ is global.

Lemma 7.5. *Let $\mathcal{H}$ be a Hilbert space, $E : \mathcal{H} \to \mathbb{C}^q$ a bounded linear map, and suppose every $c \in \mathbb{C}^q$ has a preimage f satisfying $\|f\|^2 \leq C\|c\|^2$, where $C > 0$. Then $EE^* \geq C^{-1}\,\mathrm{id}$ and $\det(EE^*) \geq C^{-q}$.*

Proof. The map is onto. Its kernel is closed, so every fibre has a unique minimum-norm element in $(\ker E)^\perp$. These elements define a linear right inverse S with $\|S\| \leq \sqrt{C}$. For $c \neq 0$,

$\|c\|^2 = \langle ESc, c\rangle = \langle Sc, E^*c\rangle$ implies $\|E^*c\| \ge C^{-1/2}\|c\|$. This is the operator inequality; multiply its q positive eigenvalue lower bounds to obtain the determinant inequality. □

Apply the lemma to the weighted Bergman space in Theorem 7.4, using the normalized evaluations $F(p)/J_g(p)$ and the point weights in (7.4). These evaluations are continuous by the local mean-value inequality and the smoothness of the positive weight. The theorem gives a lower bound on the Gram determinant of evaluation. To turn that into a slope comparison, the construction must also identify the integral evaluation lattice and its finite index. Analytic interpolation alone does not provide that identification.

The normalized base-to-sharp comparison used in the final divisibility argument is particularly transparent numerically. Put $\mu_0 = \mu_2(R_0^\vee) - \overline{\log \kappa_R}$ and $\mu_\sharp = \mu_2(R^\vee) - \overline{\log \kappa_R}$, with the common row normalization supplied by the compatibility of the two levels. The comparison input gives

$$\mu_\sharp \le \mu_0 + 3\delta_e(\mu_0 + h_0 + 28\lambda) + 28\lambda. \tag{7.5}$$

Its use requires $\delta_e \ge 0$, an upper bound $\mu_0 + h_0 + 28\lambda \le Q$, and $3\delta_e Q \le 1$ with $Q \ge 1$. Under these hypotheses, and $0 \le \lambda \le \lambda_0$, the loss is at most $1 + 28\lambda_0$. If $C_0 \ge 28$ and $m \ge 1$, it is at most $1 + C_0 m\lambda_0$. This is the explicit absorption performed in `ConstantGeTwentyEight`. The number 28 summarizes several comparison estimates already proved in that construction. It does not follow merely from the inequality $\log 280 < 28$.

We now turn to the spectral part. Let H_u be the normalized column Hilbert space of dimension d_u, let E_0 be the row Hilbert space of dimension D, and let $A_u : H_u \to E_0$ be the normalized adjoint period. The unitary and scalar identifications of Proposition 6.2 are incorporated in this normalization. Write $K_{uv} = A_u^* A_v$. The off-diagonal assertion concerns the Hilbert–Schmidt norm of this entire block, not individual matrix entries in an arbitrary basis.

Proposition 7.6 (`LinkS`). *For the spectral construction, put $\Gamma = \exp(C_\Gamma m^5 f L)$ and let $D_* > 0$ be its reference row dimension, and let $C_0 \ge 1$. Assume the row and column spaces have positive dimensions. For any admissible family whose column dimensions satisfy $C_0^{-1} \le d_u\Gamma^2/D_* \le C_0$, the source supplies numbers $\eta_0, \eta_1, e_0 \ge 0$ such that*

$$\begin{aligned} \|K_{uv}\|_{\mathrm{HS}}^2 &\le (1+\eta_0) d_u d_v / D_* \qquad (u \ne v), \\ \|K_{uu} - I_{H_u}\|_{\mathrm{HS}}^2 &\le e_0 d_u, \qquad\qquad D \le (1+\eta_1) D_*, \\ \eta_0, \eta_1, \sqrt{e_0} &\le \exp(-x^{\varepsilon/8}). \end{aligned} \tag{7.6}$$

The constants and threshold are uniform in the admissible family.

The nonnegative choice of η_0, η_1 is harmless: if a source choice is negative, replacing it by zero weakens the corresponding upper bound and preserves the displayed estimate. The construction input is stated in `GlobalTheoremInputs`; the spectral-to-family passage is `LinkSFinal`. Its hypotheses include the finite-place diagonal and off-diagonal estimates, absolute convergence of the kernel expansions, comparison with the selected row representations, and counting estimates. These conditions are what permit the spectral sums and their normalizations to be compared uniformly.

The rotated Bessel estimate entering this passage is another construction-specific result. Its uniform statement chooses constants after fixing the external weight and height bounds, and then applies to every sufficiently separated highest weight in the permitted class. The source derives it from branching support, a cohomological-induction gap, and rank-one torus estimates; see `ThmBFinal`. Classical branching and cohomological induction provide the representation-theoretic tools. The uniform bound for the particular rotated orbital integrals, together with the counting needed to use it, is the additional quantitative work. A qualitative decay theorem for one fixed representation would not yield (7.6) as both the ranks and weights grow.

The passage from these block estimates to an approximate frame is finite-dimensional linear algebra. Giving it explicitly clarifies how many columns are required.

Lemma 7.7. *Assume the first three inequalities in (7.6). For nonnegative numbers p_u with $\sum_u p_u = 1$, put $T_u = (Dp_u/d_u)^{1/2}A_u$ and $Q = \sum_u T_u T_u^*$. Then $|\operatorname{Tr} Q - D| \le D\sqrt{e_0}$, and*

$$\frac{\|Q - I_{E_0}\|_{\mathrm{HS}}^2}{D} \le (1+\eta_0)(1+\eta_1) - 1 + 2\sqrt{e_0} + D(1+\sqrt{e_0})^2 \sum_u \frac{p_u^2}{d_u}. \tag{7.7}$$

For q equally weighted columns with $d_u \ge D_*/(C_0\Gamma^2)$, the last term is at most $C_0(1+\eta_1)(1+\sqrt{e_0})^2\Gamma^2/q$.

Proof. Set $E_u = K_{uu} - I_{H_u}$. Cauchy–Schwarz for the trace gives $|\operatorname{Tr} E_u| \le \sqrt{d_u}\|E_u\|_{\mathrm{HS}} \le d_u\sqrt{e_0}$. Since $\operatorname{Tr} Q = D\sum_u p_u \operatorname{Tr} K_{uu}/d_u$, this proves the trace bound. Also $\|K_{uu}\|_{\mathrm{HS}}^2 = d_u + 2\operatorname{Tr} E_u + \|E_u\|_{\mathrm{HS}}^2 \le d_u(1+\sqrt{e_0})^2$. Expanding the square of Q and using cyclicity of the trace gives

$$\operatorname{Tr}(Q^2) = D^2 \sum_{u,v} \frac{p_u p_v}{d_u d_v} \|A_u^* A_v\|_{\mathrm{HS}}^2 \le D^2(1+\sqrt{e_0})^2 \sum_u \frac{p_u^2}{d_u} + \frac{D^2(1+\eta_0)}{D_*}.$$

Here the off-diagonal weights sum to at most one. Now use $\|Q - I\|_{\mathrm{HS}}^2 = \operatorname{Tr}(Q^2) - 2\operatorname{Tr} Q + D$ and $D/D_* \le 1 + \eta_1$. For equal weights, $\sum_u p_u^2/d_u \le C_0\Gamma^2/(qD_*)$, giving the final assertion. □

To match Proposition 5.2, use the rank weights $p_u = d_u/R$, where $R = \sum_u d_u$, and put $Q_0 = \sum_u A_u A_u^*$. Then the operator of the lemma is $Q = (D/R)Q_0$. The same trace calculation, now without varying column rescalings, gives

$$\operatorname{Tr} Q_0 \ge (1 - \sqrt{e_0})R, \qquad \operatorname{Tr}(Q_0^2) \le (1+\sqrt{e_0})^2 R + (1+\eta_0)R^2/D_*.$$

For $e_0 \le 1/4$, the hypotheses of Proposition 5.2 therefore hold with $B_u = A_u$, $c_u = 1$, $e_F = e_0$, and

$$1 + \eta_F = \frac{(D/R)(1+\sqrt{e_0})^2 + (1+\eta_0)(1+\eta_1)}{(1-\sqrt{e_0})^2}.$$

This choice has $\eta_F \ge 0$. If there are q columns, the dimension lower bound gives $D/R \le C_0(1+\eta_1)\Gamma^2/q$. Thus the rank-weighted version has the same sufficient cardinality estimate as the equal-weight version. This is the explicit connection between the block estimates and the slope-transfer theorem; it keeps the source-rank weights required in that theorem.

For that application, the full spectral row space is the lifted target $\widetilde{\mathcal{H}}$, of dimension D_{all}, whereas the arithmetic target is the retained row lattice $F = R^\vee$, of rank D_{rows}. After the Bessel identifications and the choice of auxiliary Hilbert coordinates, the lifted maps have a common projection $\pi : \widetilde{\mathcal{H}} \to F \otimes_\sigma \mathbb{C}$ satisfying $\pi\widetilde{B}_u = B_u \otimes_\sigma 1$. Concretely, the coordinates split as $(F \otimes_\sigma \mathbb{C}) \oplus \mathbb{C}^{D_{\mathrm{all}} - D_{\mathrm{rows}}}$. The row data satisfy $D_{\mathrm{rows}} = (1 - \delta_{\mathrm{row}})D_{\mathrm{all}}$, and `LinkIRR` gives $0 \le \delta_{\mathrm{row}} \le \exp(-x^{e_z/7}) < 1$, where $e_z > 0$ is fixed for the family. Thus (5.7) holds with $\delta_F = \delta_{\mathrm{row}}$. These coordinate choices supply the frame data; the arithmetic map heights retain their original metrics at all embeddings. The rank identity is part of `DataRows.lean`; the compatible projected maps are constructed in `FamRec.lean` and `FamRecP.lean`.

In particular, if the desired normalized error is $O(x^{-C_F})$, a sufficient cardinality is $q \ge C_0\Gamma^2 x^{C_F}$, up to a fixed constant. The exponential terms in (7.6) are then smaller than every fixed inverse power of x. A choice only of size $\Gamma^2 x^{c_H}$ gives a diagonal contribution of order x^{-c_H} by this calculation; it does not give an arbitrary stronger power. The required error rate must therefore be fixed before the final cardinality is chosen. This observation is particularly relevant when the frame error is multiplied by a polynomial bound for a Harder–Narasimhan slope.

The form of (7.7) also explains why pointwise off-diagonal decay cannot simply be summed without checking the normalization. There are $q(q-1)$ off-diagonal blocks and q diagonal blocks. The factors Dp_u/d_u determine their total contribution. The construction supplies estimates on these normalized blocks; an informal assertion that large auxiliary primes force individual periods to be small would not give the needed bound for $Q - I$.

The remaining integral comparisons require attention at small primes. The local jet input of Section 6 assumes $p \ge 5$. Its proof should not be reused at 2 or 3 by suppressing that hypothesis. At those primes, integral p-adic Hodge theory supplies a different comparison. Prismatic cohomology provides the general integral framework [5]; the classification of p-divisible

groups by admissible prismatic Dieudonné crystals is due to Anschütz–Le Bras [2]. These statements concern integral structures and therefore contain information lost after tensoring with $\mathbb{Q}_p$.

The source passage `PrismaticDieudonneHodge` uses a primitive compatible system of p-power roots of unity, integral period matrices, a window lift to Fontaine's integral period ring, and the connected/dual-connected comparison. It then obtains the required Hodge–Tate exactness. This explains the role of the prismatic input, but does not yet give a numerical bound for every automorphic congruence lattice. Those bounds also use the local representation data, the specified twisting and killing operators, and the cohomological defect.

For example, the small-prime interface in `BoundAtTwoThreeFamily` bounds a local multiplicity or congruence quantity $v_{i,\pi}$ by

$$v_{i,\pi} \le s_{i,\pi} + \mathcal{B}_{23} + \mathcal{V}(p) + h_E(p) + \epsilon(p), \qquad \mathcal{B}_{23} = Cfmb \log x + Cx^C L_2^2. \tag{7.8}$$

Here $s_{i,\pi}$ is the comparison quantity supplied by the local representation family, $\mathcal{V}(p)$ is the sum of the relevant cohomological lengths at p, $h_E(p)$ is the local field contribution, and $\epsilon(p)$ is the prescribed local error. The scale L_2 is the large secondary parameter of that local comparison, not the archimedean logarithmic scale L. The assertion applies at $p = 2$ and at $p = 3$ when it is not already included in the designated small-prime set. Its hypotheses include finiteness of the cohomological lengths and the compatible representation and killing data. In the source the finiteness predicate covers every prime despite its historical name containing the word "Odd".

The dependence in (7.8) is important. It is a representation-theoretic local estimate with cohomological terms, not an automatic inequality between arbitrary de Rham and étale lattices. Membership of a number field in a 2-power cyclotomic tower also does not, by itself, construct a Wach lattice for an arbitrary representation at 2 or at 3. The argument uses the integral comparison structures actually supplied by the source and their hypotheses at the relevant prime.

After local multiplicities are summed with their prescribed weights, the global quantity to control at stage j is $\eta_j = \sum_{w \in S,\ \mu_w \neq \mu_0} \ell_w e_w \eta_{j,w}$. There is no factor of two in this definition. The construction bounds it by

$$\eta_j \le D_j^0 \left[C\{fmbL^2 + (\log x)^2 N_c\} + \sum_p (\log p) \sum_{0<e<t} \operatorname{length} \operatorname{Sha}_{p,e} \right], \tag{7.9}$$

provided the nonnegative prime series converges. The source also identifies D_j^0 with the common base rank D_0 in the required family. The cohomological section establishes the remaining global length estimate. Combining it with the parameter bounds gives $\eta_j \le CD_0 x^{\varepsilon}(B_0 + N_c)$. This is how the small primes enter the common error budget. One must not insert the factor 2 from squared-norm slopes into (7.9); that factor enters only at the later arithmetic-degree conversion.

Finally, the selection of constants can be separated cleanly from the estimates they control. First fix the exponents and separation requirements of the arithmetic construction. Next obtain the polynomial bounds for column heights, ideal shifts, and maximal row slopes from the geometric and local comparison statements. Choose an exponent c_H larger than those bounds. Only then choose the frame accuracy needed by the final slope inequality, and choose the number of columns and the spectral thresholds to deliver that accuracy. The existence statement for the admissible design must allow these enlarged fixed constants. Once they have been chosen, all lower thresholds are replaced by a common maximum.

This order is more than a convenient presentation. If a frame theorem were allowed to increase the slope exponent after the frame accuracy had been fixed, the final absorption would not follow. The source uses family statements whose dependence on external constants is recorded explicitly, so the finite sequence of choices can be made in this order. The three estimates proved in this section make the remaining algebra visible: exponential bad-rank losses are absorbed by Proposition 7.3, the base-to-sharp loss is absorbed using (7.5), and spectral block estimates are converted to the required approximate frame by Lemma 7.7. Together with the cited construction inputs, they supply the geometric and spectral inequalities used in the final arithmetic comparison.

8. Cohomological bounds on a common window

The integral comparison in Section 10 leaves a specific arithmetic quantity to control: the total length of certain classes locally trivial at the places above the residue characteristic, summed over the residue characteristic and over a long initial interval of twists. This is substantially stronger than finiteness for each individual Selmer group. The construction requires a bound uniform in the auxiliary primes, in the height of a dyadic layer, and throughout one interval $[W, W^2]$. The distinction between these requirements explains both the use of several automorphic groups and the final selection of a window.

We first fix the cohomological conventions. For a number field E and a rational prime ℓ, let S_ℓ be the finite places of E above ℓ, and let G_{E,S_ℓ} be the Galois group of the maximal extension unramified outside S_ℓ and the archimedean places. Cohomology is continuous, with the usual discrete topology on $\mathbb{Q}_\ell/\mathbb{Z}_\ell$ and the continuous inverse-limit interpretation for $\mathbb{Z}_\ell$ coefficients. Write χ_ℓ for the cyclotomic character. For an integer $e \geq 1$, put

$$\begin{aligned} \mathrm{Sha}_{\ell,e}(E) &= \ker\left(H^1(G_{E,S_\ell}, \mathbb{Q}_\ell/\mathbb{Z}_\ell(-e)) \longrightarrow \prod_{v\mid\ell} H^1(E_v, \mathbb{Q}_\ell/\mathbb{Z}_\ell(-e)) \right), \\ v_{\ell,e}(E) &= \mathrm{length}_{\mathbb{Z}_\ell} \mathrm{Sha}_{\ell,e}(E). \end{aligned} \tag{8.1}$$

The twist is negative: g acts on the coefficient module by $\chi_\ell(g)^{-e}$. The localization product in (8.1) contains precisely the finite places above ℓ. These conventions agree with `Statement.lean`. The length initially belongs to $\mathbb{N} \cup \{\infty\}$; finiteness is part of every bound below, rather than an implicit convention for interpreting an infinite length numerically.

For an integer $t \geq 2$ and a nonnegative real number B, write $\mathcal{V}(E, t; B)$ if all $v_{\ell,e}(E)$ with $1 \leq e < t$ are finite and

$$\sum_{\ell\in S} \log \ell \sum_{1\leq e<t} v_{\ell,e}(E) \leq B \qquad \text{for every finite set } S \text{ of rational primes.} \tag{8.2}$$

This formulation avoids an unnecessary convergence convention. Since the summands are nonnegative, (8.2) also says that their sum, defined as the supremum of finite partial sums, is at most B. It is unchanged by replacing an algebraic closure, local decomposition groups, or the compatible cyclotomic data by isomorphic choices. The source explicitly quantifies over these choices in `ExistsPageN.lean`.

Two elementary facts will be used repeatedly. If $t' \leq t$ and $B \leq B'$, then $\mathcal{V}(E, t; B)$ implies $\mathcal{V}(E, t'; B')$, because each new sum is a subsum of an old one. Also, if the primes have been partitioned into finitely many classes and the contribution from class i is at most B_i for every finite partial sum, the full contribution is at most $\sum_i B_i$. Neither argument permits replacing finiteness by a numerical coercion of an infinite length.

For a CM field, complex conjugation acts on (8.1). At an odd prime the idempotents $(1+c)/2$ and $(1-c)/2$ split its plus and minus parts. Their lengths add to $v_{\ell,e}(E)$. At $\ell = 2$ these idempotents are not integral, and the dyadic contribution is treated separately. Poitou–Tate duality relates the length in (8.1) to the kernel of localization on $H^2(G_{E,S_\ell}, \mathbb{Z}_\ell(e+1))$. Passing from that kernel to the entire H^2 group introduces the global and local H^0 terms in the duality sequence. In the source the comparison is recorded as a length identity with these terms, not as an unconditional identification with the whole H^2 group. Its real-place qualification is automatic for CM fields, including at $\ell = 2$.

The link with the earlier lattice estimate has the following meaning. An integral Hecke algebra need not split integrally into its characteristic-zero eigenspaces. The index of its integral lattice in the sum of the projected lattices measures this failure. In a generalized matrix algebra the off-diagonal blocks encode extensions between the constituents of the associated Galois representation. The theta tower contributes character lines whose ratios are powers of χ_ℓ; extensions between two such lines therefore involve the negative twists in (8.1). Local compatibility places the relevant classes in its localization kernel. Controlling the other blocks and the integral projectors gives the lattice-index input used in Section 10. The congruence-to-extension principle and generalized matrix algebras are classical [24, 4]; the simultaneous quantitative comparison for

these many character lines is a separate construction-specific assertion. The argument below estimates the resulting cohomology lengths, without treating that comparison as a consequence of the word "congruence."

The arithmetic fields are those already used in the construction. Fix a primitive Dirichlet character ψ, let F_ψ be its fixed field, and let $o = [F_\psi : \mathbb{Q}]$. For admissible auxiliary primes q, d and an integer $h \geq 2$, write

$$E_h = F_\psi\, \mathbb{Q}(\zeta_{2^{h+2}})^+\, \mathbb{Q}(\sqrt{-q})\, \mathbb{Q}(\sqrt{-d}), \qquad N_h = \operatorname{lcm}(N_\psi, 2^{h+2})qd. \tag{8.3}$$

Here N_ψ is the conductor of ψ. Equivalently, E_h is the subfield of $\mathbb{Q}(\zeta_{N_h})$ fixed by the residues a satisfying $a \equiv \pm 1 \pmod{2^{h+2}}$, $\psi(a) = 1$, and both quadratic conditions $(a/q) = (a/d) = 1$. This equivalence specifies the field even when the factors in the displayed compositum intersect. In particular, its degree must be computed with those intersections, rather than by multiplying the four displayed degrees. Let $n_E = [E_h : \mathbb{Q}]$ and F be its conductor; then $F \mid N_h$.

The prime sum in the minus eigenspace illustrates a useful arithmetic device. It is an infinite-prime estimate obtained from finitely many nonzero algebraic integers, not by summing a pointwise bound that is uniform in ℓ.

Lemma 8.1. *Let M be a nonzero integer and let $a_\ell \geq 0$ satisfy $a_\ell \leq \operatorname{ord}_\ell(M)$ for every prime ℓ. Then every finite partial sum of $\sum_\ell a_\ell \log \ell$ is at most $\log|M|$. More generally, let a finite multiset $\mathcal{A}$ of nonzero algebraic integers be stable under the absolute Galois group, with product $M \in \mathbb{Q}$. If $a_\ell \log \ell \leq \sum_{\alpha\in\mathcal{A}} -\log|\alpha|_\ell$, where the ℓ-adic absolute value is normalized by $|\ell|_\ell = \ell^{-1}$, the same conclusion holds. If two such integers M_1, M_2 provide the bound respectively away from distinct primes r_1, r_2, then $\log|M_1 M_2|$ bounds the total prime sum.*

Proof. Unique factorization gives $\log|M| = \sum_\ell \operatorname{ord}_\ell(M) \log \ell$. All its summands are nonnegative. In the second formulation Galois invariance makes the product rational, and integrality makes it a nonzero integer; multiplicativity of the absolute value gives the first formulation. For the last assertion use M_1 at every prime other than r_1 and M_2 at r_1. The inequality for M_2 is available there because $r_1 \neq r_2$. □

For clarity, an upper bound for the absolute values of the factors is not enough on its own. One needs their nonvanishing, integrality after specified denominators are cleared, and a Galois-stable collection of factors. The normalized valuation must agree with the normalized module length. If $\mathcal{O}$ is a finite extension of $\mathbb{Z}_\ell$ of ramification index $e_\mathcal{O}$, the source divides $\mathcal{O}$-length by $e_\mathcal{O}$. This normalization is forced by the identity $\operatorname{length}_\mathcal{O}(M \otimes_{\mathbb{Z}_\ell} \mathcal{O}) = e_\mathcal{O} \operatorname{length}_{\mathbb{Z}_\ell} M$ for finite $\mathbb{Z}_\ell$-modules M.

Proposition 8.2 (Minus estimate; `thmBcFree_of`)**.** *The cyclotomic-character and integral annihilator inputs of the source imply the existence of an absolute constant $C \geq 0$ such that, for every field (8.3), every $h \geq 2$, and every integer $t \geq 2$, the odd-prime minus contribution is finite and bounded by*

$$3n_E t^2 \log(tF) + C \sum_{\substack{\ell \mid o \\ \ell \neq 2}} \ell n_E^2 t^2 \log(t\ell F)^2. \tag{8.4}$$

Consequently it is at most $C_\psi n_E^2 t^2 \log(tF)^2$ for a constant depending only on the fixed character.

Reduction to the character and annihilator inputs. When $\ell \nmid 2n_E$, integral character idempotents are available. The required minus characters have parity $\vartheta(-1) = (-1)^{e+1}$; thus the associated generalized Bernoulli values have the parity for a nonzero special value. The character input bounds their normalized lengths by valuations of the denominator-cleared special values. The Galois-stable product estimate in Lemma 8.1, applied for each $1 \leq e < t$, bounds the prime sum. The source's explicit denominator and factorial estimate yields the first term in (8.4) after summing in e.

If an odd prime divides n_E, it divides the degree o of the fixed field F_ψ: the remaining factors in (8.3) have degrees that are powers of two. There are therefore only finitely many exceptional odd primes, all fixed before q, d, h vary. At these primes the integral group-ring annihilator

replaces division by n_E. The elementary length estimate used here is that a finite ℓ-primary module killed by ℓ^r and generated by at most g elements has length at most rg. Indeed, it is a quotient of $(\mathbb{Z}_\ell/\ell^r)^g$. The annihilator and generator estimates in the source give the second term of (8.4). This establishes the displayed bound from precisely the two arithmetic inputs stated in the proposition. Finally, absorb the fixed finite set of primes dividing o and the inequalities $n_E \geq 1$, $tF \geq 2$ into C_ψ. □

The source of the exact bound is `ThmBcOfPsiBound.lean`, with its numerical form defined in `StmtFields.lean`. Cyclotomic Euler systems and Stickelberger-type annihilation belong to the established theory [18, 8]. Those references explain the provenance of the input; they do not, without its local conditions and denominator calculation, establish (8.4). The contribution relevant here is the uniform arrangement of this theory for the changing fields (8.3) and the weighted sum over all odd primes.

The plus contribution is organized by characters, with different arguments in different residual configurations. To state the division accurately, fix a coefficient DVR containing the needed roots of unity. For a character ϑ factoring through $\mathrm{Gal}(E_h/\mathbb{Q})$, first remove its ℓ-power-order part. From the remaining character remove its component of ℓ-power conductor, and denote the result by ϑ_ℓ°. Thus ϑ_ℓ° has both order and conductor prime to ℓ. It is its *order*, not merely its conductor, whose parity distinguishes two of the cases. The plus characters satisfy $\vartheta(-1) = (-1)^e$.

There is a fixed finite set $\mathcal{B}_\psi$ of odd primes. More explicitly, choose the source's auxiliary prime $p_K \equiv 3 \pmod 4$, with $p_K \geq 7$ and $(p_K, N_\psi) = 1$, put $K_* = \mathbb{Q}(\sqrt{-p_K})$, and choose the auxiliary prime p_1 with its prescribed congruence modulo the odd part of o. Then $\mathcal{B}_\psi = \{3\} \cup \{\ell \neq 2 : \ell \mid oN_\psi p_K h(K_*) p_1\}$. The choice depends only on ψ. The host-size hypothesis includes $q, d > 2t+1$ and q, d larger than these fixed constants. With $a_{\ell,e,\vartheta}$ denoting the normalized length of the source's character Selmer module, the relevant inputs are as follows.

Case	Conditions	Output used
I	$\ell \notin \mathcal{B}_\psi$ and $\ell - 1 \nmid 2e$	One bound for the prime and character sum.
Z	$\ell \notin \mathcal{B}_\psi$, $\ell - 1 \mid 2e$, and $\vartheta = 1$	Vanishing from the separate trivial-character input.
II	$e \geq 2$, $\mathrm{ord}(\vartheta_\ell^\circ)$ even, and either $\ell \in \mathcal{B}_\psi$ or $\ell - 1 \mid 2e$	A polynomial bound in $\ell N_h e$.
III	$\ell \notin \mathcal{B}_\psi$, $\ell - 1 \mid 2e$, and $\vartheta \neq 1$ has odd order	A polynomial bound in $\ell N_h e$.
IV	$\ell \in \mathcal{B}_\psi$, $e \geq 2$, and $\mathrm{ord}(\vartheta_\ell^\circ)$ odd	Reduction to a fixed cyclotomic layer.
V	$\ell \in \mathcal{B}_\psi$ and $e = 1$	Separate estimates above and below a dyadic-conductor threshold.

The host hypotheses make this division exhaustive. Outside $\mathcal{B}_\psi$, the character order is prime to ℓ. In the resonant case $\ell - 1 \mid 2e$, one has $\ell \leq 2e + 1 < q, d$ and $\ell \nmid N_\psi$, so the character also has conductor prime to ℓ and $\vartheta_\ell^\circ = \vartheta$. Resonance with $e = 1$ would force $\ell = 3 \in \mathcal{B}_\psi$. The remaining outside-$\mathcal{B}_\psi$ cases are therefore exactly Z, II, and III, with $e \geq 2$ where required. Inside $\mathcal{B}_\psi$, first separate $e = 1$ from $e \geq 2$, and then separate odd and even order. Overlapping alternatives may be assigned in any fixed order.

The precise threshold in case V is $2^{c_2(\vartheta)} \geq 16o(\ell N_\psi)^2 4^{n_0(\ell,\vartheta)}$, where $2^{c_2(\vartheta)}$ is the conductor of the 2-primary conductor component and n_0 is the 2-adic valuation of $\ell^{f_0} - 1$ for the source's residue-degree parameter f_0. Its high-conductor and low-conductor branches must both be supplied. The definitions and the exhaustive division appear in `Counts.lean` and `Cases.lean`.

Case I is where a uniform bound in a single residue characteristic would be insufficient. The rank-three argument first supplies a norm estimate for all characters with a fixed kernel. The following count explains why regrouping by kernels controls the number of factors.

Lemma 8.3. *If A is a finite abelian group, then $\sum_{H\le A,\ A/H\ \mathrm{cyclic}} \varphi([A:H]) = |A|$. In particular, the same sum over any subcollection of kernels is at most $|A|$.*

Proof. Every character $A \to \mathbb{C}^\times$ has a cyclic image. Its kernel H determines a faithful character of A/H. A cyclic group of order m has exactly $\varphi(m)$ faithful characters, because the image of a generator must be a primitive mth root of unity. Partitioning the $|A|$ characters of A by their kernels proves the identity. The final assertion uses nonnegativity. □

For the rank-three norm comparison the character configuration is $\Psi \oplus 1 \oplus \Psi^{-1}$, with $\Psi = \vartheta\chi_\ell^e$. The residue-character and imaginary-quadratic hypotheses in the source ensure that it lies in the particular deformation problem to which the comparison applies. Besides nonresonance, these include prime-to-ℓ character order, the required parity, an at-most-quadratic ℓ-power-conductor component, coprimality to the discriminant and class number of K_*, and the stated disjointness of the residual character fields from K_*. These conditions should not be replaced by nonresonance alone.

The quantitative form before regrouping is particularly informative. For a fixed character kernel H, conductor modulus N, and two distinct auxiliary primes r_1, r_2 coprime to $30N\,|\,\mathrm{disc}\,K_*|$, the input bounds the relevant finite prime sum by

$$4c_5\varphi([A:H])|\,\mathrm{disc}\,K_*|^5 N^{18} e^7(\log r_1 + \log r_2). \tag{8.5}$$

The precise character and local hypotheses are part of that input. Lemma 8.3 absorbs the orbit factors, and Lemma 8.1 explains the use of two different auxiliary primes. After increasing a constant depending only on ψ, the output is at most $(N_h t)^{C_\psi}$, uniformly over every finite set of case-I primes and every $1 \le e < t$. This is the assertion packaged by `CountI.lean`.

The mechanism producing (8.5) contains deep arithmetic input. The integral congruence algebra compares the reducible character parameter with finitely many non-Eisenstein characteristic-zero points. At a point with a $2+1$ decomposition, the two-dimensional constituent is related to a modular form and requires the particular two-twist Selmer-vanishing statement used by the source. Kato's Euler system is the classical ingredient [17]; modularity, the local conditions, and the specialization to the required twists remain additional steps. The source explicitly keeps these hypotheses in `KatoTwoW.lean`. At an irreducible three-dimensional point, the ordinary unitary deformation and automorphy argument provides a different input. Its construction from Hecke algebras, density, and arithmetic specializations is recorded in `THidaU.lean`. Hida theory and the Taylor–Wiles method supply the established background [14, 29]; they do not by their names alone imply the required tangent-space or length estimate.

One part of the nonvanishing comparison is elementary and worth isolating. If $r \ge 31$, $e \ge 1$, and $|z| = 1$, then $|zr^e + 1 + z^{-1}r^{-e}| \ge r^e - 1 - r^{-e} > 3r^{e-1/2}$. The final inequality follows on dividing by $r^{e-1/2}$: the result is at least $\sqrt{r} - r^{-1/2} - r^{-3/2} > 3$. Thus an automorphic trace bounded by $3r^{e-1/2}$ cannot equal this character trace. This proves nonvanishing once the trace estimate, normalization, and integrality have been established. The proof must still construct the finite collection of traces and their integral norm product; the numerical separation does not replace that construction.

The change of group in cases II and III has a similarly concrete reason. The four-character parameter $\alpha\Psi \oplus \Psi \oplus \alpha \oplus 1$ admits an alternating pairing with multiplier $\alpha\Psi$: pair the first coordinate with the fourth and the second with the third. For $\alpha = \chi_\ell$ its weights are $e+1, e, 1, 0$, which are distinct exactly when $e \ge 2$. This elementary observation explains the regular symplectic configuration. The deformation problem, its automorphic realization, and its polynomial length estimate are the substantive inputs. At $e = 1$ the repeated middle weight prevents invoking that regular-weight argument; the source's case V uses a separate unitary calculation. No single theorem about GSp_4 accounts for all these cases.

Outside $\mathcal{B}_\psi$, the remaining resonant primes satisfy $\ell - 1 \mid 2e$ and hence $\ell \le 2e + 1 < 2t + 1$. There are at most $2t$ possible primes, at most t twists, and at most $[E_h : \mathbb{Q}] \le N_h$ characters in

each relevant character set. Therefore a bound $(\ell N_h e)^C$ for each normalized contribution gives a total bounded by another fixed power of $N_h t$. This finite counting argument is valid only after case I has removed the unrestricted prime sum. At the fixed primes in $\mathcal{B}_\psi$, cases II and V have the same polynomial character count; case IV is the part passed to the tower argument.

The tower argument needs a deliberately crude bound at a fixed layer. Its purpose is to control a few unusually close specializations, not to optimize any arithmetic exponent.

Proposition 8.4 (Uniform fixed-layer input; `LemmaK`)**.** *The source's class-counting, odd-homology rank, and integral torsion inputs imply that there is an absolute $C_K \geq 1$ with the following property. Let L be a totally complex number field containing μ_ℓ when ℓ is odd, and containing μ_4 when $\ell = 2$. For every integer $e \geq 1$, the group $H^2(G_{L,S_\ell}, \mathbb{Z}_\ell(e+1))$ is finite and*

$$\log|H^2(G_{L,S_\ell}, \mathbb{Z}_\ell(e+1))| \leq \exp((2e[L:\mathbb{Q}])^{C_K} \log(2|\operatorname{disc} L|)). \tag{8.6}$$

The numerical interface is `LemmaK.lean`. The arithmetic-window proof obtains it through a count of equivariant ideal-class characters, followed by Poitou–Tate duality. The following account specifies the substantive inputs in that route. Their construction is retained as a deep source input; the numerical passage to the displayed bound is explained explicitly.

Fix a rational prime ℓ, an integer $e \geq 1$, and a totally complex number field L containing μ_ℓ if ℓ is odd, and μ_4 if $\ell = 2$. Write $n = [L:\mathbb{Q}]$, $r_2 = r_2(L)$, and $T_{N,j}(L) = |H_j(\mathrm{GL}_N(\mathcal{O}_L), \mathbb{Z})_{\mathrm{tors}}|$. The notation $v_{\ell,e}(L)$ is that of (8.1), with coefficient module $\mathbb{Q}_\ell/\mathbb{Z}_\ell(-e)$. One ingredient is a rank bound for a particular part of odd integral homology. Let Ψ_b be the map induced by sending a matrix to a block diagonal matrix with b identical blocks, where $b \geq 1$, and let st denote the standard stabilization. Starting with $z_0 = x \in H_{k+1}(\mathrm{GL}_N(\mathcal{O}_L), \mathbb{Z})$, form $z_{j+1} = \Psi_b z_j - b^{j+2}\mathrm{st}(z_j)$; both terms lie in $H_{k+1}(\mathrm{GL}_{b^{j+1}N}(\mathcal{O}_L), \mathbb{Z})$. The source proves that, when k is even, $k+1 < r_2$, and $N \geq 2(k+1)+4$, the subgroup spanned by all z_k has rank at most r_2. This is the assertion `BOdd`, whose proof is supplied in `BOdd.lean` through the odd-homology construction in `FinalOdd.lean`. It is a rank assertion for these specified operators; a bound on the entire odd homology group is not being asserted. In the application take $b = \ell$, $k = 2e$, and $N = 4e + 6 = 2(k+1) + 4$.

The elementary counting principle behind its use is useful independently. Suppose an abelian group H has finite torsion subgroup of order T, a subgroup J has rank at most r, and a homomorphism $f : H \to B$ takes values in a group killed by $q \geq 1$. Every finite subset of $f(J)$ has at most $q^r T$ elements. Indeed, choose finitely many preimages and let J_0 be the subgroup they generate. The structure theorem gives $J_0 \simeq \mathbb{Z}^{r_0} \oplus T_0$, with $r_0 \leq r$ and $|T_0| \leq T$. The restriction of f factors through J_0/qJ_0, which has at most $q^r T$ elements. In the source, reduction modulo ℓ^ν, the Bockstein map, and stabilization add a second torsion factor. More precisely, let B be a $\mathbb{Z}/\ell^\nu\mathbb{Z}$-module and let Φ_N be linear maps from the degree-$2e+1$ bar chains of $\mathrm{GL}_N(\mathcal{O}_L)$ with coefficients in $\mathbb{Z}/\ell^\nu\mathbb{Z}$ to B. Assume that these maps vanish on boundaries, agree under stabilization on cycles, and satisfy $\Phi_{dN}(\Psi_d z) = d\Phi_N(z)$ for every cycle z. Provided $2e+1 < r_2$, every finite subset of the values $\ell^{2e}\Phi_N(z)$, with N varying and z a cycle, has cardinality at most $(\ell^\nu)^{r_2} T_{4e+6,2e+1}(L) T_{4e+6,2e}(L)$. The assembly is `SUpperOf.lean`.

Here is the arithmetic quantity that this controls. Let $c \geq 1$, with $c \geq 2$ if $\ell = 2$, be such that L contains μ_{ℓ^c} but not $\mu_{\ell^{c+1}}$. For every integer $\nu \geq c$, consider a cyclic Galois extension F/L of degree $\ell^{\nu-c}$, a primitive ℓ^νth root $\zeta \in F$, and a generator γ of $\mathrm{Gal}(F/L)$ satisfying $\gamma(\zeta) = \zeta^{1+\ell^c}$. Put $a = (1+\ell^c)^e$. Let X_ν be the group of homomorphisms $\varphi : \mathrm{Cl}(F) \to \mathbb{Z}/\ell^\nu\mathbb{Z}$ satisfying $\varphi(\gamma^{-1}\mathfrak{a}) = a\varphi(\mathfrak{a})$. When $2e+1 < \ell^c$ and $2e+1 < r_2$, the source's cycle construction and finite-field specialization give

$$|X_\nu| \leq T_{4e+6,2e+1}(L) T_{4e+6,2e}(L) \ell^{(B_e + v_\ell(e))g + c + v_\ell(e)}. \tag{8.7}$$

Here $B_e = 2e + (2e+1)(c+2e+1)$, $g = 2n + \lfloor \log_2 h_L \rfloor + 2$, and h_L is the class number. The crucial feature is independence of ν. The construction associates homological cycles to the relevant twisted principal-divisor relations. Finite-field evaluation detects the weighted norms modulo ℓ^νth powers, while the preceding rank and torsion estimate bounds the possible values. A class-group count has a factor $(\ell^\nu)^{r_2}$ on its other side; this cancels the corresponding factor in the homological upper bound. These are specific cycle, detection, and counting theorems,

not consequences of group finiteness alone. Their exact interfaces are `SClass`, `SDetect`, `SFew`, and `SCount` in `LemmaKOfSteps.lean`. The actual proofs are inserted, together with `BOdd`, in `LemmaKOfLitN.lean`.

The passage through all cyclic layers gives an integer $s \geq 0$ such that $v_{\ell,e}(L) = s$ and ℓ^s is at most the right-hand side of (8.7). This includes finiteness of the length. When the inequalities $2e+1 < \ell^c, r_2$ do not initially hold, the source supplies a totally complex auxiliary field L' and integers c', d. The field L' contains $\mu_{\ell^{c'}}$ but not $\mu_{\ell^{c'+1}}$, with $c' \geq 1$ and $c' \geq 2$ if $\ell = 2$. They satisfy $2e+1 < \ell^{c'}$, $2e+1 < r_2(L')$, $n' = [L' : \mathbb{Q}] = dn$, $d \leq (n+1)(4e+4)$, $\ell^{c'} \leq 2n'$, and $|\operatorname{disc} L'| \leq |\operatorname{disc} L|^d (\ell^{c'})^{n'}$. Moreover $v_{\ell,e}(L) \leq v_{\ell,e}(L')$. These are the `SSha` and `SLayer` inputs in the same interface file. An integral arithmetic-group torsion estimate bounds both torsion factors: for an absolute constant C, and every $j \leq 2e+1$, $\log T_{4e+6,j}(L) \leq \exp((2en)^C \log(2|\operatorname{disc} L|))$. This includes all primes. The finite-complex and determinant construction is assembled in `Assembly.lean`; the form just used is proved in `ExponentLe.lean`. The degree and discriminant inequalities, the class-number estimate, and the torsion estimate give, after increasing an absolute exponent, $v_{\ell,e}(L) \log \ell \leq \exp((2en)^C \log(2|\operatorname{disc} L|))$. The numerical reduction is `NumThree.lean`.

Finally, Poitou–Tate duality and the local invariant calculation give

$$\operatorname{length}_{\mathbb{Z}_\ell} H^2(G_{L,S_\ell}, \mathbb{Z}_\ell(e+1)) \leq v_{\ell,e}(L) + n(2 + \lfloor \log_\ell n \rfloor + e). \tag{8.8}$$

Since $\mu_\ell \subset L$ for odd ℓ (and $\mu_4 \subset L$ when $\ell = 2$), one has $\ell \leq n+1$. Multiplying (8.8) by $\log \ell$ and increasing the absolute exponent therefore proves (8.6). This last step, including finiteness and the conversion between cardinality and module length, is `SFinal.lean`. The arithmetic-window proof calls precisely this route through `KFinal.lean` and `WindowBoundOfLiteratureLevel.lean`.

Explicit torsion estimates have substantial precedent, notably Soulé's bounds [27]. His stated estimate is up to small primes; it does not alone justify the all-prime input above. The present route uses the specified integral homology estimate, the filtered odd-rank bound, and the class-character comparison. These are the deep source inputs to this numerical deduction.

We next explain how a large fixed-layer bound can suffice. In a cyclotomic $\mathbb{Z}_\ell$-tower, the vanishing of the Iwasawa μ-invariant for an abelian base field is classical [11]. It gives a finite-rank module after the appropriate control and duality arguments. The source requires two distinct uniformity statements. At each fixed odd prime in $\mathcal{B}_\psi$, the base field is fixed by ψ and ℓ. At $\ell = 2$, the base field varies with q, d, and a separate estimate supplies a rank bound $R_0(\psi)$ independent of those primes. Its base degree is bounded by $8o$ and it embeds into a cyclotomic field of modulus $4N_\psi qd$. Thus its discriminant varies. The uniform rank assertion is the additional two-prime tower input, not a consequence of applying $\mu = 0$ separately to each varying field.

In the source the tower control is expressed by at most R numbers $s_1, \dots, s_{R'}$, $R' \leq R$, in a nonarchimedean extension of $\mathbb{Q}_\ell$. Away from a small neighborhood of these numbers, a twist e at layer m has length at most $R'(\max(m, j_0) + c + B)$ whenever $|s_i - e|_\ell \geq \ell^{-B}$ for every i. Here $j_0 = \lfloor \log_\ell R \rfloor + 3$, with the source's convention at $R = 0$, and c records the initial roots of unity. A second estimate bounds every twist by its value at layer j_0 plus $R(m - j_0)_+$. These two estimates, the fixed odd-prime and varying dyadic versions, are specified in `SecondBound.lean`. They are the actual bridge from Iwasawa theory to a finite exceptional set.

Lemma 8.5. *Let $s_1, \dots, s_r$ lie in a nonarchimedean extension of $\mathbb{Q}_\ell$, with the standard absolute value on $\mathbb{Q}_\ell$, and let $Z \geq 1$ be an integer. There is a set $\mathcal{E} \subset \{1, \dots, Z\}$ with $|\mathcal{E}| \leq r$ such that, for every $e \notin \mathcal{E}$ in this interval and every i, $|s_i - e|_\ell \geq \ell^{-\lfloor \log_\ell Z \rfloor}$.*

Proof. For each i, choose $e_i \in \{1, \dots, Z\}$ minimizing $|s_i - e_i|_\ell$, and put $\mathcal{E} = \{e_i : 1 \leq i \leq r\}$. For $e \neq e_i$, the ultrametric inequality and minimality give $|e - e_i|_\ell \leq |s_i - e|_\ell$. The nonzero integer $e - e_i$ has ordinary absolute value at most Z, so its ℓ-adic valuation is at most $\lfloor \log_\ell Z \rfloor$. This proves the assertion, including when one of the minimum distances is zero. □

For a fixed X, apply Lemma 8.5 with $Z = \lfloor X^{401} \rfloor$. This single truncation covers every twist needed at every scale $x \leq X$. The exceptional set may depend on X, q, d; its cardinality is bounded before those parameters vary. Combining the odd-prime sets with the dyadic set gives a bound R_ψ. Ordinary twists then have only logarithmic tower growth. Exceptional twists are

compared with a bounded layer and estimated by (8.6). The remaining task is to ensure that an exceptional twist which lies below x^{401} is already much smaller than x^{401}, simultaneously for every x on the selected window. This is a finite combinatorial problem.

Lemma 8.6 (Gap selection; `window1`)**.** *Let $R \in \mathbb{N}$, $B \geq 1$, $U > 0$, and let $T \subset \mathbb{R}$ have at most R elements. There is a real y with $U/(2^{R+2}B^R) \leq y \leq U/2$ such that, whenever $a \in T$ and $y \leq u \leq 2y$, the implication $a \leq u \Longrightarrow Ba \leq u$ holds.*

Proof. Induct on R. If $R = 0$, take $y = U/2$. If T has no element in $(U/(2B), U]$, the same choice works: $a \leq u \leq U$ forces $a \leq U/(2B)$, hence $Ba \leq U/2 \leq u$. Otherwise remove an element a_0 in that interval and apply the induction hypothesis to the remaining set with $U' = U/(2B)$. The resulting y satisfies $2y \leq U' < a_0$, so the removed element never satisfies $a_0 \leq u$. The other elements satisfy the desired implication by induction. The new lower bound is $U'/(2^{R+1}B^{R-1}) = U/(2^{R+2}B^R)$, as required. □

Take $U = \log X$, $T = \{\log e/401 : e \in \mathcal{E}\}$, and $B = 802C_K/\varepsilon$. If $0 < \varepsilon < 1/52$, then $B \geq 1$. With $W = e^y$ and $J = 2^{R_\psi+2}(802C_K/\varepsilon)^{R_\psi}$, Lemma 8.6 gives $X^{1/J} \leq W$, $W^2 \leq X$, and

$$e \in \mathcal{E}, \quad e \leq x^{401}, \quad W \leq x \leq W^2 \quad \Longrightarrow \quad e \leq x^{\varepsilon/(2C_K)}. \tag{8.9}$$

This is the quantitative content of the gap selection. No assertion that the exceptional twists disappear is needed.

The scale comparison must retain the auxiliary-prime discriminant. If the auxiliary scale is $Y_0 = X^{1/J}$, its prime sizes give $\log(qd) \ll Y_0^{\varepsilon/5}(1 + \log Y_0)^2$. On a selected window $x \geq W \geq Y_0$, the conductor in (8.3) therefore satisfies

$$\log N_h \ll_\psi \log x + x^{\varepsilon/5}(1 + \log x)^2, \qquad 2^h \asymp x^{2+8\varepsilon}. \tag{8.10}$$

For the bounded layers used on an exceptional twist, the degree is bounded in terms of the fixed data and $\log(2|\operatorname{disc} L|)$ satisfies the same type of bound. Substituting (8.9) into (8.6), its exponent is at most $C_\psi x^{\varepsilon/2}[\log x + x^{\varepsilon/5}(1 + \log x)^2]$, which is at most $C'_\psi x^{3\varepsilon/4}$ for sufficiently large x. Indeed, $\varepsilon/2 + \varepsilon/5 = 7\varepsilon/10 < 3\varepsilon/4$, and the remaining positive power absorbs the logarithms. Treating $\log|\operatorname{disc} L|$ as $O(\log x)$ would omit the dependence on q, d; the displayed estimate keeps it and still leaves sufficient room.

Likewise, $(N_h t)^C$ is generally not polynomial in x. What is needed, and what (8.10) proves for $t \leq x^{401}$, is $\log((N_h t)^C) = o(x^\varepsilon)$. A fixed power of x multiplying $\exp(Cx^{3\varepsilon/4})$ also has logarithm $o(x^\varepsilon)$. Consequently any fixed finite sum of these terms is eventually at most $\exp(x^\varepsilon/2)$. The threshold can be chosen uniformly in q, d because the displayed conductor bound and the number of exceptional twists were uniform. This is the reason for the generous exponential target.

Theorem 8.7 (Arithmetic window; `vwin_of_row`)**.** *Assume the source's arithmetic inputs described above, with its automorphic Galois and local compatibility input described in Section 9. Fix a primitive character ψ, a real $0 < \varepsilon < 1/52$, and any function $h : [1, \infty) \to \mathbb{N}$ satisfying $x^{2+8\varepsilon} \leq 2^{h(x)} \leq 2x^{2+8\varepsilon}$. Then there are $J \geq 1$ and X_0 such that, for every $X \geq X_0$ and every auxiliary pair q, d admissible at scale $X^{1/J}$, there is a real W with $X^{1/J} \leq W$, $W^2 \leq X$, and*

$$\mathcal{V}\left(E_{h(x)}, \lfloor x^{401} \rfloor; \sqrt{\exp(x^\varepsilon)}\right) \qquad (W \leq x \leq W^2). \tag{8.11}$$

Here the same pair q, d occurs in every field throughout the window. The constants J, X_0 are chosen before that pair and before the window.

Assembly of the arithmetic inputs. The fixed-layer input, the fixed odd-prime tower bounds, and the uniform dyadic tower bound give at most R_ψ exceptional twists. Lemmas 8.5 and 8.6, followed by the comparison just proved, bound the dyadic contribution and the fixed layers required in case IV by $\exp(x^{4\varepsilon/5})$ after enlarging a uniform threshold. The character estimates in cases I, Z, II, III, and V, together with their finite counting and the case-IV reduction, bound the plus contribution by $(N_h x^{401})^{C_\psi} + C_\psi \exp(x^{4\varepsilon/5})$. Proposition 8.2 supplies the odd-prime minus contribution, using $F \leq N_h$. Formula (8.10) absorbs all of these into the right side of (8.11).

Each input supplies finiteness as well as a bound for every finite prime sum, so the conclusion has exactly the meaning specified in (8.2).

The source performs this assembly in `JointWinOfCounts.lean`; the tower window is supplied by `ThmBabOfLemmaK.lean`. The final interface `VwinOfRow.lean` obtains precisely the statement above from its automorphic Galois row. The definitions of admissibility and of the universal dyadic selector are in `ExistsPageN.lean`. In particular, the conclusion is not restricted to a specially favorable choice of h. □

The arithmetic novelty relevant to the proof is thus concentrated in the compatible quantitative inputs: a norm product for the unrestricted prime sum, a division of residual configurations among unitary and symplectic constructions, and a uniform bound on the number of difficult tower specializations. The finite gap argument explains why these estimates are useful together. The elementary scale comparison is unusually forgiving, but only after the uniformity and the order of quantifiers have been established. Section 11 matches this arithmetic window to the analytic window and handles the fixed rescaling and the change from exponent 401 to exponent 400.

9. The automorphic Galois input

The preceding arithmetic estimates use Galois representations to interpret congruences between automorphic forms. The familiar existence theory for regular algebraic polarized cusp forms is established mathematics; for example, Chenevier and Harris construct semisimple representations with the requisite unramified and ℓ-adic Hodge properties in their setting [7]. The additional issue here is to obtain the particular local statements and integral comparisons used by the construction, with compatible normalizations. This section records that interface and explains the source's route through geometric representations, theta lifts, and quadratic descent. The route is a mathematical part of the development, whereas a claim that every dependency avoids a particular technology would require a separate audit of all imported results.

We specify the local statement before discussing its construction. Let E be CM with maximal real subfield F, let $a \geq 1$, and let π be a regular algebraic cuspidal automorphic representation of $\mathrm{GL}_a(\mathbb{A}_E)$. A polarization here consists of a Hecke character χ of F such that $\pi^c \simeq \pi^\vee \otimes (\chi \circ N_{E/F})$ and $\chi_v(-1) = (-1)^a$ at every real place v of F. Both the conjugate-duality relation and this archimedean sign condition are hypotheses of the source's `PolarizedPair`, defined in `DualIsConjByOne.lean`. Fix $\iota : \overline{\mathbb{Q}}_\ell \simeq \mathbb{C}$, and let π_v^{alg} denote the usual algebraic normalization of the local component; all Satake and Hodge normalizations below are those of this choice. We use a local parameter assignment $\mathrm{rec}_v^{\mathrm{aux}}$ with the properties specified at the end of this section. It agrees with the required character and depth-zero descriptions on the local classes used by the argument, while its laws on the remaining classes are the weaker ones stated there.

For a Weil–Deligne parameter $y = (r_y, N_y)$ over a nonarchimedean local field of residue cardinality q_v, say that its *tame multiplicity condition* holds if it is trivial on wild inertia and, for a topological generator γ of tame inertia, every eigenvalue z of $r_y(\gamma)$ of multiplicity at least two satisfies $z^{q_v-1} = 1$. The condition allows repeated eigenvalues of this specified kind; it is not a squarefreeness assertion. Say that y is *abelian* if its Weil representation factors through the abelianization. Finally, the source's large-residue-field condition is $E_v \simeq \mathbb{Q}_p$ with either $a = 1$ or $p > 16a^2$. An arbitrary finite extension of $\mathbb{Q}_p$ does not satisfy this condition merely because its residue characteristic is large.

Proposition 9.1 (Galois compatibility interface; `Row0b`)**.** *The automorphic Galois input used in Theorem 8.7 is the following assertion. For every E, π, ℓ, ι as above there exist a continuous semisimple representation $\rho_{\pi,\ell} : G_E \to \mathrm{GL}_a(\overline{\mathbb{Q}}_\ell)$, an integer w, and prescribed integral multisets H_τ for embeddings $\tau : E \hookrightarrow \overline{\mathbb{Q}}_\ell$. In the source's infinitesimal-character convention they satisfy*

$$H_\tau = \{(a-1)/2 - z : z \in \mathrm{Inf}_{\iota\tau}(\pi)\}, \qquad H_{\tau c} = \{w - h : h \in H_\tau\}, \tag{9.1}$$

with multiplicities. Put $y_v = \mathrm{rec}_v^{\mathrm{aux}}(\pi_v^{\mathrm{alg}})$. At $v \nmid \ell$, its dimension equals that of the Weil–Deligne parameter x_v attached to $\rho_{\pi,\ell}$. If $N_{y_v} = 0$ and y_v is abelian, then $x_v = y_v$. If instead $N_{y_v} = 0$,

the tame multiplicity condition and the large-residue-field condition hold, then $N_{x_v} = 0$ and the characteristic polynomials of x_v and y_v agree on inertia.

At $v \mid \ell$, the assertions are made in every compatible local realization: the local prime and period data must be tied to each other, and the comparison bridge must induce the chosen local component, in the precise sense of the source interface. In such a realization, a spherical local component gives a crystalline representation with Hodge–Tate multiset H_τ. More generally, if $N_{y_v} = 0$ and either the abelian condition or the two conditions in the tame branch hold, the local representation is de Rham and potentially semistable, has the prescribed Hodge–Tate multisets, and its Weil–Deligne parameter satisfies the same comparison with y_v just stated.

This is a statement of the precise input, not a claim that the listed local conclusions exhaust the established theory. The source definition is `WeightsConjBis.lean`. It quantifies over every permitted local parameter assignment and every compatible local realization. The spherical case includes the unramified Frobenius comparison needed to recover a representation from its characteristic polynomials. The controlled tame case supplies the inertia information used for the integral extension classes in Section 8. A proof of the stronger unrestricted local–global correspondence is not needed to apply this interface.

The geometric construction begins with compact unitary Shimura data of a controlled signature and with an automorphic eigensystem in coherent sections. Congruences obtained by multiplication with suitable Hasse invariants move the weights into a range where geometric cohomology supplies Galois representations. The construction must preserve enough Hecke data at the chosen places and enough information about Hodge jumps to identify the desired constituent later. The needed conclusion at this stage is deliberately weaker than the final rank-a representation.

More precisely, the source's geometric input assumes odd rank, its stated field and integral-model conditions, prescribed nonsplit places and levels, an auxiliary pair of places over ℓ, and a sufficiently separated weight condition. It constructs an auxiliary continuous semisimple representation with the required base property and Hodge occurrences. For each additional place away from ℓ it constructs such a representation satisfying the corresponding Hecke-polynomial relation. At suitable places above ℓ it constructs representations with the requested crystalline and Hodge properties, or with the specified crystalline-Frobenius annihilating polynomial. The auxiliary representation is allowed to depend on the place being tested. These quantifiers are explicit in `HodgeTrAt.lean`; its geometric assembly is `GeoOneOfParts.lean`.

This distinction is essential for understanding the construction. An annihilating polynomial says that the eigenvalues of an auxiliary representation belong to a specified set; it does not by itself give the multiplicities or the desired dimension. Similarly, occurrence of specified Hodge jumps does not exclude additional jumps. Extraction of a rank-a constituent, independence of the auxiliary choices, and transport of the local properties are subsequent arguments. The geometric input therefore does not compress all of those assertions into one application of a comparison theorem.

The theta step provides sufficiently many realizations in which the additional constituents are explicit. Let E/F be the fixed CM extension, let X be the definite Hermitian space of dimension n with its chosen standard lattice data, and let W be a definite skew-Hermitian space of sufficiently large dimension d with scalar Gram matrix. The splitting characters χ_X, χ_W have the prescribed restrictions to $\mathbb{A}_F^\times$ and archimedean exponents m_X, m_W. If U is an automorphic representation of the definite group $U(X)$ of dominant weight λ, the source's lifting input applies when $d + m_W(v) \le 2\lambda_k(v)$ at every infinite place and every weight coordinate. It produces an automorphic lift U' with local theta compatibility at split places and with weights given by

$$2\lambda'_k(v) = \begin{cases} 2\lambda_k(v) - d - m_W(v) + n + m_X(v), & k \le n, \\ n + m_X(v), & n < k \le d. \end{cases} \tag{9.2}$$

There is a threshold d_0 after the fixed field and Hermitian data have been chosen; the statement holds for every admissible $d \ge d_0$. The interface does not specify a universal formula such as

$d_0 = 2n$. Its exact statement, including the order of these quantifiers, is `TLiftK.lean`, and its construction is linked in `TLiftK.lean`.

The nonvanishing part is illuminated by two elementary positivity facts. First, let G be a unimodular locally compact group, π a unitary representation, and f a compactly supported test function for which the integrated operators exist. With $f^*(g) = \overline{f(g^{-1})}$, one has $\pi(f * f^*) = \pi(f)\pi(f)^*$, hence $\langle \pi(f * f^*)v, v\rangle = \|\pi(f)^*v\|^2 \geq 0$. Choosing f so that $\pi(f)^*v \neq 0$ gives strict positivity. This is the purpose of the local autocorrelation functions: positivity follows from an operator identity, rather than from the pointwise sign of a matrix coefficient.

Second, suppose an absolutely integrable kernel k and a matrix coefficient c satisfy $\int_K k(g)c(g)\,dg = A > 0$ on a measurable set K, and $\int_{G\setminus K} |k(g)c(g)|\,dg < A$. Then $\operatorname{Re}\int_G kc > 0$, by the triangle inequality. If K is a compact subgroup, $c|_K$ is the matrix coefficient $g \mapsto \langle \tau(g)v, v\rangle$ of an irreducible unitary representation τ with a unit vector v, and $k|_K = \overline{c|_K}$, Schur orthogonality gives $A = \operatorname{vol}(K)/\dim \tau$. Thus a sufficiently small tail gives a nonzero global pairing. The representation-theoretic work consists of constructing one set of local and archimedean data for which the theta inner-product identity, integrability, the specified compact contribution, and the tail estimate all hold. Classical theta correspondence and inner-product methods provide the background [15, 23]; the uniform choices for the present weight and lattice data form the source's additional construction.

At split unramified places the theta lift has the desired Satake parameters together with explicitly known character lines. It is tempting to subtract those lines from the trace and declare the result a representation. The required positivity of multiplicities has a separate algebraic proof, for which two compatible realizations are useful.

Lemma 9.2. *Let R_1, R_2, A_1, A_2 be finite-dimensional semisimple representations of a group over an algebraically closed field of characteristic zero. Suppose A_1, A_2 have no common irreducible constituent and $[R_1] - [A_1] = [R_2] - [A_2]$ in the Grothendieck group. Then this common virtual representation is the class of a semisimple representation V. If the representations are continuous, V can be chosen continuous.*

Proof. For an irreducible representation S, let $m(S)$ be its multiplicity in the common virtual class. If S does not occur in A_1, then $m(S)$ is its nonnegative multiplicity in R_1. If it occurs in A_1, it does not occur in A_2, so $m(S)$ is its nonnegative multiplicity in R_2. Only finitely many $m(S)$ are nonzero. Their direct sum defines V. Every constituent occurs in one of the continuous representations R_i, so its action and the action on the finite direct sum are continuous. □

In an application, equality of the virtual classes must itself be proved, usually by compatible Frobenius characteristic polynomials, density, and semisimplicity. The known extra lines must also be separated as actual constituents, with the twists correctly normalized. Lemma 9.2 explains the extraction once these hypotheses are available; an equality of numerical dimensions would not suffice. Uniqueness of semisimple Galois representations with the same Frobenius data then identifies the extractions made for different test places. This is how local properties of auxiliary geometric representations can become properties of one final representation. The compatibility statements needed for that transport remain part of the geometric and theta inputs.

The removal of auxiliary field restrictions is a descent problem. We give the underlying representation argument in some detail, because it clarifies both the role of semisimplicity and the precise meaning of patching a family of quadratic base changes.

Lemma 9.3. *Let H be an open normal subgroup of index two in a topological group G, and let k be an algebraically closed topological field of characteristic zero. A continuous finite-dimensional semisimple representation r of H extends to a continuous semisimple representation of G if r is isomorphic to its conjugate by every element of G. There are only finitely many isomorphism classes of semisimple extensions.*

Proof. Choose $t \in G \setminus H$. Conjugation by t permutes the irreducible constituents of r in orbits of size one or two and preserves their multiplicities. For a fixed irreducible constituent s, choose an

intertwiner A satisfying $As(h)A^{-1} = s(tht^{-1})$. Schur's lemma gives $A^2 = c\,s(t^2)$ for a nonzero scalar c. Replacing A by $c^{-1/2}A$ makes $A^2 = s(t^2)$, so s extends to G. Its two extensions differ by the nontrivial character of G/H. If its multiplicity is m, an extension of the isotypic summand is obtained by assigning multiplicities $m_+, m_- \geq 0$ to these two extensions with $m_+ + m_- = m$.

For an orbit $\{s, s^t\}$ of size two, $\mathrm{Ind}_H^G s$ is irreducible and restricts to $s \oplus s^t$. Its required multiplicity therefore extends that orbit. The direct sum over all orbits gives an extension, and the displayed finite choices exhaust semisimple extensions. Continuity follows on the open subgroup H and on its other coset, where the action is obtained by multiplication by a fixed matrix. □

Proposition 9.4 (Quadratic patching; `IndexTwoPatch.main`)**.** *Let G be a topological group and let $\eta_i : G \to \{\pm 1\}$ be infinitely many distinct continuous characters, with $H_i = \ker \eta_i$. Let k be an algebraically closed topological field of characteristic zero. Suppose that for each i there is a continuous semisimple representation $r_i : H_i \to \mathrm{GL}_a(k)$ whose trace is invariant under conjugation by G, and that the traces of r_i and r_j agree on $H_i \cap H_j$ for every i, j. Then there is a continuous semisimple $\rho : G \to \mathrm{GL}_a(k)$ whose restriction is isomorphic to r_i for all but finitely many i.*

Proof. The possible trivial character can be discarded. Trace invariance and semisimplicity imply conjugation invariance up to isomorphism, so Lemma 9.3 gives an extension of every r_i. Fix $i = 0$ and let $R_1, \ldots, R_s$ represent all semisimple extensions of r_0 to G.

We use a finite-index observation. Let R be any finite-dimensional representation of G, and let M be the preimage of the identity component of the Zariski closure of $R(G)$. This is a normal subgroup of finite index. If a quadratic character η is nontrivial on M, the Zariski closures of $R(\ker \eta)$ and $R(G)$ agree. To see this, the former closure contains the identity component, because its index in the latter is at most two. Moreover $G = (\ker \eta)M$, so its image meets every component. Only finitely many quadratic characters are trivial on M, since they factor through the finite quotient G/M. The same argument applies to any fixed finite-index subgroup of G. When that subgroup is an index-two subgroup H, restriction of quadratic characters has fibers of size at most two: two characters with the same restriction differ by a character of G/H. Thus finitely many exceptional restricted characters also exclude only finitely many original indices.

Apply this observation first to an extension of r_0, restricted to H_0. For all but finitely many i, restriction from H_0 to $L_i = H_0 \cap H_i$ preserves the irreducibility and distinctness of every constituent of r_0. These properties depend only on its Zariski image. Restriction of a semisimple representation to a normal subgroup of finite index is semisimple, by the finite-index constituent argument used in Lemma 9.3. The representations $r_i|_{L_i}$ and $r_0|_{L_i}$ are therefore isomorphic, by their equal traces and semisimplicity. The index-two constituent description in the preceding lemma now shows that r_i is the restriction to H_i of one of the R_j. Indeed, a constituent fixed by the nontrivial coset has exactly two extensions, differing by $\eta_0|_{H_i}$; a pair interchanged by that coset has its unique induced extension. The multiplicities on L_i specify the sums of the two multiplicities in the fixed case and the multiplicity in the induced case. These are exactly the choices already present among the R_j.

Apply the observation to the direct sum $R_1 \oplus \cdots \oplus R_s$. Choose an index i_1 outside the finite exceptional set so that restriction to H_{i_1} preserves all intertwiners between the R_j. Let ρ be a member of this finite list whose restriction is r_{i_1}. For all but finitely many i, the observation applied again, now on H_{i_1}, shows that $H_i \cap H_{i_1}$ has the same Zariski image under the direct sum as G. Choose R_j restricting to r_i. Pairwise compatibility makes R_j and ρ isomorphic on $H_i \cap H_{i_1}$. Equality of the Zariski images makes this an isomorphism on G, so $\rho|_{H_i} \simeq r_i$. The chosen ρ is continuous and semisimple by construction. □

For Galois groups, quadratic base change supplies the H_i, and matching Satake parameters supplies the equality of traces on intersections. Those two assertions require the base-change and Frobenius-compatibility theorems; they are not consequences of merely having many representations of the same dimension. The finite exceptional set in Proposition 9.4 is also significant. To recover a prescribed local property at a place v, the source constructs infinitely

many allowed quadratic extensions in which v splits. One can then choose such an extension outside the exceptional set. Its local Galois group identifies with the original local group, so the property descends without changing the local field. The generic patching declaration is proved in `EqOneOrEqOfKer.lean`; its arithmetic application is recorded in `ResInjective.lean`; the passage to local conclusions is part of `CoreAllOfRowStarCoreInp.lean`. This descent has classical precedents, and Sorensen's patching lemma gives an explicit formulation with a sufficiently general family of splitting extensions [26].

We finish by specifying the local parameter assignment used in Proposition 9.1. For a nonarchimedean completion k, call a supercuspidal representation *good* if it has dimension one, or if $k \simeq \mathbb{Q}_p$ and it is a character twist of a depth-zero supercuspidal representation. Call an irreducible admissible representation good when every member of its supercuspidal support is good. In this definition depth zero is expressed through fixed vectors for the first principal congruence subgroup in the chosen maximal compact. The requirement $k \simeq \mathbb{Q}_p$ is part of the source's class, not a convention that all local fields are prime fields.

On characters, $\mathrm{rec}_v^{\mathrm{aux}}$ is local reciprocity. On depth-zero supercuspidals over $\mathbb{Q}_p$, it is obtained from the regular finite-field character and the unramified-extension Weil parameter. The inertia eigenvalues are the Frobenius conjugates of that character; the finite reductive-group representation supplies the matching type. The classical finite-group and local reciprocity ingredients are separate from the extension to all supercuspidal supports [10, 20]. Langlands segments then prescribe the parameter of a general irreducible admissible representation as the corresponding sum of segment parameters.

Proposition 9.5 (The required local assignment; `RecTExists`)**.** *The local construction supplies an assignment* $\mathrm{rec}_v^{\mathrm{aux}}$ *with the following properties. It has the correct dimension, agrees with local reciprocity in dimension one, respects the segment description, sends supercuspidals to irreducible Weil–Deligne parameters, and respects real unramified twists. It has the prescribed absolute value of the determinant for every central character. On the good class it also respects arbitrary character twists and duals, has the exact determinant character, is injective against a good parameter, and satisfies the stated depth-zero tameness and finite-field type laws.*

The exact list is the structure of laws in `IsRecTOfIsRec.lean`; its existence is proved by the successive cuspidal and segment constructions in `RecTExists.lean`. Outside the good class the source is free to choose auxiliary wild parameters subject to the weaker global laws just listed. It does so using parameters that remain distinguishable from the good tame range. Such a completion of a partial assignment is a useful organizational device, but it does not assert a new full local Langlands correspondence on the arbitrary wild classes. In particular, exact determinant compatibility, arbitrary twisting, duality, and injectivity have the good-class qualifications in this interface.

The successive roles are now clear. Geometric cohomology supplies auxiliary representations with Hecke and Hodge information. Theta lifting gives realizations with explicit additional character lines. Constituent extraction produces the required dimension and makes the local realizations compatible. Quadratic patching removes field restrictions, while the splitting families restore each prescribed local assertion. The restricted local assignment expresses exactly the compatibility needed by the later congruence and Selmer estimates. What is specific to the development is this combination with its integral and quantitative constraints; the existence of automorphic Galois representations, theta correspondence, and quadratic descent themselves have substantial prior literature. With this input, the arithmetic conclusion available to the final proof is precisely Theorem 8.7.

10. Cancellation and the square-root budget

We now collect the output of the constructions into a numerical inequality. This step has two purposes. First, it shows why the local twisting factors and the archimedean periods must be normalized consistently. Second, it replaces an extensive table of guessed asymptotic equalities by the inequalities actually needed to obtain the exponent $1/2 + O(\varepsilon)$. The argument uses no further representation theory once those inequalities have been supplied.

Fix a thick translate and suppress it from the notation. Let $\mathcal{T}_1$ be the finite top set, let $m \geq 1$ be the tower length, and let Φ be a CM type of cardinality f. Write $a_t > 0$ for the top weights, $A = \sum_{t \in \mathcal{T}_1} a_t$, and $w_j(\rho) > 0$ for the archimedean weights at stage j. Let η be the imaginary quadratic character distinguished by the host. The archimedean expression that must pair with the finite gain is

$$\mathcal{M}(a, w) = a \left(\log a - 1 - \gamma_E + \operatorname{Re} \ell_\eta - \frac{1}{f} \sum_{\rho \in \Phi} \log w(\rho) \right). \tag{10.1}$$

It is real in this notation. The finite contribution $\mathcal{A}(a, t, w)$ is the Frobenius gain defined in the host construction. The balance identity of Section 4 expresses $\mathcal{A} - \mathcal{M}$ as a linear combination of the regularized prime sums G_χ.

The period normalization has the form

$$\log \Omega = \gamma_E + \log(\pi/v) - \operatorname{Re} \ell_\eta, \qquad v > 0. \tag{10.2}$$

This identity is the normalized form of the quadratic period formula used by the construction; the classical background is the formula of Selberg and Chowla [25]. The definition of the period and its relative-different normalization are part of the theta input. Only (10.2), rather than an unnormalized period symbol, can be substituted into the cancellation below.

Let i index the word spaces for the chosen translate, with positive weights p_i summing to 1. In the slope application these are normalized ranks. Write μ_i for their source slopes, h_i for their map-height bounds, and μ_T for the slope of the common target. Let $T_i > 0$ denote the auxiliary parameters occurring in the word count. The next proposition is a numerical version of the source's cancellation identity. It isolates every error term so that no equality is justified by an asymptotic estimate.

Proposition 10.1 (`key_identity`). *Suppose $g_0, v, \Omega, a_t, w_j(\rho), T_i$ are positive and (10.2) holds. Suppose there are real numbers M, G, F, E_i, L_i and nonnegative remainder bounds $E_{\mathrm{HN}}, E_D, E_C, E_T, E_P$ for which*

$$\mu_i \geq -M + G + F - E_i - L_i, \tag{10.3}$$

$$\sum_i p_i(\mu_i - h_i) \leq \mu_T + E_{\mathrm{HN}}, \tag{10.4}$$

$$\mu_T + h_i - F \leq m \sum_t a_t(\log a_t - 1 - \log \Omega) - A \log T_i + E_D, \tag{10.5}$$

$$E_i \leq mA \log g_0 + E_C, \tag{10.6}$$

$$\log T_i \geq \sum_{j=1}^{m} \frac{1}{f} \sum_{\rho \in \Phi} \log(w_j(\rho) v/\pi) + m \log g_0 - E_T, \tag{10.7}$$

$$\sum_{j=1}^{m} \sum_t \mathcal{A}(a_t, t, w_j) \leq G + E_P. \tag{10.8}$$

Then

$$\sum_{j=1}^{m} \sum_t (\mathcal{A}(a_t, t, w_j) - \mathcal{M}(a_t, w_j)) \leq M + E_{\mathrm{HN}} + \sum_i p_i L_i + E_D + E_C + AE_T + E_P. \tag{10.9}$$

Proof. Average (10.3) and combine it with (10.4). Rearranging gives $G \leq M + E_{\mathrm{HN}} + \sum_i p_i(E_i + L_i + \mu_T + h_i - F)$. Use (10.5) and (10.6) on the right. The result is

$$G \leq M + E_{\mathrm{HN}} + \sum_i p_i L_i + E_D + E_C + m \sum_t a_t(\log a_t - 1 - \log \Omega) - A \sum_i p_i \log T_i + mA \log g_0.$$

The common finite twisting term F has already disappeared exactly; it has not been bounded in absolute value. Next apply (10.7). Its level term cancels $mA \log g_0$. The remaining main expression is

$$m \sum_t a_t(\log a_t - 1 - \log \Omega) - A \sum_{j=1}^{m} \frac{1}{f} \sum_{\rho \in \Phi} \log(w_j(\rho) v/\pi).$$

Since $f = |\Phi|$, expanding the logarithm contributes $-mA\log(v/\pi)$. Substitution of (10.2) cancels this term against $-mA\log(\pi/v)$ from the period. What remains is exactly $\sum_{j,t}\mathcal{M}(a_t, w_j)$. The errors added in this substitution are AE_T. Finally use (10.8) and move the main expression to the left. This proves (10.9). □

In the arithmetic application, F is the common row-twisting contribution, G is the local divisibility gain, and E_i is the signed level contribution. Only its upper bound (10.6) is used; no positivity of E_i is required. The term M bounds the negative part of the initial column slope. The slope transfer in Section 5 supplies (10.4); the local and archimedean theta estimates supply (10.3), (10.5), and (10.6); and the size of the word family supplies (10.7). This identifies the exact place where each construction is used. The source identity appears in `WeightedMeanData.lean`, while its surrounding hypotheses are collected in `MeanBoundSecondForm.lean`.

There is a quantifier issue at this point. An estimate averaged over all thick translates does not imply the same estimate for every translate. The Fourier extraction used here needs a bound for each fixed translate before that family is averaged with its prescribed coefficients. The requisite source statement is the pointwise unit bound used in `ChTop.lean`. The weaker mean statement is useful for organizing the numerical algebra, but it cannot silently replace this pointwise input. Proposition 10.1 was formulated for one fixed translate precisely to keep that distinction visible.

Let $x \geq 1$ be the scale and let $\varepsilon > 0$ be fixed. We write n for the size of the host factor A, $r_T = |\mathcal{T}_1|$ for the number of top nodes, and n_2 for the size of the thick set. The notation r_T is separate from the rank $r = k + m - 1$ of a unitary space. Put $\delta_2 = n_2/n$. The weight b is a stretched exponential, typically $b = e^{x^\varepsilon}$; the budget below only needs its positivity. We use a_E for a common upper bound on the nominal source amplitude and the relevant top weights, with $a_E \leq 2xb$.

The construction supplies a total error bounded by $x^{C\varepsilon}\mathcal{B}$, after a threshold depending on its fixed constants, where

$$\mathcal{B} = r_T a_E\left(1 + \frac{m^2}{k}\right) + m\left(2nmbx^\varepsilon + \frac{n_2 kbx^\varepsilon}{n(nx^{-\varepsilon})}\right). \tag{10.10}$$

The first term includes the source-slope and finite divisibility errors. The second contains the Hodge and interpolation contributions, with their tower multiplicity. A large minimum-slope range does not appear as an uncontrolled product: Lemma 5.4 reduces its logarithmic contribution, and the trace errors are chosen small enough to control the remaining product with that range. When a trace estimate contains a term $\Gamma^2/|W|$, the size of the family W must be chosen to make this particular ratio sufficiently small; an arbitrary improvement of its decay exponent is not automatic.

The following elementary estimate is the part of the error budget that produces the final exponent. It uses inequalities rather than claimed exact scales for all auxiliary parameters.

Proposition 10.2 (`proposition_size_bound`)**.** *Suppose* $x \geq 1$, $\varepsilon > 0$, $n, r_T, m, k, b > 0$, $0 < n_2 \leq n$, $0 \leq a_E \leq 2xb$, *and*

$$\begin{aligned} r_T &\geq \tfrac{1}{2}nx^{-\varepsilon/8}, \qquad x^{1/2-\varepsilon} \leq m \leq x^{1/2+\varepsilon/2}, \\ xm &\leq k \leq n^2x^{1/2-\varepsilon}, \qquad \delta_2 = n_2/n \geq x^{-\varepsilon}. \end{aligned} \tag{10.11}$$

Then $\mathcal{B} \leq 10r_T mb\,x^{1/2+13\varepsilon/8}$. *In particular, if* $E \leq x^{C\varepsilon}\mathcal{B}$, *then*

$$\frac{2E}{r_T m\delta_2 b} \leq 20x^{1/2+(C+21/8)\varepsilon}. \tag{10.12}$$

Proof. Divide (10.10) by $r_T mb$. The weight bound and $k \geq xm$ give $a_E/(mb) \leq 2x/m$ and $a_E m/(kb) \leq 2$. The two remaining terms are $2(n/r_T)mx^\varepsilon$ and $n_2 kx^{2\varepsilon}/(r_T n^2)$. Thus

$$\begin{aligned} \frac{\mathcal{B}}{r_T mb} &\leq 2x^{1/2+\varepsilon} + 2 + 4x^{1/2+13\varepsilon/8} + 2x^{1/2+9\varepsilon/8} \\ &\leq 10x^{1/2+13\varepsilon/8}. \end{aligned}$$

For the last term we used $n_2 \le n$, the upper bound for k, and $n/r_T \le 2x^{\varepsilon/8}$. Every comparison is valid for $x \ge 1$. Multiplication by $2x^{C\varepsilon}/\delta_2$ and the lower bound for δ_2 prove (10.12). □

The formal statement `proposition_size_bound` packages this calculation with the Fourier inequalities and the admissible parameter predicates. Proposition 10.2 displays its numerical core, whose four terms can be checked independently. The source is `LipschitzWeightAndMinimum.lean`. No equality such as $k \asymp x^{9/2+10\varepsilon}$ is required. The sufficient range (10.11) is both more transparent and less restrictive.

To see how the residual slack is used, choose x so large that $x^{3\varepsilon/8} \ge 26$. Then (10.12) is at most $(20/26)x^{1/2+(C+3)\varepsilon}$. The remaining Fourier-rounding and Lipschitz errors have the form $O_{\psi,\varepsilon}(x^{\varepsilon}\log x) + O_{\psi,\varepsilon}(x^{\varepsilon-1})$. For fixed $C \ge 0$, they are at most $(6/26)x^{1/2+(C+3)\varepsilon}$ after enlarging the threshold. Indeed, each exponent gap is strictly positive, and a fixed logarithm is smaller than every fixed positive power of x eventually. The constants 20 and 6 are only a convenient division of the available budget.

It follows from the positive Fourier selector that

$$H_\psi(x) \le \Sigma_{\mathrm{small}}(x) + x^{1/2+(C+3)\varepsilon}, \tag{10.13}$$

where the discarded part consists of a fixed number of character sums controlled by the auxiliary-prime estimates. In the source's final choice the divisibility and loss exponents are $e_D = 2$ and $e_L = 0$, and the organization of the error terms gives $C = e_D + e_L + 3 = 5$. These values are visible in `ShapeWin.lean`; they are not inferred from the informal sizes of individual parameters. The discarded sums satisfy $\Sigma_{\mathrm{small}}(x) \le 4o\, x^{1/2+\varepsilon}$ for the fixed order o of ψ. Taking x still larger so that $1 + 4o \le x^{\varepsilon}$ yields

$$H_\psi(x) \le x^{1/2+9\varepsilon}. \tag{10.14}$$

For a fixed admissible family, these estimates hold at every allowed scale beyond one threshold. At each scale the analytic, cohomological, and geometric inputs must refer to compatible data. The individual objects, including auxiliary primes, may vary with the scale; the family and its comparison constants are fixed. The next section uses the family's arbitrarily distant windows to choose an entire interval beyond that threshold.

11. A common window and the conclusion

The estimates proved so far depend on several choices: a primitive character, a small exponent, two auxiliary primes, a dyadic layer, a host field, and a family of geometric data. The conclusion requires these choices in a specific order. It is not enough that each estimate can be made true by some choice of its own parameters. We first explain how the analytic and cohomological requirements are made to hold for the same prime pair, and then how the finitely many geometric requirements are imposed simultaneously.

Fix a primitive character ψ and $0 < \varepsilon < 1/52$. The analytic prime-selection theorem of Section 3 has the following quantifier form: for each fixed $J \ge 1$, after a threshold depending on ψ, ε, J, every large T admits primes q, d satisfying all the pair conditions on the entire interval $[T, T^J]$. In particular, their congruences and their size bounds are chosen once, at the scale T. The bounds for the relevant twisted character sums then hold for every argument in that interval.

The cohomological theorem has a different order of quantifiers. Let $h : [1,\infty) \to \mathbb{N}$ be any selector satisfying

$$x^{2+8\varepsilon} \le 2^{h(x)} \le 2x^{2+8\varepsilon} \qquad (x \ge 1). \tag{11.1}$$

Such a selector exists, for example by rounding $(2+8\varepsilon)\log_2 x$ upward. The arithmetic window theorem provides $J \ge 1$ and X_0 such that, for every $X \ge X_0$ and every prime pair satisfying its congruence and size conditions at $X^{1/J}$, there is a W with $X^{1/J} \le W$, $W^2 \le X$, and

$$\mathcal{V}(E_{h(x'),q,d}, \lfloor (x')^{401} \rfloor; \exp((x')^{\varepsilon}/2)) \qquad (W \le x' \le W^2). \tag{11.2}$$

Here $\mathcal{V}$ is the finite-prime-sum bound defined in Section 8, and the assertion holds for every host field with the specified fixed-field description. Thus J and X_0 are selected before q, d; the favorable interval is selected after them. This order permits the two results to be combined.

The fixed character field contributes a factor $c_\psi \geq 1$ to the size of the host. The analytic construction uses the variable x for which this total size is comparable to $x^{2+8\varepsilon}$, whereas (11.1) has been stated for the dyadic factor alone. The following explicit rescaling reconciles them without changing the prime pair.

Proposition 11.1 (`winPairs`)**.** *Assume the analytic pair-selection statement and the cohomological window statement above. For each fixed $c \geq 1$ there is a $J \geq 1$ such that arbitrarily large Z admit a single prime pair q, d with the following property. For every $x \in [Z, Z^{3/2}]$, there is a dyadic exponent e and a scale T such that the analytic pair conditions hold on $[T, x]$, $x^{1/J} \leq T \leq x$,*

$$x^{2+8\varepsilon} \leq c2^e \leq 2x^{2+8\varepsilon}, \tag{11.3}$$

and every associated host field satisfies

$$\mathcal{V}(E_{e,q,d}, \lfloor x^{400} \rfloor; \exp(x^\varepsilon/2)). \tag{11.4}$$

The prime pair and T can be kept fixed throughout this interval. The dyadic exponent e and the ensuing geometric data may vary with x.

Proof. Choose h as in (11.1), and obtain J, X_0 from the cohomological statement. Use this same J in the analytic pair-selection theorem, and let T_0 be its threshold. Put $\kappa = c^{1/(2+8\varepsilon)} \geq 1$. Given a desired lower bound for Z, choose M above that bound, T_0, 1, κ^3, and κ^{400}. Choose $X \geq \max(X_0, M^J)$ and put $T = X^{1/J}$. Then $T \geq M$, and analytic pair selection supplies q, d on $[T, X]$.

Apply the cohomological theorem to this pair, giving $W \geq T$ with $W^2 \leq X$. Set $Z = \kappa W$. Since $W \geq \kappa^3$, one has $(\kappa W)^{3/2} \leq W^2$. Hence, if $Z \leq x \leq Z^{3/2}$ and $x' = x/\kappa$, then

$$T \leq W \leq x' \leq x \leq W^2 \leq X. \tag{11.5}$$

The analytic conditions remain true on $[T, x]$, and $x^{1/J} \leq X^{1/J} = T \leq x$. Taking $e = h(x')$ and using $\kappa^{2+8\varepsilon} = c$ proves (11.3).

The cohomological estimate (11.2) is available at x'. Since $x' \geq W \geq \kappa^{400}$, we have $x^{400} = \kappa^{400}(x')^{400} \leq (x')^{401}$, and therefore $\lfloor x^{400} \rfloor \leq \lfloor (x')^{401} \rfloor$. Also $x' \leq x$ gives $\exp((x')^\varepsilon/2) \leq \exp(x^\varepsilon/2)$. Monotonicity of the finite-prime-sum bound in its twist cutoff and its upper bound proves (11.4). The pair was never changed, and $Z \geq W \geq M$ is as large as required. □

This proof explains two otherwise arbitrary-looking choices in the development. The exponent 401 reserves one power for a fixed rescaling, while the final cohomological input only requires twists below x^{400}. The interval exponent $3/2$ leaves room inside the quadratic arithmetic window after the same rescaling. Neither choice is an assertion that 401 or $3/2$ is optimal. Their purpose is to make the inclusions (11.5) uniform.

It remains to impose the construction requirements on one family of data. These requirements have many auxiliary constants, so the source formulates them as properties closed under increasing an admissibility threshold. The elementary logical principle is as follows. Suppose finitely many properties $P_1, \ldots, P_s$ satisfy: for sufficiently small $\varepsilon > 0$, there is a threshold $K_j(\varepsilon)$ such that every admissible family beyond that threshold has P_j. Suppose also that for every prescribed threshold there exists an admissible family. Choose ε below all the finitely many allowed upper bounds, and then choose the family beyond $\max_j K_j(\varepsilon)$. It satisfies all the properties simultaneously. If a threshold is a function of several fixed parameters, its maximum is taken pointwise; the argument is unchanged.

The order in this principle is important. The exponent is fixed first, each threshold is obtained with that exponent, their finite maximum is formed, and the existence theorem is invoked last. It is not legitimate to choose a family separately for each property and then identify the resulting objects. The formal closure machinery makes this distinction explicit.

Proposition 11.2 (`topDesWin`)**.** *Assume the design-existence theorem for the host class, the cohomological condition* (11.4)*, and the nine groups of eventual construction properties listed below. In the design-existence hypothesis, require that for every threshold the returned family belongs to a class $\mathcal{C}'$ contained in both the geometric class $\mathcal{C}$ and the arithmetic class $\mathcal{E}$; it must also supply the smooth-denominator and compatible-twisting conditions, a nonnegative comparison constant C_H with the chosen-evaluation witness* `ChosenE`*, and the size witnesses* `TuSize` *and* `DGamSize`*. Assume that every family in $\mathcal{E}$ satisfies the cohomological bound* (11.4) *eventually, for every datum at the relevant scale, and that its admissible scale set contains intervals $[Y, Y^{3/2}]$ with arbitrarily large Y.*

Then, for every sufficiently small $\varepsilon > 0$, there is one family $\mathcal{F}_\varepsilon \in \mathcal{E}$ satisfying the divisibility, frame, rank, slope, and global comparison hypotheses required for (10.13)*, with fixed comparison constants and one threshold valid at every admissible scale beyond it. The admissible scale set of this same family still contains arbitrarily distant intervals $[Y, Y^{3/2}]$.*

Source property	Mathematical role
`LinkN`	Bounds the integral comparison defect by the cohomological lengths and explicit size errors.
`LinkEta`	Supplies the local covering condition compatible with the chosen twists.
`LinkIRR`	Bounds the discarded-row proportion by $\exp(-x^{e_z/7})$, giving the retained-target rank share.
`SmallLevelRankEq`	Identifies the relevant rank after the small-level passage.
`LinkDIVt`	Converts the local covering and defect bounds into integral evaluation maps, height, period, level, and slope comparisons.
`NuFullF`, `MuRank`	Gives the required full lattice and positive-rank assertions.
`LinkG`	Supplies the finite-place gain, local lattice-index bounds, and the comparison of served Frobenius orbits.
`LinkS`	Supplies the spectral and approximate-frame estimates.
`LinkGLOBAL`	Combines these inputs into the global slope budget.

Proof. Apply the finite maximum principle to these nine groups of eventual properties and to the threshold in the design-existence theorem. Choose a family beyond the resulting threshold, retaining its witnesses. The inclusion $\mathcal{C}' \subset \mathcal{C}$ permits the geometric properties to be applied; $\mathcal{C}' \subset \mathcal{E}$ supplies the cohomological hypothesis and preserves the full-window property for that same family. The smooth-denominator and twisting conditions place the family in the design class. The size witnesses enable `LinkDIVt`, `LinkS`, and `LinkGLOBAL`, while `ChosenE` enables the full-lattice and rank properties. The cohomological condition, together with the integral comparison and the small-level rank identity, supplies the normalized defect bound. With the local covering condition, this enables the integral evaluation-map and height comparisons in `LinkDIVt`. The full-lattice and positive-rank assertions are still needed for the slope comparison. The additional quantitative retained-rank share comes from `LinkIRR`; positive rank alone would not supply it.

The geometric, spectral, retained-rank, and divisibility properties now refer to the same family. Insert them into the global comparison. This gives the source and row bounds, the slope-transfer estimate, and the error budget used in Section 10, with fixed constants and one eventual threshold for this family. The explicit proof of this finite join is `DesWin.lean`; the order "family, constants, threshold, all admissible scales" is recorded by `FamilyClassConnected` in `DivinNOfHypV.lean`. The specialization `ShapeWin.lean` supplies the values $C_{\mathrm{abs}} = 2$, $C_0 = 3$, $e_D = 2$, and $e_L = 0$ used in the final numerical estimate. □

A family here comprises data that may vary with the scale x, including its auxiliary primes. Proposition 11.1 supplies fixed pairs on input windows, but the abstract family selection need only preserve the pointwise requirements and the existence of full windows. The analytic function H_ψ is independent of these auxiliary choices. The proposition is an assembly theorem: its nine

groups of hypotheses are substantive construction results. Sections 6 and 7 explain the local, geometric, and frame inputs; Sections 8 and 9 explain the arithmetic and Galois inputs. The cited source modules prove their eventual closure and the required existence statement. The finite maximum principle puts them on one family before a final window is selected.

Proof of Theorems 1.1–1.3. Fix a primitive character ψ. The Galois construction in Section 9 supplies the precise local-global row used both by the theta comparison and by the cohomological argument. The unconditional source declaration closing that input is in `RowSrc.lean`. The cohomological argument gives the arithmetic window theorem of Section 8. The auxiliary-prime theorem of Section 3, combined with Proposition 11.1, gives arbitrarily distant intervals $[Y, Y^{3/2}]$ on which the same pair satisfies the analytic and arithmetic requirements.

Fix a sufficiently small $\varepsilon > 0$. The host and positive Fourier construction of Section 4 fit these data into the design class. Apply Proposition 11.2 once to obtain one family $\mathcal{F}_\varepsilon$, depending on ψ and ε, whose admissible scale set W_ε contains intervals $[Y, Y^{3/2}]$ with arbitrarily large Y. The comparison constants are now fixed. Section 10 therefore supplies x_ε such that

$$H_\psi(x) \le x^{1/2+9\varepsilon} \qquad (x \ge x_\varepsilon,\ x \in W_\varepsilon).$$

Given $A > 0$, choose $Y \ge \max(A, x_\varepsilon, 1)$ with $[Y, Y^{3/2}] \subset W_\varepsilon$. The estimate holds at every point of this interval, proving (1.3). The data at each x may vary; the family and its constants were fixed before Y was chosen. The source makes this last choice in `TopWin.lean`, using the whole-interval property in `FamCore.lean`.

The residue computation and Theorem 2.2 now exclude every zero of $L(s, \psi)$ with real part greater than $1/2$. Equation (2.12) extends this assertion to every Dirichlet character, proving Theorem 1.1. The functional equations and the final reductions in Section 2 prove Theorems 1.2 and 1.3. □

At the final source interface, `MainConclusion.lean` first derives the all-character theorem from the Galois row and then inserts the constructed row. The local parameter existence theorem is provided separately; it is not hidden in a blanket appeal to local-global compatibility. The wrapper `MainTheorems.lean` passes the resulting theorem to the public declarations in `Main.lean`. This is the point at which the analytic and arithmetic branches of the argument meet.

Appendix A. Reading the Lean development

All source links in this paper are pinned to commit

`7cbf7a0c32fca1badf05931879a37c3f8a2b03b4`

of the repository supplied with the manuscript [1]. The project specifies Lean 4.33.1 and Mathlib commit `0df444a360eaa60ab8c11dca51a86af692955474`. Pinning the links is useful here because a theorem name alone does not specify its hypotheses, normalizations, or imported constructions.

The public statements use Mathlib's Dirichlet characters, L-functions, zeta function, and primitive-character predicate. In Lean, these are total functions, so their value at a classical pole is assigned by a convention. The exclusions at $s = 1$ in the public statements prevent that convention from being mistaken for an analytic assertion. The mathematical formulation throughout the paper uses meromorphic functions and their usual domains.

Public declaration in `Main.lean`	Mathematical statement
`riemannHypothesis`	Mathlib's Riemann-hypothesis predicate; Theorem 1.3.
`riemannZeta_eq_zero_re`	The same zeta-zero assertion written using the local predicate.
`LFunction_ne_zero_of_half_lt_re`	Nonvanishing in $\operatorname{Re} s > 1/2$, excluding a principal pole; Theorem 1.1.
`LFunction_eq_zero_re`	All-character critical-strip theorem; Theorem 1.2.
`LFunction_eq_zero_re_of_isPrimitive`	Its restriction to primitive characters; the hypotheses still include $0 < \operatorname{Re} s < 1$.

The following concordance gives entry points for the principal mechanisms. A file may state an interface and prove it by assembling several imported results; its name is not evidence that the corresponding mathematical assertion is an elementary lemma. Conversely, the large number of supporting files should not obscure the small number of deductions that connect the main parts.

Mathematical passage	Source entry points
Mellin transforms and residues; Proposition 2.1	`poleData.lean`. The family statement also records the positive weighted multiplicities of poles.
One-sided pole exclusion; Theorem 2.2	`landauWindow.lean`, with `landauTails.lean` and the nearest-pole lemmas.
Auxiliary prime pair and thin estimates; Section 3	`pairEx.lean`; the exact pair predicate is in `PairGood.lean`.
Fourier balance and extraction; Section 4	`ChBalance.lean`, `ChCore.lean`, and `ChTop.lean`.
Frame codimension estimate; Proposition 5.2	`FrameLinearAlgebraFirst.lean`.
Sharp mean-slope transfer; Theorem 5.3	`StatementFrameF.lean`, `TheoremFSharp.lean`, and the public wrapper in `ShiftInvariance.lean`.
Theta spaces and integral data; Section 6	`DataSpaces.lean`, `DataKernels.lean`, and `DataArith.lean`; the section gives the more specialized local input links.
Geometric covolume and spectral inputs; Section 7	`TopCovolSlim2Use.lean`; the interpolation and frame interfaces are linked at their statements.
Uniform fixed-layer cohomology; Proposition 8.4	`ClassCounting/LemmaKOfLitN.lean`, through `KFinal.lean`. The used route is the class-character count and its integral homology inputs.
Cohomological lengths and arithmetic windows; Section 8	`Statement.lean`, `JointWinOfCounts.lean`, and `VwinOfRow.lean`.
Galois row and local parameter; Section 9	`RowSrc.lean`, `RecTExists.lean`.
Exact normalization cancellation; Proposition 10.1	`WeightedMeanData.lean`. The pointwise input is kept distinct from the average formulation.
Final exponent calculation; Proposition 10.2	`LipschitzWeightAndMinimum.lean`.
Common auxiliary pair; Proposition 11.1	`WinPairs.lean`, with the quantified window definitions in `ExistsPageN.lean`.
Finite join and final theorem; Proposition 11.2	`DesWin.lean`, `ShapeWin.lean`, and `MainConclusion.lean`.

Several conventions deserve particular attention when moving between prose and source. The slope μ used in the arithmetic comparison is twice the conventional normalized Arakelov slope, and the map height is doubled with it. A CM type has half as many elements as there are

complex embeddings; one must not divide an embedding sum by the size of a CM type without the matching factor of two. The retained row space is formed from the appropriate algebraic row subspace and its kernel; it is not obtained by replacing that kernel with the whole kernel in an unrelated ambient Hilbert space. The predicate `MuRank` asserts positive rank and does not, by itself, assert that only a small fraction of the ambient rank was discarded. A word is the source's prescribed tuple of local and archimedean data, not necessarily a binary string. These distinctions affect the hypotheses of the frame and covolume estimates.

The status of a mathematical statement also matters. A `def` ending in `: Prop` can merely specify the assertion to be proved. A later theorem proves that assertion from explicit hypotheses. The source-input statements in Sections 6–9 preserve those hypotheses; they should be read together with the construction modules that discharge them. The elementary lemmas proved in this paper explain how those statements are used, rather than supplying substitutes for the omitted deep constructions.

For reproducibility, a local formal check should first use the pinned toolchain and dependency manifest, then build the main target, inspect the axioms of the five public declarations, and replay the resulting environment with an appropriate checker. Those are separate operations. A text search for placeholders does not establish compilation or correctness of the elaborated proof, and successful kernel checking does not automatically verify a new prose translation.

Independent verification is reported by Andrew Sutherland's MIT verification page, updated 21 November 2026 at 12:55 UTC [28]. It reports compilation of all 175,187 modules, only the three standard axioms `propext`, `Classical.choice`, and `Quot.sound` for the five public theorems, complete module replay by `leanchecker` and `lean4lean`, and acceptance of the exported proof by the independent kernels `nanoda` and `con-leche`. The independent statement comparison in `con-leche` is also reported as passing. At that update, the full `lake comparator` run and the Lean-kernel replay of the export were still running. These are reported external checks of the published development; a new full Lean build was not performed for this editorial revision. The mathematical correspondence between this exposition and the cited source statements remains a separate subject for review.

References

[1] L. Alpöge, with Claude, *Lean development of the Riemann hypothesis and the critical-line theorem for Dirichlet L-functions*, 2026. Source mirror supplied with the manuscript, commit `7cbf7a0c32fca1badf0593 1879a37c3f8a2b03b4`, https://github.com/CocoSgt/riemann-hypothesis-lean.

[2] J. Anschütz and A.-C. Le Bras, *Prismatic Dieudonné theory*, Forum Math. Pi **11** (2023), e2, 1–92. https://doi.org/10.1017/fmp.2022.22.

[3] F. A. Behrend, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Natl. Acad. Sci. USA **32** (1946), no. 12, 331–332. https://doi.org/10.1073/pnas.32.12.331.

[4] J. Bellaïche and G. Chenevier, *Families of Galois representations and Selmer groups*, Astérisque **324**, Soc. Math. France, Paris, 2009, 326 pp. https://www.numdam.org/item/AST_2009__324__R1_0/.

[5] B. Bhatt and P. Scholze, *Prisms and prismatic cohomology*, Ann. of Math. (2) **196** (2022), no. 3, 1135–1275. https://doi.org/10.4007/annals.2022.196.3.5.

[6] J.-B. Bost, *Périodes et isogénies des variétés abéliennes sur les corps de nombres (d'après D. Masser et G. Wüstholz)*, Astérisque **237** (1996), Sém. Bourbaki exp. 795, 115–161. https://www.numdam.org/item/SB_1994-1995__37__115_0/.

[7] G. Chenevier and M. Harris, *Construction of automorphic Galois representations, II*, Cambridge J. Math. **1** (2013), no. 1, 53–73. https://doi.org/10.4310/CJM.2013.v1.n1.a2.

[8] J. Coates and W. Sinnott, *An analogue of Stickelberger's theorem for the higher K-groups*, Invent. Math. **24** (1974), 149–161. https://doi.org/10.1007/BF01404303.

[9] H. Davenport, *Multiplicative Number Theory*, third ed., revised by H. L. Montgomery, Grad. Texts in Math. **74**, Springer, New York, 2000. https://link.springer.com/book/9780387950976.

[10] P. Deligne and G. Lusztig, *Representations of reductive groups over finite fields*, Ann. of Math. (2) **103** (1976), no. 1, 103–161. https://doi.org/10.2307/1971021.

[11] B. Ferrero and L. C. Washington, *The Iwasawa invariant μ_p vanishes for abelian number fields*, Ann. of Math. (2) **109** (1979), no. 2, 377–395. https://doi.org/10.2307/1971116.

[12] W. T. Gan, Y. Qiu, and S. Takeda, *The regularized Siegel–Weil formula (the second term identity) and the Rallis inner product formula*, Invent. Math. **198** (2014), no. 3, 739–831. https://doi.org/10.1007/s00222-014-0509-0.

[13] D. R. Grayson, *Reduction theory using semistability*, Comment. Math. Helv. **59** (1984), 600–634. https://doi.org/10.1007/BF02566369.
[14] H. Hida, *p-Adic Automorphic Forms on Shimura Varieties*, Springer Monogr. Math., Springer, New York, 2004. https://doi.org/10.1007/978-1-4684-9390-0.
[15] R. Howe, *θ-series and invariant theory*, in Automorphic Forms, Representations and L-Functions, Part 1, Proc. Sympos. Pure Math. **33**, Amer. Math. Soc., Providence, RI, 1979, pp. 275–285. https://doi.org/10.1090/pspum/033.1/546602.
[16] M. N. Huxley, *An imperfect hybrid zero-density theorem*, J. London Math. Soc. (2) **13** (1976), no. 1, 53–56. https://doi.org/10.1112/jlms/s2-13.1.53.
[17] K. Kato, *p-adic Hodge theory and values of zeta functions of modular forms*, Astérisque **295** (2004), 117–290. https://www.numdam.org/item/AST_2004__295__117_0/.
[18] V. A. Kolyvagin, *Euler systems*, in The Grothendieck Festschrift, Vol. II, Progr. Math. **87**, Birkhäuser, Boston, MA, 1990, pp. 435–483.
[19] E. Landau, *Über einen Satz von Tschebyschef*, Math. Ann. **61** (1905), no. 4, 527–550.
[20] J. Lubin and J. Tate, *Formal complex multiplication in local fields*, Ann. of Math. (2) **81** (1965), 380–387. https://doi.org/10.2307/1970622.
[21] A. Moriwaki, *Continuity of volumes on arithmetic varieties*, J. Algebraic Geom. **18** (2009), no. 3, 407–457. https://doi.org/10.1090/S1056-3911-08-00500-6.
[22] T. Ohsawa and K. Takegoshi, *On the extension of L^2 holomorphic functions*, Math. Z. **195** (1987), no. 2, 197–204. https://doi.org/10.1007/BF01166457.
[23] S. Rallis, *Injectivity properties of liftings associated to Weil representations*, Compos. Math. **52** (1984), no. 2, 139–169. https://www.numdam.org/item/CM_1984__52_2_139_0/.
[24] K. A. Ribet, *A modular construction of unramified p-extensions of $\mathbb{Q}(\mu_p)$*, Invent. Math. **34** (1976), no. 3, 151–162. https://doi.org/10.1007/BF01403065.
[25] A. Selberg and S. Chowla, *On Epstein's zeta-function*, J. Reine Angew. Math. **227** (1967), 86–110. https://doi.org/10.1515/crll.1967.227.86.
[26] C. M. Sorensen, *A Patching Lemma*, in Shimura Varieties, T. Haines and M. Harris, eds., Cambridge University Press, Cambridge, 2020, pp. 297–305. https://doi.org/10.1017/9781108649711.012.
[27] C. Soulé, *A bound for the torsion in the K-theory of algebraic integers*, Doc. Math., Extra Volume: Kazuya Kato's Fiftieth Birthday (2003), 761–788. https://doi.org/10.4171/DMS/3/21.
[28] A. V. Sutherland, *Verification files for the Lean proof of the Riemann Hypothesis*, Massachusetts Institute of Technology, update of 21 November 2026, 12:55 UTC. https://math.mit.edu/~drew/grh/.
[29] R. Taylor and A. Wiles, *Ring-theoretic properties of certain Hecke algebras*, Ann. of Math. (2) **141** (1995), no. 3, 553–572. https://doi.org/10.2307/2118560.