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 formal development was announced in [2] and has since been checked independently. At the time of writing, the verification page [31] reports that all 175,187 modules compile, that the five public theorems depend only on the axioms `propext`, `Classical.choice`, and `Quot.sound`, and that the compiled modules and the exported proof are accepted by several independent checkers; two further kernel replays were still running. The details and caveats are recorded at the end of Section A. Those reports concern the Lean source. They do not certify the prose below, whose correspondence with the source is for the reader to test; this is why every principal statement carries a pinned source location.

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. The contribution from source spread depends logarithmically on the target slope range; a separate term multiplies that range by the frame defects and is controlled by the spectral estimates. This is the role of Theorem 3.5.

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 [11, 21]. The formulation on a selected sequence of power 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 [4]; 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 [15, 7]; 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 7–10 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 main distinctions can be summarized more precisely as follows. The third column identifies what has to be established for this argument; it should not be read as a claim that the classical results in the middle column already imply the required specialization. The source locations for these assertions appear with their statements and in the final concordance.

Mechanism	Established apparatus	Construction-specific assertion
Pole detection	Mellin transforms, logarithmic derivatives, and positivity	A one-sided estimate on arbitrarily distant whole windows excludes each pole with positive weighted multiplicity.
Character selection	Finite Fourier inversion and large configuration-free sets	A parity-compatible host and positive selector recover exactly the coefficient $1/(3o+1)$ while allowing the unwanted twists to be controlled.
Slope transfer	Arithmetic slopes and Harder–Narasimhan filtrations	The pooled frame estimate bounds codimension at every slope threshold, giving a logarithmic source-spread loss and a separately controlled frame-defect term.
Theta comparison	Local theta theory, period formulas, and geometric norm estimates	One integral tower satisfies the divisibility, common-target, metric, counting, and frame estimates simultaneously.
Cohomological control	Group homology, arithmetic cohomology, and characteristic ideals	The required class counts, torsion control, and root bounds hold with the uniformity needed as the host layer varies.
Final assembly	Elementary comparison of powers and finite maxima	One admissible family is fixed before its constants, threshold, and final full window are used.

The distinction between a mechanism and its realization is especially important for the theta part. A theorem about the norm of a single lift does not establish that many lifts have nearly orthogonal images. A theorem about their complex images does not establish that the maps preserve the required integral lattices. A theorem about one host field does not establish uniformity when its degree changes with the scale. The separate constructions meet these separate obligations. The fixed-data comparison in Section 4 shows exactly where each obligation enters and permits the elementary deductions to be checked before those constructions are read.

Several symbols recur in different branches of the proof. The following list records their roles at a fixed scale; each object is defined in the section where it is constructed. In particular, none of the approximate sizes in the discussion is intended to replace the inequalities of Proposition 4.3. One letter is used in two ways that do not meet: h_i and $h(B)$ denote map heights in Sections 3–8, whereas in Sections 9 and 11 the letter h denotes the dyadic layer of the host field.

Notation	Role in the comparison
ψ, o, ε	The fixed primitive character, its order, and the small positive exponent controlling all allowed losses.
x, m, k, b	The prime-sum scale, the theta-tower length, a rank parameter, and the common weight scale; m is near $x^{1/2}$.
q, d, E, Φ	The two auxiliary primes, the resulting CM host, and its CM type. The field of coefficients of a lattice is a separate object.
n, r_T, n_2	The size of the host factor, the top-set cardinality, and the thick-set cardinality. The density δ_2 is n_2/n.
μ, h_i, p_i	The doubled arithmetic slope, the corresponding map-height bound, and the source-rank probability weights.
δ_F, η_F, e_F	The discarded-target proportion and the two collective frame defects; these are distinct from ε and δ_2.
$\mathcal{A}, \mathcal{M}, H_\psi$	The finite divisibility gain, the matching archimedean main term, and the fixed analytic function obtained by character extraction.

In outline, the logical dependencies are as follows. Theorem 2.2 reduces Theorems 1.1–1.3 to the window estimate (1.3). Proposition 6.4 reduces (1.3) to an upper bound for the comparison gap $\sum(\mathcal{A} - \mathcal{M})$ on a host field, given the estimates of Section 5 for the discarded characters and for the thin arguments (Propositions 5.1 and 5.3). Proposition 4.1 reduces that bound to six inequalities. One of them, the transfer (4.4), is Theorem 3.5 applied with the frame estimates of Proposition 8.9. The others are supplied by the theta tower of Section 7, through the local divisibility estimate (Proposition 7.7) and the global interface (Proposition 7.11), together with the covolume and interpolation estimates of Section 8. The integral comparison inside these inputs consumes the cohomological bound of Theorem 9.10, which is available only on windows and which uses the Galois interface of Section 10. Propositions 11.1 and 11.2 place all requirements on one family and one window, and Proposition 4.3 converts the resulting error into the exponent $1/2 + 9\varepsilon$.

The argument may be read in two passes. Sections 2, 3, 4, and 11 give the analytic, linear-algebraic, and quantitative deductions. In this order the reader first learns why the desired window estimate is enough, then why an approximate frame transfers mean slopes, and finally why the correctly normalized comparison has a square-root budget. The remaining sections construct the hypotheses needed for those deductions. They distinguish the elementary bridges proved here from the deeper source inputs and explain which uniformity is retained when one construction feeds another.

Only principal propositions retain the source's declaration names in their headings. Auxiliary lemmas use ordinary mathematical descriptions. The name of a source predicate is not used as a substitute for its hypotheses: where an input is needed, its mathematical content is stated before its source location is given. This also permits the paper to be read without following each hyperlink. A reader returning to the formal development can use the concordance to locate

the declaration, inspect its precise hypotheses, and then follow the construction that discharges them.

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 [11]. 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.

The smoothing in this definition has a useful elementary interpretation. Put $\Psi_\chi(t) = \sum_{n\le t}\Lambda(n)\chi(n)$. Then

$$S_\chi(y) = y\int_1^y \frac{\Psi_\chi(t)}{t^2}\,dt, \qquad S'_\chi(y) = \sum_{n\le y}\frac{\Lambda(n)\chi(n)}{n} \quad (y \notin \mathbb{N}). \tag{2.2}$$

To prove the first identity, interchange the finite sum with the integral and use $y\int_n^y t^{-2}\,dt = y/n - 1$. The second follows by differentiating on each interval between consecutive integers. At an integer $n \ge 2$, the newly entering summand has value zero, which proves continuity; the one-sided derivatives differ by $\Lambda(n)\chi(n)/n$. Thus the prime coefficients are still completely present in the smoothed function. The smoothing replaces the jumps of Ψ_χ by continuous ramps and makes the Mellin transform particularly simple.

The same identity can be read on the logarithmic scale. If $y = e^u$, then $S_\chi(e^u) = \sum_{\log n\le u}\Lambda(n)\chi(n)(e^{u-\log n} - 1)$. Each prime power contributes a translate of the single kernel $e^v - 1$ for $v \ge 0$. Its Laplace transform is $1/[s(s-1)]$ in $\operatorname{Re} s > 1$, by integrating $e^{-(s-1)v} - e^{-sv}$. This explains the denominator in Proposition 2.1 before any continuation is used. The linear regularizations in (2.1) then remove the singularity associated with $s = 1$. Their constants are determined by that cancellation; changing the constant multiplying y would reintroduce a real pole at 1.

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*

$$F_\chi(s) = -\frac{L'(s,\chi)/L(s,\chi) - s\ell_\chi}{s(s-1)} \qquad (\chi \neq 1), \tag{2.3}$$

$$F_1(s) = -\frac{\zeta'(s)/\zeta(s) + (s-1)^{-1} + 1 - (1+\gamma_E)s}{s(s-1)}. \tag{2.4}$$

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 (2.3).

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 (2.4). The terms 1 and $(1+\gamma_E)s$ in its numerator must both be retained.

For $\chi \neq 1$, the numerator in (2.3) 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 (2.4). These are removable singularities. At a zero $\rho \neq 0, 1$, write $L(s,\chi) = (s-\rho)^m u(s)$ with u holomorphic and $u(\rho) \neq 0$. Then $L'/L = m/(s-\rho) + u'/u$, giving the asserted residue. The pole is simple even when the zero has higher multiplicity. □

There are two domains in this calculation. In $\operatorname{Re} s > 1$ the original Mellin integral is absolutely convergent and equals the displayed formula. In $\operatorname{Re} s > 0$ the formula gives its meromorphic continuation. Regularity of that continuation at a point does not by itself assert convergence of the original improper integral there. The window argument keeps this distinction explicit: every differentiated integral is evaluated on a line $\operatorname{Re} s = \sigma > 2$, where the elementary growth bound suffices. Only the subsequent Cauchy estimates use continuation toward the critical strip. There is consequently no need at this stage to shift an integration contour across infinitely many zeros, sum an explicit formula over all zeros, or interchange such a sum with the moving window. The local principal part of one pole and a holomorphic remainder will supply all the required information.

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

$$\Phi(s) = \int_0^\infty g(u)e^{-su}\,du = \frac{1}{2}\sum_\chi a_\chi(F_\chi(s) + F_{\bar\chi}(s)). \tag{2.5}$$

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 (2.5) has residue

$$\operatorname{Res}_{s=\rho}\Phi(s) = -\frac{b_\rho}{\rho(\rho-1)}, \qquad b_\rho = \frac{1}{2}\sum_\chi a_\chi(m_\chi(\rho) + m_{\bar\chi}(\rho)) > 0 \tag{2.6}$$

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.

For example, if two represented characters have a common zero ρ of multiplicities m_1, m_2, their principal parts are positive real multiples of the same complex number $-1/[\rho(\rho-1)]$. Their sum has multiplicity weight $a_1m_1 + a_2m_2 > 0$, regardless of the argument of that complex number. A proof that considered only the real part of an individual residue could lose this information. Taking the real part of the function instead gives the conjugation formula (2.5), which remains a holomorphic identity in the transform variable. Explicitly, $\overline{G_\chi(y)} = G_{\bar\chi}(y)$ for real y, since the finite prime sums and the logarithmic derivatives have this conjugation property. Consequently the transform of $\operatorname{Re} G_\chi$ is $(F_\chi + F_{\bar\chi})/2$, rather than the real part of $F_\chi(s)$ at a complex argument.

The finite family and its weights are fixed while the windows move to infinity. This gives one function g and one meromorphic function Φ to which pole localization can be applied. In the arithmetic construction the auxiliary field and its discarded characters may change with the scale; the retained expression H_ψ does not. This separation is what allows the analytic conclusion to concern a zero of the original fixed L-function.

We isolate the localization argument as a general statement. The poles of a meromorphic function on a domain form a closed discrete subset, so every compact subset of the domain meets only finitely many of them; we call such a set *locally finite*.

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

$$\int_U^{(1+\delta)U} R_U(u) e^{-\theta u}\, du \le e^{\eta U}. \tag{2.7}$$

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 [21]; 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.

Two elementary models explain the hypotheses and the use of whole windows. For $0 < \theta < \beta < 1$, the function $g(u) = e^{\beta u}$ has transform $(s-\beta)^{-1}$. Its positive pole records a positive exponential term, and an upper bound of size $e^{\theta u}$ eventually fails at every point. With the opposite sign, $g(u) = -e^{\beta u}$ satisfies that upper bound everywhere while its transform has a negative residue at β. Thus the sign assumption at real poles is essential to a theorem based on one-sided information.

A nonreal pair behaves differently. For $\gamma \neq 0$, take $g(u) = 2e^{\beta u}\cos(\gamma u)$. Direct integration gives

$$\Phi(s) = \frac{1}{s-\beta-i\gamma} + \frac{1}{s-\beta+i\gamma}, \qquad A_k(\sigma) = 2r^{-k-1}\cos((k+1)\alpha), \tag{2.8}$$

where $\sigma - \beta - i\gamma = re^{i\alpha}$ and $A_k = (-1)^k \Phi^{(k)}/k!$. At arbitrarily large points $u = (2j+1)\pi/|\gamma|$ the function is negative, so bounds on a sequence of individual values cannot exclude these poles. Every interval of length at least $2\pi/|\gamma|$ contains a point $u = 2j\pi/|\gamma|$ where the function is $2e^{\beta u}$. Hence an entire sufficiently distant window of length δU cannot satisfy the smaller pointwise upper bound. The derivative formula also shows why a fixed real center can be inconvenient: the cosine can cancel the two equal-distance pole contributions. Moving the center vertically separates their distances. The proof below performs that separation for an arbitrary locally finite pole set, without assuming that there is a globally rightmost zero or a gap between the real parts of all zeros.

The same model distinguishes a power interval from an interval of fixed multiplicative length in the original variable. Set $\gamma = 1$, fix $0 < h < \pi/2$, and put $U_j = (2j+1)\pi - h$. On $[U_j, U_j + 2h]$ one has $\cos u \le -\cos h < 0$, so the upper bound $g(u) \le e^{\theta u}$ holds throughout every one of these intervals, although the two poles lie to the right of θ. In the variable $y = e^u$ these are the intervals $[e^{U_j}, e^{2h}e^{U_j}]$. Their logarithmic widths remain bounded. By contrast, $[Y, Y^J]$ with a fixed $J > 1$ becomes $[\log Y, J\log Y]$, exactly the window in Theorem 2.2 with $\delta = J - 1$. Its logarithmic width grows, and the concentrating derivative can fit inside it. This is why uniform control on a whole power interval is a substantive output of the arithmetic construction.

The concentration mechanism is the following. For $a > 0$ and an integer $k \ge 0$, the function $a^{k+1}u^k e^{-au}/k!$ is a probability density on the positive half-line, with mean $(k+1)/a$ and standard deviation $\sqrt{k+1}/a$. A derivative of order k proportional to U therefore places its mass near a point proportional to U, with a width proportional to $\sqrt{U}$, and a window of length δU contains almost all of this mass once the center lies suitably inside it. Exponential tail estimates, rather than the variance alone, are needed, because a pole contributes a geometric factor in the derivative order. There is also a mismatch of parameters to absorb: the uncontrolled part of g can grow like e^{2u}, so the tail kernel has parameter $\sigma - 2$, whereas comparison with a pole uses the center σ. The following lemma makes these two parameters compatible and gives the precise geometric bound required later.

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

$$\int_{(0,\infty)\setminus[U,(1+\delta)U]} \frac{u^k e^{-(\sigma-2)u}}{k!}\, du \leq \frac{C_\delta}{\sigma^k}. \tag{2.9}$$

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

$$\int_{(0,\infty)\setminus[U,V]} \frac{u^k e^{-su}}{k!}\, du \leq \frac{e^{-\ell V}}{(s-\ell)^{k+1}} + \frac{e^{\ell U}}{(s+\ell)^{k+1}}. \tag{2.10}$$

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 $k \geq U$ 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 $a > 0$, $\log(\sigma/a) \leq (\sigma - a)/a$. Hence $e^x a^{-k-1} \leq e^c/(a\sigma^k)$ whenever $x + k(\sigma-a)/a \leq c$. For the first term of (2.10), with $V = (1+\delta)U$, one has $(\sigma - a_-)/a_- = 1/(M-1)$ and $b\sigma = M(M-1)$. Since $k \leq b\sigma U + 1$, it follows that $-\ell V + k/(M-1) \leq 1$. For the second term, $(\sigma - a_+)/a_+ = -(\ell-2)/a_+$. The inequality $\ell a_+ \leq (\ell+1)\sigma$, together with $k \geq b\sigma U$, gives $\ell U - k(\ell-2)/a_+ \leq 0$. Applying these two bounds to (2.10) proves (2.9) with $C_\delta = e/a_- + 1/a_+$. □

For $k \geq 0$ and $\operatorname{Re} z > 2$, differentiation under the integral in (2.5) is justified by the integrable majorant $C_0 u^k e^{-(\operatorname{Re} z - 2)u}$. Set

$$A_k(z) = \frac{(-1)^k}{k!}\Phi^{(k)}(z) = \int_0^\infty g(u)\frac{u^k e^{-zu}}{k!}\, du. \tag{2.11}$$

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

$$|A_k(\sigma + it)| + \operatorname{Re} A_k(\sigma) \leq 2\int_0^\infty \max(g(u), 0)\frac{u^k e^{-\sigma u}}{k!}\, du. \tag{2.12}$$

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

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

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

$$|A_k(\sigma + it)| + \operatorname{Re} A_k(\sigma) \leq \frac{2e^{\eta U}}{(\sigma-\theta)^k} + \frac{2C_0 C_\delta}{\sigma^k}. \tag{2.13}$$

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 $\operatorname{Re} s \leq 1$. 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 $0 < D \leq \sigma - \theta$ 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 $r \leq \sigma - \theta$ 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 $D = |\sigma - \rho| \geq \sigma - \operatorname{Re}\rho \geq D$. 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

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

The disk $B(\sigma, D)$ is pole-free, so, after possibly decreasing r while keeping $r > d$, the same estimate gives $|A_k(\sigma)| \leq M' r^{-k}$. Consequently the left side of (2.13) is at least $|c| d^{-k-1} - (M + M') r^{-k}$.

Choose $\eta > 0$ with $e^{\eta} d/E < 1$. Multiply (2.13) by d^{k+1} and let admissible U tend to infinity. Since $k \geq U$, 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 $d < r' < r$, $\operatorname{Re} A_k(\sigma) \geq c_\beta d^{-k-1} - M (r')^{-k}$. The left side of (2.13) 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 $g(u) \leq e^{\theta u}$ 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 (2.7) is δU, and $\delta U \leq e^{\eta U}$ for all sufficiently large U, for each fixed $\eta > 0$. □

The conclusion is stable under the fixed constants and logarithmic factors that often remain after an arithmetic estimate. The following version is occasionally more convenient than absorbing each such factor into a small power at an earlier stage.

Corollary 2.6. *Assume the analytic hypotheses of Theorem 2.2. Fix $C, D, A \geq 0$ and $\kappa < \theta$. Suppose that on arbitrarily distant windows $[U, (1+\delta)U]$ one has $g(u) \leq C e^{\theta u}(1+u)^A + D e^{\kappa u}$. Then Φ has no pole in $\operatorname{Re} s > \theta$.*

Proof. Take the nonnegative right side as R_U. Its weighted integral in (2.7) is at most $C\delta U(1 + (1+\delta)U)^A + D e^{-(\theta-\kappa)U}/(\theta - \kappa)$. For every fixed $\eta > 0$ this is at most $e^{\eta U}$ when U is sufficiently large: if the first term is nonzero its logarithm is $O(\log U)$, and the second tends to zero. Theorem 2.2 applies. □

In the original variable this permits an upper bound $G(y) \le Cy^{\theta}(1+\log y)^{A} + Dy^{\kappa}$ on arbitrarily distant power intervals $[e^{U}, e^{(1+\delta)U}]$. The constants and the relative logarithmic width δ are fixed before those intervals are selected. The result asserts exclusion in the open half-plane $\operatorname{Re} s > \theta$; a pole on its boundary is compatible with the bound. For instance $g(u) = e^{\theta u}$ has exactly that boundary pole and satisfies the hypothesis with equality. This example also explains why the arithmetic exponent is allowed an arbitrarily small positive excess over $1/2$: each fixed excess excludes an open half-plane, and the intersection of those conclusions gives the desired boundary.

Apply the corollary to $g(u) = H_{\psi}(e^{u})$. All coefficients in (1.2) are positive, so (2.6) 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 of this fixed 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. This argument does not take a limit of fields, characters, or meromorphic functions as ε tends to zero. It assumes a hypothetical zero and then chooses one sufficiently small ε for which the window theorem already excludes it.

For a quadratic primitive character the retained expression is especially simple: $o = 2$, so $H_{\psi} = G_1 + G_{\psi}/7$ and its transform is $F_1 + F_{\psi}/7$. At a common zero of multiplicities m_1, m_{ψ}, the residue is $-(m_1 + m_{\psi}/7)/[\rho(\rho-1)]$, which cannot vanish. For a primitive character of order three, the two nontrivial powers are conjugate. Hence $H_{\psi} = G_1 + (1/5)\operatorname{Re} G_{\psi}$, while its transform is $F_1 + (F_{\psi} + F_{\bar\psi})/10$. Its residue has numerator $m_1 + (m_{\psi} + m_{\bar\psi})/10$. These identities follow by substituting the powers into (1.2) and using conjugation; they illustrate how the real expression still records a zero of either complex character. In general the finite quotient calculation in Lemma 6.1 gives precisely the common coefficient $1/(3o+1)$, and (2.6) gives the same noncancellation argument for every order. The coefficient need not be large. Its strict positivity is enough, because in the derivative comparison any fixed nonzero residue eventually dominates the holomorphic remainder.

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

$$L(s,\chi) = L(s,\chi^{*}) \prod_{\substack{p \mid N \\ p \nmid f}} (1 - \chi^{*}(p)p^{-s}). \tag{2.14}$$

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 (2.14) 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. 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

$$\widehat{\deg}_{\mathrm{Ar}}(E) = \frac{1}{d}\left(\sum_{\mathfrak{p}} \mathrm{ord}_{\mathfrak{p}}(v;\det E)\log N\mathfrak{p} - \sum_{\sigma:K\hookrightarrow\mathbb{C}} \log\|v\|_{\sigma}\right). \tag{3.1}$$

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 (3.1) 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

$$\mu(E) = 2\mu_{\mathrm{Ar}}(E), \qquad h(B) = \frac{2}{d}\sum_{\sigma:K\hookrightarrow\mathbb{C}} \log\|B_{\sigma}\|_{\mathrm{op}} \quad (B \neq 0). \tag{3.2}$$

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

$$\mathcal{H}_{\mathrm{cov}}(E) = 2^{r_2 r}|\Delta_K|^{-r/2}\,\mathrm{covol}_{\mathbb{R}}(E), \qquad \mu(E) = -\frac{2}{dr}\log \mathcal{H}_{\mathrm{cov}}(E). \tag{3.3}$$

The first equality removes the covolume of $\mathcal{O}_K$ in the Minkowski space, and the second agrees with (3.1)–(3.2). 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 $\mathcal{H}_{\mathrm{cov}}$ is used. The affine discriminant term appearing in the source's conversion from raw euclidean slopes has already been absorbed in (3.3).

In the totally complex case, the determinant calculation behind this normalization is useful when the coefficient field or an integral model changes. Choose a pseudobasis $E = I_1 e_1 \oplus \cdots \oplus I_r e_r$, where the I_j are fractional ideals and the e_j form a K-basis. Let G_{σ} be the Gram matrix of this basis. Choosing one embedding from each complex conjugate pair, the real covolume is

$$\mathrm{covol}_{\mathbb{R}}(E) = 2^{-r_2 r}|\Delta_K|^{r/2}\left(\prod_j N(I_j)\right)\prod_{\sigma \bmod \mathrm{conj}} \det G_{\sigma}. \tag{3.4}$$

Indeed, the covolume of an ideal in the usual complex Minkowski space is $2^{-r_2}|\Delta_K|^{1/2}N(I)$, initially for integral ideals by their index in $\mathcal{O}_K$ and then for fractional ideals by scaling. The direct sum multiplies these covolumes. At a complex place a linear change of coordinates has real determinant equal to the squared absolute value of its complex determinant, which is precisely the determinant of the associated Gram matrix. This proves (3.4). On the other hand, the finite order of $e_1 \wedge \cdots \wedge e_r$ is the negative of the order of $\prod_j I_j$, and its norm is $(\det G_{\sigma})^{1/2}$. Substitution in (3.1) proves (3.3) with every factor of two accounted for. A determinant norm is a complex volume; its square is the corresponding real volume at a complex place.

The same calculation gives a useful invariance of the eventual comparison.

Lemma 3.1. *Suppose the norm of E_{σ} is multiplied by $e^{a_{\sigma}}$ and the norm of F_{σ} by $e^{b_{\sigma}}$, with real numbers compatible with complex conjugation. Write primes for quantities computed with the*

new metrics. For any nonzero map $B : E_K \to F_K$,

$$\mu'(E) = \mu(E) - \frac{2}{d}\sum_\sigma a_\sigma, \qquad h'(B) = h(B) + \frac{2}{d}\sum_\sigma (b_\sigma - a_\sigma), \qquad \mu'(F) = \mu(F) - \frac{2}{d}\sum_\sigma b_\sigma.$$

Consequently $\mu(E) - h(B) - \mu(F)$ is unchanged. The maximum and minimum slopes of either lattice change by the same amount as its mean slope.

Proof. The determinant norm of a rank-r subspace of E_σ is multiplied by e^{ra_σ}. The finite lattice is unchanged, so the degree formula gives the first assertion and gives the same slope shift for every sublattice and quotient. The corresponding assertions for F follow in the same way. For each nonzero vector v the ratio $\|Bv\|'_\sigma/\|v\|'_\sigma$ is $e^{b_\sigma - a_\sigma}\|Bv\|_\sigma/\|v\|_\sigma$; taking the supremum gives the height formula. The three changes cancel in the stated difference. Taking extrema over sublattices and quotients proves the assertion about extremal slopes. □

This invariance explains why a common period normalization can be moved between the source, the target, and the displayed height, provided all three changes are made together. It does not permit an independent change of the norm of one term in an inequality. In particular, the auxiliary Hilbert metrics used for a frame at one embedding will not replace the arithmetic metrics used to average map heights.

A common tensor twist has an analogous effect at finite and infinite places simultaneously. If L is a hermitian line lattice, equip $E \otimes L$ with the tensor metric. Then

$$\mu(E \otimes L) = \mu(E) + \mu(L), \qquad \mu_{\max,\min}(E \otimes L) = \mu_{\max,\min}(E) + \mu(L). \tag{3.5}$$

For the mean, the determinant identity $\det(E \otimes L) = (\det E) \otimes L^{\otimes r}$ proves the formula by degree additivity for lines. This identity can also be checked locally: tensoring by a local generator of L multiplies all r basis vectors by that generator, and its determinant metric is multiplied by its norm to the power r. To obtain the extremal statements, tensoring with the invertible module L identifies the saturated sublattices and torsion-free quotients on the two sides. Their degrees change by their rank times the degree of L, so their slopes all undergo the same translation. Taking the maximum or minimum proves the formulas.

For a nonzero map $B : E \to F$, the map $B \otimes I_L : E \otimes L \to F \otimes L$ has exactly the same operator norm at every embedding: after choosing any nonzero vector of the one-dimensional space L_σ, its norm factor cancels between source and target. Its height is therefore unchanged, and integrality is preserved by tensoring. The comparison $\mu(E) - h(B) - \mu(F)$ is again unchanged. This observation applies to a fractional-ideal line as well as to a line obtained by rescaling an archimedean metric, so it treats finite and infinite normalizations within one determinant calculation.

For a bilinear pairing $N \times R \to \mathcal{O}_K$, the compatible twist is $N \otimes L^\vee$ on the column and $R \otimes L$ on the row. Contraction of $L^\vee \otimes L$ gives the new pairing, and its operator norm equals the old one because the evaluation pairing on a hermitian line and its dual has norm one. Viewed as a map into the row dual, both its source and target have been tensored by $L^\vee$. Thus both slopes decrease by $\mu(L)$ and the map height is unchanged. This is a concrete model for a common contribution that cancels in a slope comparison. To use such a cancellation in an actual tower, the same line and the compatible metrics must occur on both sides. An unrelated finite index or a scalar attached to only one column remains a real contribution and has to be estimated. The explicit common term in Section 4 is retained until the corresponding compatibility identity has been applied.

A rank-one example also fixes the direction of the finite-place gain. Over $\mathbb{Q}$, take $E = a\mathbb{Z}$ and $F = c\mathbb{Z}$ with their usual absolute-value metrics, where $a, c > 0$ are rational. Multiplication by a nonzero rational number b is integral from E to F precisely when ba/c is a nonzero integer. The slopes are $-2\log a$ and $-2\log c$, and the map height is $2\log|b|$. Thus $\mu(E) - h(B) - \mu(F) = -2\log|ba/c| \le 0$. Replacing $a\mathbb{Z}$ by a larger lattice raises its slope, but integrality of the same map must then be rechecked. The local jet calculation later in the paper provides exactly the divisibility needed for that replacement. The example also shows why negative map heights are

natural: an integral map between differently scaled lattices can have operator norm less than one.

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 $\lambda_1(E) \geq \cdots \geq \lambda_r(E)$ with $r^{-1}\sum_j \lambda_j(E) = \mu(E)$. The filtration subspace $E_{\geq t} \subset E \otimes K$ is the sum of the successive pieces with slopes at least t. Thus $\mu_{\min}(E_{\geq t}) \geq t$ whenever it is nonzero, whereas the nonzero quotient $E/E_{\geq t}$ has maximum slope strictly less than t. These statements, including attainment and the canonical filtration, belong to arithmetic reduction theory [15, 7]. We will use only these explicitly stated properties.

Lemma 3.2. *Let $B : E \to F$ be a nonzero integral map of hermitian lattices. Then $\mu_{\min}(E) \leq \mu_{\max}(F) + h(B)$. More precisely,*

$$B(E_{\geq t}) \subset F_{\geq t-h(B)} \qquad (t \in \mathbb{R}). \tag{3.6}$$

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 $\|B_\sigma\|_{\mathrm{op}}^{\mathrm{rk}\, I}$, when the domain has the quotient norm. Formula (3.1) therefore gives $\mu(E/\ker B) \leq \mu(I) + h(B)$. Saturating I in F can only increase its degree, so $\mu(I) \leq \mu_{\max}(F)$. Since $\mu_{\min}(E) \leq \mu(E/\ker B)$, the first assertion follows.

For the filtered assertion, compose the restriction of B to $E_{\geq t}$ with the quotient map $F \to F/F_{\geq t-h(B)}$. 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 (3.6). □

A diagonal lattice provides a concrete model for the filtration. Let $E = \mathbb{Z}^r$ with squared norm $\sum_{j=1}^r e^{-a_j}|x_j|^2$, where $a_1 \geq \cdots \geq a_r$ are real. Its Harder–Narasimhan slopes in the doubled convention are exactly $a_1, \ldots, a_r$. To verify this, let $S \subset E$ have rank q and choose an integral basis of S, written as the columns of an integer matrix V. Cauchy–Binet gives

$$\det(V^{\mathsf{t}} \operatorname{diag}(e^{-a_j})V) = \sum_{\substack{J \subset \{1,\ldots,r\} \\ |J|=q}} (\det V_J)^2 e^{-\sum_{j\in J} a_j} \geq e^{-\sum_{j=1}^q a_j}. \tag{3.7}$$

At least one minor is a nonzero integer, and every exponential in the sum is at least the one on the right. Thus the doubled degree of S is at most $\sum_{j=1}^q a_j$. The coordinate lattice on the first q basis vectors attains this bound for every q. The polygon of maximal degrees consequently has these successive slopes; grouping equal slopes gives the canonical filtration. The same argument shows directly that a diagonal lattice with all $a_j = a$ is semistable: every rank-q sublattice has mean slope at most a.

In this example the arithmetic filtration is visible as a coordinate flag, but its role in the general proof is identical. A threshold selects directions according to their arithmetic slopes, and a map height limits how far those directions can move down the target filtration. For instance, $a = (4, 0, -5)$ gives maximum slope 4, minimum slope -5, and mean slope $-1/3$. The maximum alone gives no lower bound for the mean. A determinant bound sees their sum, while the filtration records how that sum is distributed among directions. The frame argument uses this distribution because its sources can have a small low-slope part that must be removed before their images are compared with a target filtration piece.

The matrix calculation also indicates what an integral hypothesis contributes. Over $\mathbb{R}$, the chosen basis of a subspace could be rescaled arbitrarily, and its Gram determinant would give no arithmetic degree bound. The nonzero integral minor in (3.7) prevents that rescaling for a fixed sublattice. In a number field, the corresponding protection is the product formula and the finite determinant ideal. This is why an archimedean family of well-conditioned maps is not, by

itself, a slope theorem: integrality is required to relate its dimensions and operator norms to arithmetic degrees.

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 3.3 (`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*

$$\sum_i \|c_i B_i^* B_i - I_{V_i}\|_{\mathrm{HS}}^2 \le e_F R, \qquad \operatorname{Tr}(Q^2) \le (1+\eta_F)\frac{(\operatorname{Tr} Q)^2}{D}, \tag{3.8}$$

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

$$\operatorname{codim}_{\mathcal{H}} \sum_i B_i N_i \le (\eta_F + 4\varepsilon_F + 4\sqrt{e_F})D. \tag{3.9}$$

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, $|\operatorname{Tr}(P_iK_i)| \le \sqrt{\operatorname{rk} P_i}\,\|K_i\|_{\mathrm{HS}}$, 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

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

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 (3.9). □

The two assumptions in (3.8) 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.

Two elementary examples distinguish the hypotheses. Let the target be $\mathbb{C}^2$ and let both source spaces be $\mathbb{C}$, with $c_1 = c_2 = 1$. If $B_1 z = (z, 0)$ and $B_2 z = (0, z)$, then each source Gram operator is the identity and $Q = I_{\mathbb{C}^2}$. Both errors can be zero, and the images span the target. If instead $B_1 z = B_2 z = (z, 0)$, the two source Gram operators are still identities, but Q has eigenvalues 2 and 0. Now $\operatorname{Tr}(Q^2) = 4$ and $(\operatorname{Tr} Q)^2/D = 2$, so the second hypothesis requires $\eta_F \ge 1$. Perfect individual norm identities have supplied no information about the missing target direction. This is why the correlation estimate in Section 8 is a separate part of the proof.

The second hypothesis has a direct interpretation in terms of the number of directions detected by the maps. If a nonzero positive operator Q has r_Q positive eigenvalues, Cauchy–Schwarz on these eigenvalues gives $(\operatorname{Tr} Q)^2 \le r_Q \operatorname{Tr}(Q^2)$. Consequently $r_Q \ge D/(1+\eta_F)$ under (3.8). The quotient $(\operatorname{Tr} Q)^2/\operatorname{Tr}(Q^2)$ therefore measures a lower bound for the dimension occupied by the frame. Proposition 3.3 strengthens this observation by allowing parts of the sources to be removed. It controls how much trace those parts could carry, using the first hypothesis, and then applies the same dimension argument to the remaining trace. There is no need for a separate operator-norm bound on every source error. A squared Hilbert–Schmidt bound summed over the sources is sufficient because the removed dimensions are summed with the same rank weights.

The logarithm in the slope conclusion comes from a distribution estimate. We record the elementary calculation separately, since it is responsible for preserving the scale of the arithmetic gain.

Lemma 3.4. *Let R, D be positive integers. Let $a_1, \dots, a_R$ be real numbers with mean at least s and $a_j \le s+u$, where $u \ge 0$. For $t > 0$, let $q(t)$ be the fraction of the a_j that are less than $s-t$. Then $q(t) \le u/(u+t)$ if $u > 0$, and $q(t) = 0$ if $u = 0$. If $b_1, \dots, b_D \ge s - H$ with $H \ge 0$, and if their fraction below $s-t$ is at most $A + Bu/(u+t)$ for $0 < t \le H$, where $A, B \ge 0$, then*

$$\frac{1}{D}\sum_j b_j \ge s - AH - Bu\log(1+H/u).$$

At $u = 0$, the last term is interpreted as zero and the hypothesis on the target fraction is interpreted as the bound A.

Proof. The mean of the a_j is at most $q(t)(s-t) + (1-q(t))(s+u) = s + u - q(t)(u+t)$, which proves the first assertion. If $u = 0$, every $a_j \le s$ and the mean is at least s, so all the numbers equal s. For the second assertion, the length of $\{t \in (0,H) : b_j < s - t\}$ is $(s-b_j)_+$. Averaging these lengths gives the exact identity $D^{-1}\sum_j \min(b_j, s) = s - \int_0^H q_b(t)\,dt$, where q_b is the target fraction. Insert its stated bound and integrate $u/(u+t)$. Since $b_j \ge \min(b_j, s)$, this proves the inequality. The same identity gives the assertion when $u = 0$ or $H = 0$. □

The reciprocal decay $u/(u+t)$ cannot be improved using only the two source hypotheses. For a distribution supported at $s-t$ and $s+u$ with mean s, the mass at $s-t$ is exactly $u/(u+t)$. Replacing the lower value by $s - t - \varepsilon$ gives the corresponding strict-tail example, whose mass tends to that value as $\varepsilon \downarrow 0$. Rational masses can be realized by repeating entries in a finite list. The additional arithmetic information enters through the transfer of this source distribution to the target. A bound for the mean alone would allow a very small fraction of extremely large source slopes to compensate for a large low-slope part; the common upper bound $s+u$ rules this out.

Theorem 3.5 (`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 (3.8), and suppose

$$\operatorname{rk} F \ge (1-\delta_F)\dim\widetilde{\mathcal{H}}, \qquad 0 \le \delta_F < 1. \tag{3.10}$$

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

$$\sum_i p_i(\mu(E_i) - h_i) \ge s, \qquad \mu_{\max}(E_i) - h_i \le s+u, \qquad \mu_{\min}(F) \ge s - H_0. \tag{3.11}$$

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

$$\mu(F) \ge s - \frac{4u\log(1+H_0/u) + (\eta_F + 4\sqrt{e_F})H_0}{1-\delta_F}. \tag{3.12}$$

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 $s+u$. By the first part of Lemma 3.4, the fraction $q(\tau)$ of these slopes that are less than $s-\tau$ satisfies $q(\tau) \le u/(u+\tau)$ for every $\tau > 0$, and $q(\tau) = 0$ if $u = 0$.

Set $N_i = (E_i)_{\ge s-\tau+h_i} \otimes \mathbb{C}$. Lemma 3.2 shows that B_iN_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 3.3 bounds the codimension of $\sum_i \widetilde{B}_iN_i$ in $\widetilde{\mathcal{H}}$. Passing to the quotient $F \otimes \mathbb{C}$ cannot increase that codimension. Using (3.10), we find that the fraction of target slopes below $s-\tau$ is at most

$$\frac{1}{1-\delta_F}\left(\eta_F + 4\sqrt{e_F} + \frac{4u}{u+\tau}\right). \tag{3.13}$$

The auxiliary Hilbert structures at the chosen embedding enter only this dimension calculation. The map heights in (3.11) still involve all embeddings and the original arithmetic metrics.

Every target slope is at least $s - H_0$, and by (3.13) the fraction of target slopes below $s - \tau$ is at most $A + Bu/(u + \tau)$ for $0 < \tau \le H_0$, where $A = (\eta_F + 4\sqrt{e_F})/(1 - \delta_F)$ and $B = 4/(1 - \delta_F)$. The second part of Lemma 3.4, applied to the list of target slopes with $H = H_0$, gives (3.12), including the conventions at $u = 0$ and at $H_0 = 0$. □

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 (3.12). 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 3.6. *If $u, U, H_0, b, X \ge 0$, $u \le U$, and $H_0 \le Xb$, then*

$$u \log(1 + H_0/u) \le (U + b) \log(1 + X), \tag{3.14}$$

with the same convention at $u = 0$.

Proof. For $u > 0$, concavity of log at $1+X$ gives $\log(1+H_0/u) \le \log(1+X)+(H_0/u-X)/(1+X)$. After multiplication by u and use of $H_0 \le Xb$, the right side is at most $u \log(1+X)+Xb/(1+X)$. Since $X/(1 + X) \le \log(1 + X)$ and $u \le U$, this proves (3.14). 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 $A, B, c > 0$, $x^A e^{-x^c} \le x^{-B}$ 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.

For example, suppose $\delta_F \le 1/2$, $u \le C_1 b$ and $H_0 \le x^A b$, with $x \ge 2$, $b > 0$, and fixed $C_1, A \ge 0$. Lemma 3.6 bounds the logarithmic part of (3.12) by $8(C_1 + 1)b \log(1 + x^A)$. Thus even a very crude polynomial lower bound on the target minimum slope costs only $O(b \log x)$ in this part of the argument. The frame term behaves differently. If $\eta_F + 4\sqrt{e_F} \le x^{-A-C}$ for a fixed $C > 0$, its contribution is at most $2x^{-C}b$. The error requirement depends on the polynomial exponent A in the target slope bound. Choosing an accuracy merely because it tends to zero, without comparing it with that exponent, would not suffice.

There is also a useful refinement when the minimum-slope estimate is very weak. The target fraction is always at most one, so the proof actually permits integrating the minimum of one and the right side of (3.13). We use the simpler displayed bound because the chosen parameters already make its error acceptable. Nothing in the argument requires estimating the distribution of target slopes individually or constructing a basis of short target vectors. The frame controls dimensions of filtration pieces, and the integral of those dimensions determines the mean. This is the point at which a large collection of modestly controlled maps can yield a useful arithmetic estimate that no single map supplies.

Finally, the number-field embedding used to test the frame has a limited and precise role. A K-subspace and its scalar extension at any embedding have the same dimension. Once a high-slope source subspace is chosen arithmetically, its dimension and the dimension of its image can therefore be tested at one embedding using convenient Hilbert coordinates. The filtration itself and the height inequality remain arithmetic statements involving all embeddings. This separation allows the spectral calculation to use its natural unitary models without silently changing the arithmetic degree.

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.

4. The arithmetic comparison

The arithmetic construction is designed to meet a fixed list of inequalities. We state their numerical consequence before constructing the objects that satisfy them. This separates the comparison itself from the existence and uniformity questions that occupy the later sections. The comparison 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

$$\mathcal{M}(a, w) = a \left(\log a - 1 - \gamma_E + \operatorname{Re} \ell_\eta - \frac{1}{f} \sum_{\rho \in \Phi} \log w(\rho) \right). \tag{4.1}$$

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 6 expresses $\mathcal{A} - \mathcal{M}$ as a linear combination of the regularized prime sums G_χ.

To interpret this formula, first ignore the indices j, t. The finite quantity $\mathcal{A}$ measures the logarithmic divisibility forced by Frobenius orbits. The expression $a \log a - a$ is the leading archimedean cost of a coefficient of order a. The remaining terms record the period and the metric at infinity. The comparison will bound the difference between the finite gain and this archimedean cost. Neither term separately has square-root size. It is their difference, after the exact normalizations have been matched, that is recognizable as a regularized prime sum. This is why estimating each displayed main term by its absolute value would destroy the argument.

The period normalization has the form

$$\log \Omega = \gamma_E + \log(\pi/v) - \operatorname{Re} \ell_\eta, \qquad v > 0. \tag{4.2}$$

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 [28]. The definition of the period and its relative-different normalization are part of the theta input. Only (4.2), 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 4.1 (`key_identity`). *Suppose $g_0, v, \Omega, a_t, w_j(\rho), T_i$ are positive and (4.2) 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*

$$\mu_i \geq -M + G + F - E_i - L_i, \tag{4.3}$$

$$\sum_i p_i(\mu_i - h_i) \leq \mu_T + E_{\mathrm{HN}}, \tag{4.4}$$

$$\mu_T + h_i - F \leq m \sum_t a_t(\log a_t - 1 - \log \Omega) - A \log T_i + E_D, \tag{4.5}$$

$$E_i \leq mA \log g_0 + E_C, \tag{4.6}$$

$$\log T_i \geq \sum_{j=1}^{m} \frac{1}{f} \sum_{\rho \in \Phi} \log(w_j(\rho) v/\pi) + m \log g_0 - E_T, \tag{4.7}$$

$$\sum_{j=1}^{m}\sum_{t}\mathcal{A}(a_t,t,w_j)\le G+E_P. \tag{4.8}$$

Then

$$\sum_{j=1}^{m}\sum_{t}(\mathcal{A}(a_t,t,w_j)-\mathcal{M}(a_t,w_j))\le M+E_{\mathrm{HN}}+\sum_i p_iL_i+E_D+E_C+AE_T+E_P. \tag{4.9}$$

Proof. Average (4.3) and combine it with (4.4). Rearranging gives $G\le M+E_{\mathrm{HN}}+\sum_i p_i(E_i+L_i+\mu_T+h_i-F)$. Use (4.5) and (4.6) on the right. The result is

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

The common finite twisting term F has already disappeared exactly; it has not been bounded in absolute value. Next apply (4.7). 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 (4.2) 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 (4.8) and move the main expression to the left. This proves (4.9). □

In the arithmetic application, g_0 stands for the product of the two level primes, which the source writes g_0g_1; 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 (4.6) 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 3 supplies (4.4); the local and archimedean theta estimates supply (4.3), (4.5), and (4.6); and the size of the word family supplies (4.7). 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`.

A remark on quantifiers is in order. Proposition 4.1 is formulated for one fixed thick translate, and the source proves the resulting bound for every translate separately; this is the pointwise unit bound used in `ChTop.lean`. The Fourier extraction of Section 6 then averages over the translates $u\in A_0$ with equal weights, so mathematically it consumes only the uniform mean of these bounds over A_0. What it does require is that the weights at the different translates form one compatible family with error bounds uniform in u, because those errors are estimated before the average is taken; see the discussion following Proposition 6.4. A bound for a mean with other weights, or over another index set, would not suffice.

We can now state the comparison in a form in which the frame hypothesis, rather than an unexplained slope inequality, is the input. This is a useful stopping point for a reader who wants to understand what the geometric construction must accomplish. The following corollary is an explicit combination of Theorem 3.5 and Proposition 4.1; it makes no existence assertion for the arithmetic objects.

Corollary 4.2 (Comparison for fixed arithmetic data)**.** *Let I be finite and nonempty, and let $\mathcal{E}_i$ ($i\in I$) and $\mathcal{F}$ be nonzero hermitian lattices, let $B_i:\mathcal{E}_i\to\mathcal{F}$ be integral maps with height bounds h_i for the nonzero maps, and let $p_i=\mathrm{rk}(\mathcal{E}_i)/\sum_j\mathrm{rk}(\mathcal{E}_j)$. Suppose their lifts to one complex Hilbert space satisfy the frame hypotheses and the retained-rank hypothesis of Theorem 3.5, with parameters e_F,η_F,δ_F. Write*

$$s=\sum_i p_i(\mu(\mathcal{E}_i)-h_i),\qquad u=\max_i(\mu_{\max}(\mathcal{E}_i)-h_i)-s,$$

and choose $H \ge \max\{0, s - \mu_{\min}(\mathcal{F})\}$. *Suppose the period identity* (4.2) *and inequalities* (4.3), (4.5)–(4.8) *hold with* $\mu_i = \mu(\mathcal{E}_i)$ *and* $\mu_T = \mu(\mathcal{F})$. *Then*

$$\sum_{j,t}(\mathcal{A} - \mathcal{M}) \le M + \sum_i p_i L_i + E_D + E_C + AE_T + E_P + \frac{4u\log(1 + H/u) + (\eta_F + 4\sqrt{e_F})H}{1 - \delta_F}. \tag{4.10}$$

The arguments of $\mathcal{A}$ *and* $\mathcal{M}$ *in the left side are those of* (4.9); *the first term in the numerator is understood to vanish when* $u = 0$.

Proof. Each mean slope is no larger than its maximum slope, so $s \le \max_i(\mu_{\max}(\mathcal{E}_i) - h_i)$ and $u \ge 0$. The definition of H gives the required lower bound for the minimum slope of $\mathcal{F}$. Theorem 3.5 therefore gives (4.4) with E_{HN} equal to the last term of (4.10). This quantity is nonnegative because $\delta_F < 1$, $\eta_F, e_F, H \ge 0$, and $u\log(1 + H/u) \ge 0$. Proposition 4.1 applies without any further change of normalization. □

There are three different measures of quality in this statement. The number u measures the spread of the shifted source slopes above their pooled mean. The numbers e_F and η_F measure the two frame defects. The number δ_F measures the proportion of the ambient target discarded when passing to the retained arithmetic target. Their appearances are different: u is accompanied by a logarithm of the target range, the frame defects are multiplied by that range, and the discarded proportion divides the entire transfer estimate. Thus an estimate that is adequate for one of these quantities need not be adequate for another. In particular, proving positive target rank does not give a useful bound for $1/(1 - \delta_F)$.

The geometric hypotheses may now be read from left to right. Integral maps permit the use of the arithmetic filtration. Their global heights shift the source slopes. A common target identifies the images whose span is to be measured. The collective frame bounds transfer the slope distribution to that target. Finally, the target covolume, the word count, and the period formula identify the main term of the transferred inequality. The theta construction has to supply all of these facts for the same maps. A separate family of maps with the correct norm formula would not suffice.

Here is a convenient quantitative consequence. Suppose $0 \le \delta_F \le \delta_* < 1$, $H \le x^c b$, and $u \le U$, with $c > 0$ fixed and $x, b > 0$. Lemma 3.6 gives

$$E_{\mathrm{HN}} \le \frac{4(U + b)\log(1 + x^c) + (\eta_F + 4\sqrt{e_F})x^c b}{1 - \delta_*}. \tag{4.11}$$

For $x \ge 2$, the logarithm is at most $c\log x + \log 2$. If, for some $d > 0$, one has $\eta_F \le x^{-c-d}$ and $e_F \le x^{-2c-2d}$, the second term in the numerator is at most $5bx^{-d}$. The square on the required bound for e_F comes from its square root in the frame estimate. This calculation explains why the trace estimates must be stated with enough quantitative room before the geometric parameters are fixed. It also shows why merely writing both defects as $o(1)$ does not justify the slope budget when H grows with x.

The loss in (4.11) is unchanged if every source and target slope and every main slope comparison are translated coherently by a common constant: u and the difference $s - \mu_{\min}(\mathcal{F})$ are unchanged. This elementary invariance is useful for checking conventions. By contrast, changing the arithmetic metric in only one part of the construction changes a map height or a period and must be recorded. The exact cancellation in Proposition 4.1 is a stronger check: it tests that the integral, metric, level, and period conventions fit each other, rather than only that the individual estimates have the right order of magnitude.

Let $x \ge 1$ be the scale and let $\varepsilon > 0$ be fixed. We write n for the order of the finite abelian host factor of Section 6, $r_T = |\mathcal{T}_1|$ 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 $a_E \le 2xb$.

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

$$\mathcal{B} = r_T a_E \left(1 + \frac{m^2}{k}\right) + m\left(2nmbx^{\varepsilon} + \frac{n_2 kbx^{\varepsilon}}{n(nx^{-\varepsilon})}\right). \tag{4.12}$$

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 3.6 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 4.3 (`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*

$$\begin{aligned} &r_T \geq \tfrac{1}{2} n x^{-\varepsilon/8}, \qquad x^{1/2-\varepsilon} \leq m \leq x^{1/2+\varepsilon/2}, \\ &xm \leq k \leq n^2 x^{1/2-\varepsilon}, \qquad \delta_2 = n_2/n \geq x^{-\varepsilon}. \end{aligned} \tag{4.13}$$

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

$$\frac{2E}{r_T m \delta_2 b} \leq 20 x^{1/2+(C+21/8)\varepsilon}. \tag{4.14}$$

Proof. Divide (4.12) 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 $n_2 k x^{2\varepsilon}/(r_T n^2)$. Thus

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

For the last term we used $n_2 \leq n$, the upper bound for k, and $n/r_T \leq 2x^{\varepsilon/8}$. Every comparison is valid for $x \geq 1$. Multiplication by $2x^{C\varepsilon}/\delta_2$ and the lower bound for δ_2 prove (4.14). □

The formal statement `proposition_size_bound` packages this calculation with the Fourier inequalities and the admissible parameter predicates. Proposition 4.3 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 (4.13) is both more transparent and less restrictive.

The balance underlying the first and third terms can be made explicit. After division by the total tower weight, a cost of order x incurred once contributes a term ax/m, whereas a cost that grows through the tower contributes a term dm. Since $ax/m + dm - 2\sqrt{adx} = (\sqrt{ax/m} - \sqrt{dm})^2 \geq 0$, the sum is smallest at $m = \sqrt{ax/d}$, where it equals $2\sqrt{adx}$. This is the scale $m \approx x^{1/2}$ to which the much more elaborate construction is directed. In the actual calculation the factors n/r_T and $1/\delta_2$ are not bounded by absolute constants; they cost small powers of x, so the allowed interval for m has a small exponent width around $x^{1/2}$. The parameter k must also satisfy two opposing inequalities. Its lower bound controls the contribution $a_E m/(kb)$, while its upper bound controls the term containing $n_2 k$; increasing k indefinitely improves one error and worsens the other. Proposition 4.3 keeps both restrictions visible.

The stretched-exponential factor b plays another role. It makes certain arithmetic defects small relative to the available dimension or weight, but it cancels from the normalized estimate. There is no deduction that a large value of b directly improves H_ψ. Indeed, replacing b by a larger number changes the admissible source weights and the sizes of several error terms at once. The useful assertion is that, after the construction has made these choices coherently, the

quotient in (4.14) has only a polynomial dependence on x. This is one reason to perform the division explicitly before collecting powers.

The exponent $13\varepsilon/8$ in the bound for $\mathcal{B}$ comes from $\varepsilon/8$ in n/r_T, $\varepsilon/2$ in the upper bound for m, and ε in the displayed Hodge error. The passage from $13/8$ to $21/8$ costs the reciprocal density $1/\delta_2$. This allocation is useful when checking an improved construction: a stronger density estimate or a sharper rank share could improve a coefficient of ε, but the analytic conclusion only requires that this coefficient remain fixed. Reducing the coefficient 9 in the final bound is therefore secondary to proving all the estimates with the correct uniformity.

The residual slack can be used without dividing it among artificial numerical constants. Fix one admissible family and its comparison constants, and put $\alpha = 1/2 + (C + 21/8)\varepsilon$, with $C \geq 0$. The remaining Fourier-rounding and Lipschitz errors have the form $O_{\psi,\varepsilon}(x^{\varepsilon}\log x) + O_{\psi,\varepsilon}(x^{\varepsilon-1})$. Since $\alpha - \varepsilon = 1/2 + (C + 13/8)\varepsilon > 0$, these errors, together with (4.14), are bounded by $K_{\psi,\varepsilon}x^{\alpha}$ for a fixed finite constant $K_{\psi,\varepsilon}$. Increasing the threshold until $x^{3\varepsilon/8} \geq \max(1, K_{\psi,\varepsilon})$ bounds their sum by $x^{1/2+(C+3)\varepsilon}$. The family is fixed before this threshold is chosen, so the estimate holds at every admissible scale beyond it.

It follows from the positive Fourier selector that

$$H_{\psi}(x) \leq \Sigma_{\mathrm{small}}(x) + x^{1/2+(C+3)\varepsilon}, \tag{4.15}$$

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. There are $3o$ discarded characters, namely the twists of the o retained characters by χ_{-q}, by χ_{-d}, and by their product, so the discarded sums satisfy $\Sigma_{\mathrm{small}}(x) \leq 3o\, x^{1/2+\varepsilon}$ for the fixed order o of ψ. Taking x still larger so that $1 + 3o \leq x^{\varepsilon}$ yields

$$H_{\psi}(x) \leq x^{1/2+9\varepsilon}. \tag{4.16}$$

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. Section 11 uses the family's arbitrarily distant windows to choose an entire interval beyond that threshold.

5. 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 [11] and Huxley [18]. 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.

The section delivers two statements, both consumed by the character extraction of Section 6. Proposition 5.1 supplies the pair q, d together with square-root bounds for the twisted characters on a whole power interval; these bound the discarded sums $\Sigma(x)$ in (6.11) and give the logarithmic-derivative bounds required there. Proposition 5.3 supplies a uniform relative bound at the exponentially large arguments produced by the thin weights, which is the hypothesis $|G_{\theta}(y)| \leq \vartheta y$

of Proposition 6.4. Neither statement involves the geometric construction, so this section can be read independently of Sections 7–10.

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 5.1 (`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

$$\frac{1}{4}(\log Y)^2 \log m_0 \le \log U, \qquad Y^{\varepsilon/6} \le \log U \le Y^{\varepsilon/5}(\log Y)^2. \tag{5.1}$$

The following three uniform inputs are sufficient. First, throughout (5.1), $|\mathcal{P}| \ge U^{9/10}$ 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

$$\left|\sum_{p \in \mathcal{P}} \left(\frac{p}{s}\right)\right| \le Y^{-B}|\mathcal{P}| \quad (1 < s \le Y^A,\ s \text{ odd and squarefree},\ (s, m_0) = 1,\ s \ne s_*). \tag{5.2}$$

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 $T = Q$ [18].

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 (5.1), $\log m/\log U \le 4/(\log Y)^2$, 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 (5.1): 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 $y \ge 3$, $Z \ge y$, and $|\kappa(n)| \le 1$, set

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

Assume $p > Z^2$ for every $p \in \mathcal{P}$. If $n, n' \le Z$, 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

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

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 (5.2) therefore yields

$$\frac{1}{|\mathcal{P}|}\sum_{p\in\mathcal{P}}|S_p(y)|^2 \le R^2+\sum_{\ell}A_\ell^2+2B_0R+2B_0^2+Y^{-B}\left(\sum_{n\le Z}c_y(n)\right)^2. \tag{5.3}$$

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

The shape of these mixed terms can be seen directly from the support of the von Mangoldt function. An odd power of a prime $\ell\nmid m_0$ has outside squarefree part ℓ, whereas an even prime power has outside part 1. Every power of a prime dividing m_0 also has outside part 1. Thus, for example, a pair ℓ^3, r^2 with $\ell, r\nmid m_0$ belongs to the outside class ℓ; two odd powers of distinct outside primes ℓ, r belong to class ℓr; and two odd powers of the same outside prime belong to class 1. These are all possibilities because each nonzero coefficient is supported on one prime power.

Consequently a single exceptional squarefree value in (5.2) has a very limited effect. If it is a prime ℓ, its coefficient is at most $2RA_\ell$. If it is the product of two distinct primes ℓ, r, its coefficient is at most $2A_\ell A_r$. If it has three or more prime factors, its coefficient is zero. This proves the bound $2B_0R+2B_0^2$ without any knowledge of which exceptional value actually occurs. The ordinary class 1 similarly has coefficient $R^2+\sum_\ell A_\ell^2$. The phase factors $\kappa(n)\overline{\kappa(n')}$ and the reciprocity signs have absolute value at most one, so these coefficient bounds remain valid uniformly in the base character. This is the specific reason that one exceptional real character is compatible with the second moment. The argument uses the arithmetic support of the coefficients before applying cancellation in the prime parameter.

For completeness, all coefficient bounds needed in (5.3) follow from an elementary Chebyshev estimate $\sum_{n\le u}\Lambda(n)\ll u$. For a fixed prime ℓ, the powers $\ell^k\le y$ contribute at most $\log y$, and the geometric tail contributes $O(\log\ell)$ when $\ell\le y$. 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

$$\sum_{\ell,k\ge1}\log\ell\,\min(1,y/\ell^{2k}) = 2y\int_{\sqrt{y}}^{\infty}\frac{\sum_{n\le t}\Lambda(n)}{t^3}\,dt \ll \sqrt{y}.$$

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 $Y\le y\le Y^J$, this is $o(\sqrt{y})$, and hence $R\ll\sqrt{y}$.

To bound $\sum_\ell A_\ell^2$, first take $\ell\le y$. 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 $\sum_{\ell\le y}A_\ell\ll y$, and therefore $\sum_{\ell\le y}A_\ell^2\ll y\log(2y)$. For $\ell>y$, partial summation with $\sum_{\ell\le t}\log\ell\ll t$ 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 $\sum_{n\le Z}c_y(n)\ll y(1+\log Z)$. With $Z=X^{400}$, choose $A=800J$ and $B=2J+2$ in (5.2). The last term of (5.3) is then $O(1)$, uniformly for $Y\le y\le X$. We have proved

$$\sum_{p\in\mathcal{P}}|S_p(y)|^2\ll y\log(2y)\,|\mathcal{P}|,\qquad Y\le y\le X, \tag{5.4}$$

uniformly for every sequence $|\kappa(n)|\le1$. 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 (5.4) applies to $tS_p'(t)$ as well. Here is a general form of the argument,

including its dependence on the exponent. It explains why a longer power interval does not introduce a factor proportional to the number of test points.

Lemma 5.2. *Let $\mathcal{P}$ be finite, $2 \le Y \le X$, and let $F_p : [Y, X] \to \mathbb{C}$ be absolutely continuous. Suppose that for a constant $C \ge 0$ one has $\sum_p |F_p(t)|^2 \le C|\mathcal{P}|t\log(2t)$ for every t, and $\sum_p |tF_p'(t)|^2 \le C|\mathcal{P}|t\log(2t)$ for almost every t. For $h, \tau > 0$, the number of p for which $|F_p(t)| > ht^{1/2+\tau}$ somewhere on $[Y, X]$ is at most*

$$Ch^{-2}|\mathcal{P}|Y^{-2\tau}\left\{\left(1+\frac{1}{\tau}\right)\log(2Y)+\frac{1}{2\tau^2}\right\}. \tag{5.5}$$

The bound is independent of X.

Proof. Set $Q_p(t) = t^{-1-2\tau}|F_p(t)|^2$. This function is absolutely continuous on $[Y, X]$, and almost everywhere

$$Q_p'(t) = \frac{2\,\mathrm{Re}(F_p'(t)\overline{F_p(t)})}{t^{1+2\tau}} - \frac{(1+2\tau)|F_p(t)|^2}{t^{2+2\tau}} \le \frac{2|F_p(t)|\,|tF_p'(t)|}{t^{2+2\tau}}.$$

Integrating from Y to any $v \in [Y, X]$ and then taking the supremum gives

$$\sup_{Y\le v\le X} Q_p(v) \le Q_p(Y) + 2\int_Y^X \frac{|F_p(t)|\,|tF_p'(t)|}{t^{2+2\tau}}\,dt.$$

Sum over p. Cauchy–Schwarz and the two moment hypotheses bound $\sum_p |F_p(t)|\,|tF_p'(t)|$ by $C|\mathcal{P}|t\log(2t)$ almost everywhere. Consequently

$$\sum_p \sup_{Y\le v\le X} Q_p(v) \le C|\mathcal{P}|Y^{-2\tau}\log(2Y) + 2C|\mathcal{P}|\int_Y^\infty t^{-1-2\tau}\log(2t)\,dt.$$

The functions F_p and their derivatives have been used only on $[Y, X]$; the extension to infinity applies solely to the nonnegative comparison integral. Integration by parts evaluates that integral as $Y^{-2\tau}\{\log(2Y)/(2\tau) + 1/(4\tau^2)\}$. Every index violating the claimed bound contributes more than h^2 to the summed supremum. Divide by h^2 and substitute the integral to obtain (5.5). □

Apply the lemma with $F_p = S_p$ and $h = 1/2$. The constant in (5.4) is uniform in the coefficient sequence, so it also applies to the sequence defining $tS_p'(t)$, even though that sequence changes with t. The finitely many breakpoints do not affect the integral estimate. For each fixed $\tau > 0$ the right side of (5.5), divided by $|\mathcal{P}|$, tends to zero with Y. A finite union handles all the base characters. The argument therefore selects one prime for a simultaneous bound at every real $y \in [Y, X]$. The allowable exceptional proportion is quantified before the prime is chosen, and no discretization error remains between test points.

Uniformity in the base character is particularly useful at the second stage. Although q is allowed to vary through an exponentially long interval, the possible base functions are obtained from the fixed characters modulo $8M$ by also allowing multiplication by χ_{-q}. For any fixed q there are at most $2\varphi(8M)$ such functions. Every one has absolute value at most one, which is the only property of the sequence κ used in the finite second-moment calculation. Thus its constants do not increase with the conductor of the varying base character. The progression in the prime parameter does depend on q, but its modulus satisfies the same admitted range uniformly in q. Applying (5.5) to each of these functions and adding their exceptional proportions costs only the fixed factor $2\varphi(8M)$. This gives a common large-subset estimate for all the required twists while retaining the dependence of the progression on the first chosen prime.

Two remaining analytic operations convert this estimate into Proposition 5.1. 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

$$G_\vartheta(y) = -\sum_{n\le Z} \Lambda(n)\vartheta(n)\min(1, y/n) - yT_\vartheta(Z). \tag{5.6}$$

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

$$X\,|T_\vartheta(X^{400})| \le Y^{\varepsilon}. \tag{5.7}$$

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 (5.6) 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.

For clarity, the injectivity behind this counting step holds even if the base character varies with the prime. A product of primitive characters of coprime conductors f, p is primitive of conductor fp: on the Chinese remainder product of the two unit groups, restriction to either factor recovers the corresponding original character, so a smaller conductor would contradict its primitivity. Since $f \le F < p$, the new prime is the unique prime divisor of fp greater than F. Thus the same resulting primitive character cannot be assigned to two different new primes. Choose one offending character for each bad prime; these choices give an injection into the zero-density exceptional set. There is no need to multiply its cardinality by the number of possible base characters. The displayed estimates even give the explicit exceptional proportion $O(U^{-21/40})$, since $3/8 - 9/10 = -21/40$. The fixed gap between these exponents is ample for all the finite exclusions in this step.

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 (5.4). The lower bound on $\log U$ also gives $p > Z^2$. The estimates on $\log Q$ used in (5.7) 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.

The order of these choices is usefully expressed as a finite selection principle. Suppose $\mathcal{Q}$ is a nonempty finite set, and each $q \in \mathcal{Q}$ has a nonempty finite set $\mathcal{D}_q$ of second-stage choices. Fix $0 < \eta < 1$. If a set $\mathcal{Q}_0$ of good first choices has cardinality at least $(1-\eta)|\mathcal{Q}|$, and every $q \in \mathcal{Q}_0$ has at least $(1-\eta)|\mathcal{D}_q|$ good second choices, then a compatible good pair exists. Both lower bounds are strictly positive, so first choose $q \in \mathcal{Q}_0$ and then a good element of its own fiber $\mathcal{D}_q$. More quantitatively, choose q uniformly from $\mathcal{Q}$ and then d uniformly from $\mathcal{D}_q$. The probability of a good pair is at least $(1-\eta)^2$, by summing the conditional probabilities over $\mathcal{Q}_0$. This assertion requires no independence between the two stages and permits their second-stage sets to have different sizes.

In the present application $\mathcal{D}_q = \mathcal{P}_2(q)$ depends on q through its congruence conditions. All second-stage estimates must therefore hold beyond one threshold for every admissible q. The ceiling $F = 8Me^{z}$ supplies exactly that uniformity. Indeed $q \le e^{z}$ bounds every base conductor by F, and $\log F = \log(8M) + z = o(z')$, so $F < e^{z'/2} = \sqrt{U}$ for all sufficiently large Y, independently of the selected q. The estimates for the progression modulus and for the truncated second moment were also stated uniformly in this range. Their finitely many thresholds can therefore be replaced by their maximum before making either choice. This is the large-subset statement in `PairShare`. It is stronger than the existence of unrelated good first and second

primes: the selected second prime belongs to the progression determined by the selected first one.

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 5.1 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 5.3 (`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_0qd, and suppose $\log(Q_0qd) \le Dx^{\varepsilon/5}(\log x)^2$. 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

$$|\psi_0(\vartheta,u)| \le Cu(\log u + \lambda)^2\exp\Bigl(-\frac{c\log u}{\lambda + \sqrt{\log u}}\Bigr), \qquad u \ge e. \tag{5.8}$$

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

$$G_\vartheta(y) = -y\int_y^\infty \frac{\psi_0(\vartheta,u)}{u^2}\,du. \tag{5.9}$$

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 (5.8). 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 $\lambda \le Dx^{\varepsilon/5}(\log x)^2$. 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 (5.8) and (5.9) 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 (5.8).

The exponent comparison has a general form that clarifies the roles of the three scales. Suppose a family satisfies (5.8) with uniform constants, its logarithmic conductor obeys $\lambda \le Dx^a(\log x)^b$, and the test point satisfies $\log y \ge c_0 x^d$, where $D, c_0 > 0$, $b \ge 0$, and $0 < a < d$. For every fixed r with

$$0 < r < \min(d-a, d/2) \tag{5.10}$$

the same integration gives $|G_\vartheta(y)|/y \le e^{-x^r}$ for all sufficiently large x, uniformly in the family and in these test points. To prove this, choose $\max(a/d, 1/2) < \mu < 1 - r/d$; condition (5.10) is exactly what makes that interval nonempty. For $v \ge \log y$ and large x, the fixed power of $\log x$ is absorbed by $x^{d\mu-a}$, so $\lambda \le v^\mu$ and $\sqrt{v} \le v^\mu$. The integral in (5.9) is therefore bounded by a constant times $\int_{\log y}^{\infty} v^2 e^{-c_1 v^{1-\mu}}\, dv$. After the substitution $t = v^{1-\mu}$ this is the integral of a fixed power of t times $e^{-c_1 t}$. For sufficiently large t that fixed power is at most a constant times $e^{c_1 t/2}$: taking logarithms reduces the assertion to $\log t = o(t)$. Integration then bounds the original integral by $C_1 e^{-c_2(\log y)^{1-\mu}}$. Since $d(1-\mu) > r$, this is at most e^{-x^r} after increasing the threshold. This proves the claim directly from the stated analytic input.

In Proposition 5.3, the exponents are $a = \varepsilon/5$ and $d = \varepsilon/3$. Hence the available range is $r < 2\varepsilon/15$. The target $r = \varepsilon/8$ lies strictly inside that range; the explicit choice $\mu = 61/100$ in the proof gives $d(1-\mu) = 13\varepsilon/100 > \varepsilon/8$. This explains the numerical margins. The conductors can be exponentially large in a small power of x because the thin-place arguments have logarithms that are larger powers of x. The weak classical zero-free region then produces a strong enough relative error at those thin arguments. Its use here is separate from the square-root estimate for selected twists on the much shorter interval $[Y, Y^J]$.

For the host constructed next, the required real-character hypothesis is supplied by Proposition 5.1. Indeed, a real primitive character of conductor dividing $Q_0 qd$ 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, $(i,j) \ne (0,0)$. 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`.

6. 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.

The section has one output, Proposition 6.4: if the comparison gap $\sum(\mathcal{A} - \operatorname{Re}\mathcal{M})$ is at most E for every thick translate, then $H_\psi(x)$ is bounded by the discarded character sums, by $2E/(r_T m \delta_2 b)$, and by small explicit errors. Its ingredients appear in the order in which they are used. Lemma 6.1 is the finite calculation behind the coefficient $1/(3o+1)$. Proposition 6.2

provides the thick set. Proposition 6.3 is the exact identity between the gap and a sum of regularized prime sums over all characters of the host. The extraction then removes the thin places, averages over the translates, and reads off the retained characters.

The coefficient that must survive the construction can be determined before any field is chosen. We record that finite calculation first. For a finite abelian group A of order N, write $\widehat{A}$ for its unitary characters and normalize Fourier coefficients by $\widehat{f}(\lambda) = N^{-1}\sum_{a\in A} f(a)\overline{\lambda(a)}$. If $H \subset A$ is a subgroup, its annihilator $H^{\perp} = \{\lambda : \lambda|_H = 1\}$ consists of the characters of A/H. This notation will be used only for the finite calculation; the host subgroup will be $H = A_0$.

Lemma 6.1. *Let $[A : H] = o$, and let $\alpha \geq \beta \geq 0$ with $\alpha > 0$. Put $d = \beta + (\alpha - \beta)/o$ and $W(a) = (\beta + (\alpha - \beta)\mathbf{1}_H(a))/d$. Then $W \geq 0$, $N^{-1}\sum_a W(a) = 1$, and, for every $f : A \to \mathbb{C}$ and $u \in A$,*

$$\frac{1}{N}\sum_{a\in A} W(a)f(u+a) = \widehat{f}(1) + \frac{\alpha-\beta}{\alpha+(o-1)\beta}\sum_{\substack{\lambda\in H^{\perp}\\ \lambda\neq 1}} \widehat{f}(\lambda)\lambda(u). \tag{6.1}$$

In particular, $\alpha = 1$ and $\beta = 3/4$ give coefficient $1/(3o+1)$ for every nontrivial character in $H^{\perp}$.

Proof. We use orthogonality of characters in two forms. The sum of a character λ over a subgroup K is $|K|$ if λ is trivial on K and zero otherwise, and Fourier inversion reads $f(u+a) = \sum_{\lambda}\widehat{f}(\lambda)\lambda(u)\lambda(a)$.

The mean of the unnormalized weight is d, since $|H| = N/o$. Its average against the trivial character is also d. Against a nontrivial character it is $(\alpha-\beta)/o$ when that character belongs to $H^{\perp}$, and zero otherwise, by subgroup orthogonality. Insert Fourier inversion into the left side of (6.1) and divide these averages by d. This gives the asserted formula. Substitution of $\alpha = 1$, $\beta = 3/4$ gives $(1/4)/(1 + 3(o-1)/4) = 1/(3o+1)$. □

This identity separates three features of the desired coefficient. Its positivity comes from putting greater mass on the distinguished coset. Its equality for all nontrivial retained characters comes from putting the same mass on every other coset. Its normalization comes from dividing by the total mass. Thus the number $3o + 1$ has a finite counting origin. No estimate for an L-function enters its derivation. The particular choice $1 : 3/4$ is a convenient bias compatible with the top-parameter construction; other positive biases give the ratio in (6.1).

For real f, conjugate Fourier coefficients occur in conjugate pairs, so the right side of (6.1) is real. Moreover the left side is a positive average: an upper bound valid for all its arguments remains an upper bound after averaging. Both facts will be needed. The first makes the retained prime-sum expression real, and the second permits an upper geometric estimate to pass through the extraction. Positivity of the averaging measure and positivity of its retained Fourier coefficients are separate properties; this particular measure has both.

The preliminary subgroup average is equally explicit. Define $P_H f(u) = |H|^{-1}\sum_{h\in H} f(u+h)$. It keeps precisely the Fourier coefficients in $H^{\perp}$, by the subgroup sum proved above. It also satisfies $P_H^2 = P_H$: each $h \in H$ occurs exactly $|H|$ times as a sum of two elements of H. Since W is constant on cosets of H, its weighted averaging operator T_W satisfies $T_W P_H = P_H T_W = T_W$. Indeed, in either double sum change the variable a to $a + h$ and use $W(a + h) = W(a)$. These identities explain the two averages in Proposition 6.4. The translates first remove characters that vary on A_0; the biased top distribution then gives the specified coefficient to the surviving retained characters. The additional quadratic characters also survive the first projection and are controlled by the auxiliary-prime estimate.

There is a simple quantitative stability statement for this calculation. If ν and ν_0 are probability measures on A/H and $\sum_C |\nu(C) - \nu_0(C)| \leq e$, then their averages against each character differ by at most e. This follows by expanding the difference and using that every character has absolute value one. Consequently, for any complex numbers z_λ indexed by a subset $\mathcal{S}$ of $\widehat{A/H}$, replacing the Fourier coefficients of ν_0 by those of ν changes the resulting linear combination by at most $e\sum_{\lambda\in\mathcal{S}} |z_\lambda|$. The latter assertion is the triangle inequality applied once more. This elementary estimate shows exactly what a rounding error in coset proportions costs.

It will be applied to the finitely many retained sums using a crude bound for those sums, so it requires no cancellation for the L-functions whose zeros are being studied.

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

$$Q_e = \operatorname{lcm}(2^{e+2}, M), \qquad H_e = (\mathbb{Z}/2^{e+2}\mathbb{Z})^\times/\{\pm1\},$$
$$A = \operatorname{im}\bigl(\pi : (\mathbb{Z}/Q_e\mathbb{Z})^\times \longrightarrow H_e \times B\bigr), \qquad \pi(a) = (\operatorname{cl}(a), \bar\psi(a)).$$

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 $e + 2 \ge v_2(M)$, 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

$$x^{2+8\varepsilon} \le n \le 2x^{2+8\varepsilon}. \tag{6.2}$$

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

Take the primes q, d from Proposition 5.1, with Y large enough that they are coprime to Q_e. Define $w : \{\pm1\} \to A$ 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 $Q_e qd$, set

$$\operatorname{art}(a) = \bigl(\chi_{-q}(a), \chi_{-d}(a), \pi(a) + w(\chi_{-q}(a))\bigr) \in \{\pm1\}^2 \times A. \tag{6.3}$$

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 $\mathbb{Q}(\zeta_{Q_e qd})$ fixed by $\ker(\operatorname{art})$. The elementary Galois correspondence identifies $\operatorname{Gal}(E/\mathbb{Q})$ with $\{\pm1\}^2 \times A$ and gives $[E : \mathbb{Q}] = 4n$. Complex conjugation is $\operatorname{art}(-1) = (-1, -1, 0)$, 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 (6.3) 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 $\theta_\chi(\operatorname{art}(v)) = \lambda_\chi(\pi(v) + 2w(\chi_{-q}(v))) = \chi(v)$. 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 $(i, j) \ne (0, 0)$ are therefore exactly the characters to which the auxiliary square-root bounds apply.

One can see the parity correction explicitly without additive notation. For a power χ of ψ, write $\chi(-1) = (-1)^{e_\chi}$ with $e_\chi \in \{0, 1\}$. Then $\lambda_\chi(w(s_q)) = s_q^{e_\chi}$ and $\theta_\chi(s_q, s_d, a) = s_q^{e_\chi}\lambda_\chi(a)$. After pullback, the last factor is $\chi(v)\chi_{-q}(v)^{e_\chi}$; the first factor contributes the same quadratic sign, and their product is $\chi(v)$. For an even character the correction is invisible. For an odd character it inserts the quadratic sign twice, so that the original character is still retained. At complex conjugation the value of θ_χ is $(-1)^{e_\chi} = \chi(-1)$, as required.

This computation explains a useful asymmetry between the two quadratic coordinates. The parity correction is assigned to the q-coordinate, leaving every retained character trivial on

the d-coordinate. After the coordinate interchange below, this becomes (6.4), the sign needed for positive recovery on the thick places. It imposes no evenness assumption on ψ: complex conjugation changes both sign coordinates, while the thick-place extraction tests only the first one. The same construction therefore handles even and odd primitive characters with one formula.

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

$$\theta(1,0,0) = 1 \quad (\theta \in \mathcal{K}). \tag{6.4}$$

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 6.2 (`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*

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

and the following property: whenever $s_1, s_2, s_3 \in T_2$ are pairwise distinct and $i, j \in \mathbb{Z}$ satisfy $0 < |i|, |j| \le L_{\mathrm{orb}}$, $i \ne j$, then $j(s_2 - s_1) \ne i(s_3 - s_1)$. 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 [4], 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 $r = \operatorname{ord}(h) \le 4L_{\mathrm{orb}}$, 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.

For later counting purposes the geometric statement can be expressed in terms of visit times. Fix an initial point t, a step h of order r, a translate u, and an integer $N \ge 1$. The number of indices $0 \le j < N$ for which $t + jh \in T_2 + u$ is at most $\lceil N/r \rceil$ if $r \le 4L_{\mathrm{orb}}$, and at most $2\lceil N/L_{\mathrm{orb}} \rceil$ if $r > 4L_{\mathrm{orb}}$. In the first case the orbit has at most one distinct point in the set, and its visits form at most one residue class modulo r. In the second case partition the indices into consecutive blocks of length L_{orb}, with a possibly shorter final block. Every block has at most two visits by the proved orbit property. This proves both estimates, including all endpoints. For long orbits the proportion of visits is at most $2/L_{\mathrm{orb}} + 2/N$. Short orbits retain the bound determined by their period, which is why the arithmetic splitting conditions and the separate exceptional-orbit estimate remain part of the construction. In particular, the case $r = 1$ can have a visit at every time and is accounted for by the first bound. The construction and orbit statements are proved in `ThickExists` and `ThickOrbit`.

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

$$|N_C - \mu_C| \le \sqrt{2\delta_1 n} \quad \text{for every } C, \qquad \mathcal{B}(T) \le 4b_0. \tag{6.5}$$

For a finite prime set $\mathcal{P}$, the source applies this with $\mathcal{B}(T) = \sum_{p \in \mathcal{P}} \omega_p |\{t \in T : \mathrm{Bad}(p, t, T)\}|$, where the weights are nonnegative and Bad is its specified failure of the orbit interval and tile conditions. Conditional on the separate estimate $\Pr(t \in T, \mathrm{Bad}(p, t, T)) \le C_b x^{-2}\delta_1$ for every $p \in \mathcal{P}$ and $t \in A$, one has $b_0 \le C_b x^{-2}\delta_1 n \sum_{p \in \mathcal{P}} \omega_p$. The variance computation proves the distribution of top parameters; it does not prove this additional probability estimate.

The scope of that probability input is essential. In `topsGroup`, fix constants C_o, L_b and put $H_\varepsilon = 2 + 8\varepsilon$. The hypotheses include

$$\tfrac{1}{2}x^{H_\varepsilon} \le n \le C_o x^{H_\varepsilon}, \qquad o \le C_o, \qquad L_2 \le L_b, \qquad \operatorname{ord}(h_p) \ge x^{\varepsilon/4} \quad (p \in \mathcal{P}),$$

where $h_p \in A$ is the orbit step. The starting point used to index the orbit must lie in the same orbit as t. The good-orbit predicate has the source parameters $K = \lceil 8(H_\varepsilon + 2)/\varepsilon \rceil$, $R_{\mathrm{tile}} = K + 1$, and sampling density δ_1. Here L_2 is the local interval parameter of that predicate; it is distinct from the Behrend parameter L_{orb}. For all sufficiently large x, uniformly under these hypotheses, the failure estimate follows from `NotGoodLe` and `orbitTerms`. The finite selection principle is `select`.

These restrictions are verified in the host application. With the thin cutoff X' specified below, its prime set consists only of inert-class primes $p < X'$ for which the A-component h_p of

Frobenius has $\operatorname{ord}(h_p) \ge x^{\varepsilon/4}$; its weight is $\omega_p = \log p/p$. Thus the order condition is imposed by restricting the sum, not asserted for every inert prime. The rank bound (6.2) and fixed index o give the first two conditions for a fixed $C_o \ge \max(2, J, o)$. For the rank $k \le 2x^{9/2+10\varepsilon}$ chosen later, set

$$L_1(p) = \left\lceil \log_p\big((kx^{26})^4\big) \right\rceil, \qquad L_2 = 2 \sup_{p \text{ in the inert class}} L_1(p),$$

with supremum zero if the set is empty. The splitting conditions in Proposition 5.1 and $x \le Y^J$ give $p > Y^{\varepsilon/5} \ge x^{\varepsilon/(5J)}$ for every such prime. Hence, for large x, $\log((kx^{26})^4) \le 124 \log x$ and $L_2 \le 2(620C_o/\varepsilon + 1)$, a fixed bound. These deductions are proved in `topsBase` and used in `topsChoiceH`; the restricted sum is exactly the predicate `BadOK` in `ThinScalesIffHostBound`. The host construction invokes this theorem in `hostOfFieldOf`. Primes with smaller orbit order are handled by an elementary congruence estimate. Write $N = |H_e| = 2^e$. If an odd prime p has order at most R in H_e, with $1 \le R < N$, put $j = \lfloor \log_2 R \rfloor$. Its order is a power of 2, so $p \equiv \pm 1 \pmod{2^{e+2-j}}$. Bounding the primes in these two progressions by their integers at least 2, and summing the harmonic series, gives

$$\sum_{\substack{p \le Z,\ p \ne 2 \\ \operatorname{ord}_{H_e}(p) \le R}} \frac{\log p}{p} \ll \frac{R}{N} \log^2(NZ) \qquad (Z \ge 3).$$

The projection of h_p to H_e is the class of p, so its order divides $\operatorname{ord}(h_p)$. Taking $R = x^{\varepsilon/4}$ and the larger cutoff $Z = X'' = xn \exp(x^{\varepsilon/3})$, with $N \asymp_\psi n \asymp x^{2+8\varepsilon}$, bounds the short-orbit sum by $O_\psi(x^{-2-85\varepsilon/12}) \le x^{-2-7\varepsilon}$ for sufficiently large x. The source proves the progression bound in `GenericFramesLetters` and its host specialization in `TopChoiceBounds`. The local gain calculation retains this short-orbit contribution as a separate error and combines it with the long-orbit bad-top bound in `LatticeChainSuccOfPartial`. Thus the restriction in the probability estimate is accounted for in the subsequent prime sum.

Let T_1 be a chosen subset satisfying (6.5). For fixed o, the errors there are $o(\delta_1 n)$. Thus $\frac{1}{2}\delta_1 n \le r := |T_1| \le 2\delta_1 n$ (this is the quantity written r_T in Section 4, where r is reserved for a rank) 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 (6.2) 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

$$\left| \frac{1}{r} \sum_{t \in T_1} \lambda(t) - \frac{1}{3o+1} \right| \le o\, x^{-1+\varepsilon}. \tag{6.6}$$

The same statement holds with complex conjugates. The trivial character has average exactly 1. The positive coefficient in (6.6) 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.

In the notation of Lemma 6.1, the ideal probability of the zero coset is $4/(3o+1)$ and that of each other coset is $3/(3o+1)$. The actual probability is N_C/r. Thus (6.5) constructs a deterministic approximation to the exact positive measure used in that lemma, while also imposing the geometric condition $\mathcal{B}(T_1) \le 4b_0$. The available accuracy can be read directly from the count: since $r \ge \delta_1 n/2$, each coset error is at most $6\sqrt{2}/\sqrt{\delta_1 n}$. Equations (6.2) and $\delta_1 = x^{-\varepsilon/8}$ give $\sqrt{\delta_1 n} \ge x^{1+63\varepsilon/16}$. In particular the weaker and simpler bound $x^{-1+\varepsilon}$ used in the main extraction is available with a fixed power of room to spare. Summing the coset errors gives the error for any quotient character by the stability calculation above. This accounts for the coefficient approximation without requiring an exact cardinality in each coset or an exact rational ratio between those cardinalities.

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.

The inertia average is itself a projection of the same kind as P_H. Replacing F_p by another representative modulo I_p just translates the summation index, so $v_\theta(p^a)$ is well defined. Orthogonality on I_p gives $v_\theta(p^a) = 0$ for a ramified character and $v_\theta(p^a) = \theta(F_p)^a$ otherwise. Hence $v_\theta(p^a) = v_\theta(p)^a$ in both cases, and the Euler factor is $(1 - v_\theta(p)p^{-s})^{-1}$. A prime ramified in the host field can therefore still have a nontrivial Euler factor for a character unramified at that prime. The relevant ramification belongs to the character, through its restriction to inertia.

At $p = d$ this distinction is concrete. The inertia subgroup is generated by $(1,0,0)$, as used below. The character η takes value -1 there, so its average and prime-power coefficients are zero. Every retained character takes value 1 there by (6.4); its Euler factor is present and equals the factor for the primitive character associated with ψ^k, since d is coprime to M. Thus the balance identity simultaneously handles characters with different primitive conductors inside one field. Using their primitive L-functions and this inertia projection keeps the Euler factors consistent at every prime.

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

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

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

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.

Both functionals have an exact homogeneity that helps identify the later normalization. For any $c > 0$, $\mathcal{A}(ca,t,cw) = c\mathcal{A}(a,t,w)$ and $\mathcal{M}(ca,cw) = c\mathcal{M}(a,w)$. The first equality follows term by term from $(cu)_+ = cu_+$. For the second, replacing a by ca adds $\log c$ inside the parenthesis, while replacing each of the f weights by $cw(\tau)$ adds the same $\log c$ to their average logarithm; these additions cancel. The remaining exterior factor is multiplied by c, proving the equality.

Thus the prime-sum arguments depend on the ratios a/w, while the comparison gap is homogeneous of degree one in the common size of the weights. At a thick place with $w = b$ and $a \simeq xb$, the argument is x, regardless of how large b is. A much smaller thin weight tests the same identity at a much larger argument, where Proposition 5.3 applies. After the rm top and level choices and the thick-place density have been averaged, the gap must therefore be divided by $rm\delta_2 b$. This is the normalization appearing in (6.11). The exact homogeneity explains why the exponential size of b does not by itself worsen the retained prime-sum estimate; what matters is the comparison gap relative to that size.

Proposition 6.3 (`ch_balance`)**.** *For every positive a, w as above,*

$$\mathcal{A}(a,t,w) - \mathcal{M}(a,w) = \frac{1}{f} \sum_{g \notin \Phi} \sum_{\theta \in \widehat{G}} \overline{\theta(g)}\, w(\rho_t(g)) G_\theta\!\left(\frac{a}{w(\rho_t(g))} \right). \tag{6.7}$$

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 (6.7). □

For this host, the invariant expression for $\mathcal{A}$ is exactly the more concrete divisibility sum needed later. Primes dividing Q_eq are split in $\mathbb{Q}(\sqrt{-d})$, by the congruences of Section 5; 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 (6.7) uses two elementary continuity estimates. For a nonprincipal character, differentiation away from the finitely many breakpoints gives

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

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

$$|wG_\theta(a/w) - vG_\theta(a/v)| \leq \frac{3a}{\min(v,w)}|w-v|. \tag{6.8}$$

The bound $\sum_{n\leq u}\Lambda(n) \leq 3u$ is a form of Chebyshev's elementary estimate: the central binomial coefficient argument gives $\sum_{n\leq u}\Lambda(n) \leq (4\log 2)u$, and $4\log 2 < 3$.

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

$$|G_\theta(y) - G_\theta(x)| \leq V_x \quad (\theta\in\Psi), \qquad V_x = 4(1+\log(x+1)) + |\gamma_E| + K_0 + C_0\log x. \tag{6.9}$$

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

$$|G_\theta(x)| \leq P_x \quad (\theta\in\mathcal{K}), \qquad P_x = x\{4(1+\log(x+1)) + 2 + |\gamma_E| + K_0\}. \tag{6.10}$$

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

Proposition 6.4 (`ch_core`). *Let $T_1 \subset A$ and $T_2 \subset A_0$ be nonempty, put $r = r_T = |T_1|$, $\delta_2 = |T_2|/n$, and let $m \geq 1$ be an integer. Assume $b, b', X' > 0$, $\varkappa, \vartheta \geq 0$, $x \geq 3$, and*

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

For each $t \in T_1$, choose $a_t > 0$ with $|a_t/b - x| \leq 1$. For every $u \in A_0$ and $1 \leq j \leq m$, let $w_{t,j,u} : \Phi \to (0,\infty)$ satisfy $b \leq w_{t,j,u} \leq b+\varkappa$ on $\{(0,1,s+u) : s \in T_2\}$, and $w_{t,j,u} \leq 2b'$ elsewhere. Suppose $|G_\theta(y)| \leq \vartheta y$ for every $\theta \in \widehat{G}$ and $y \geq X'/3$, and suppose the logarithmic-derivative bounds preceding (6.9) 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 $\Sigma(x) = \sum_{\theta\in\Psi\setminus\mathcal{K}}|G_\theta(x)|$,

$$G_1(x) + c_0\sum_{\theta\in\mathcal{K}\setminus\{1\}}\operatorname{Re}G_\theta(x) \le \Sigma(x) + 4o(V_x+\rho P_x) + \frac{2E}{rm\delta_2 b} + \frac{4}{\delta_2 x}. \tag{6.11}$$

Proof. Let $\mathcal{D}$ denote the right side of (6.7). 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 (6.8) bounds the change in the whole sum by $6f(x+1)\varkappa \le b/x$. 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 (6.9) bounds the change by $|\Psi|V_x = 4oV_x$. On $\mathcal{K}$, (6.4) removes the sign coordinate and (6.6), or the stated coefficient hypothesis, replaces the top average by q_θ. Equation (6.10) 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 (6.11). □

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 6.4 proves the Fourier reduction itself without presupposing that geometric bound.

The order of the operations in its proof also matters when the geometric data vary with the translate. The weights $w_{t,j,u}$ are allowed to depend on u, so their character sum is not initially a fixed function times $\theta(u)$. First the uniform thin estimate removes all thin places with its stated error. Next the thick weights are replaced by the common value b, using (6.8). Only then does the remaining dependence on u occur in a single character factor, to which orthogonality applies exactly. The errors were bounded before averaging and uniformly in u, so their averages obey the same bounds. Finally the nearby arguments a_t/b are moved to x before the top Fourier coefficients are used. These steps explain why the comparison theorem must supply a compatible family of weights for every translate, with uniform numerical bounds; independently favorable choices at isolated translates would not give this identity.

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

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

Proposition 5.3 allows $\vartheta = \exp(-\frac{1}{2}x^{\varepsilon/8})$. Since $f \asymp x^{2+8\varepsilon}$, both numerical smallness conditions in Proposition 6.4 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 5.1 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 (6.11) is therefore the normalized geometric gap $2E/(rm\delta_2 b)$. Sections 3 and 4 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 $1/(3o+1)$.

7. 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 [22, 14]. 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.

In terms of Proposition 4.1, this section accounts for two things. The local coefficient estimate, Proposition 7.7, is the origin of the finite gain G in (4.8). The global interface, Proposition 7.11, supplies the integral evaluation maps, their heights, and the row inequality (4.5). Between them, Proposition 7.3 identifies the pairing that is studied integrally with the Bessel period whose correlations are estimated in Section 8. These three propositions are source inputs and are stated without proof. The lemmas around them are proved here: they are the elementary steps by which the inputs are combined, and they show which normalization each input must carry.

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, $m \ge 2$, 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 $0 \le j \le m$ for X_j, Y_j and $1 \le j \le m$ for W_j, V_j. 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 $1 \le j < m$, 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. They introduce no extra size factor in estimates formed from absolute values of the kernel, but their integrality at a finite place must still be checked.

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 weights in this finite sum determine its Hilbert-space adjoint as well as its denominators. To see the two roles in a general finite model, let A, B be finite sets with positive measures m_a, n_b, and give their function spaces the norms $\sum_a m_a|f(a)|^2$ and $\sum_b n_b|g(b)|^2$. For a kernel $(Tf)(b) = \sum_a m_a k(a,b)f(a)$, direct expansion of the inner product gives $(T^*g)(a) = \sum_b n_b\overline{k(a,b)}g(b)$. In the orthonormal coordinate bases the matrix of T is $\big(\sqrt{m_a n_b}\,k(a,b)\big)_{b,a}$, and hence

$$\|T\|_{\mathrm{HS}}^2 = \sum_{a,b} m_a n_b |k(a,b)|^2. \tag{7.1}$$

The identity follows also by summing the squared norms of the images of the vectors $m_a^{-1/2}\mathbf{1}_{\{a\}}$. Thus the stabilizer factors cannot be omitted when a kernel is composed with its adjoint. They are already part of the measure used in the spectral calculation.

For composition, the elementary matrix identity has the same content. If S is a second weighted kernel, then the coefficient of ST is a sum over the intermediate classes with their intermediate measure. When all these measures have denominators dividing D_Γ, each composition introduces the stated additional denominator. A norm estimate and a divisibility estimate can consequently be compared only when their kernel conventions include the same intermediate weights. A change to orthonormal coordinates is entirely legitimate for the complex calculation, but the factors $\sqrt{m_a}$ used in that change are not an assertion about an integral basis. The integral model continues to be the original class-function lattice with its explicitly recorded stabilizer denominators.

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 7.1. *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 $TT^*|_{E_\pi} = \lambda_\pi \,\mathrm{id}$ 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.

The product of the row eigenvalues requires compatible spectral summands at the successive stages. The following calculation describes the needed compatibility without making a commutativity assumption on unrelated operators.

Lemma 7.2. *Let $\mathcal{P}$ be a finite nonempty set. Let $V_0, \dots, V_m$ be finite-dimensional Hilbert spaces with orthogonal decompositions $V_j = \bigoplus_{\pi\in\mathcal{P}} V_{j,\pi}$, and let $T_j : V_{j-1} \to V_j$ preserve these summands. Suppose every $V_{j,\pi}$ is nonzero and $T_jT_j^*|_{V_{j,\pi}} = \lambda_{j,\pi} I$ with $\lambda_{j,\pi} > 0$. For $T = T_m \cdots T_1$ one has*

$$TT^*|_{V_{m,\pi}} = \left(\prod_{j=1}^m \lambda_{j,\pi}\right) I.$$

Consequently the minimum-norm inverse of T has squared operator norm $\exp(-\min_\pi \sum_j \log \lambda_{j,\pi})$.

Proof. Because the decompositions are orthogonal and T_j preserves them, its adjoint sends $V_{j,\pi}$ into $V_{j-1,\pi}$. Inductively, if $S = T_{j-1}\cdots T_1$, then SS^* is multiplication by $\prod_{i<j} \lambda_{i,\pi}$ on $V_{j-1,\pi}$. For $v \in V_{j,\pi}$ this gives $T_jSS^*T_j^*v = (\prod_{i<j}\lambda_{i,\pi})T_jT_j^*v$, which proves the product formula. Every resulting eigenvalue is positive, so T is onto. Apply the inverse formula in Lemma 7.1 with $V = V_0$. □

In the tower, the summands are selected representation spaces, and the inner product identities supply the scalars on those spaces. Their identification between stages is part of the row construction. For arbitrary surjections T_1, T_2, the separate spectra of $T_jT_j^*$ do not determine the spectrum of $T_2T_1(T_2T_1)^*$: the relative positions of the intermediate eigenspaces also matter. The lemma explains which part of the representation theory makes the multiplicative formula available. It also shows that additional kernel directions at an intermediate stage do not cause a problem; positivity is required on the selected target summands, and the inverse always uses the orthogonal complement of the actual kernel.

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

$$\log \|B_u\|_\sigma^2 \le \log \|P_u\|_\sigma^2 + \sum_j \log \|K_j\|_\sigma^2 - \log \kappa_{R,\sigma} - \min_\pi \sum_j \log \lambda_{j,\pi,\sigma}. \tag{7.2}$$

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 7.3 (`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 8. 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 K be the coefficient field, $d_K = [K : \mathbb{Q}]$, and let $\ell_w = \log Nw/d_K$ 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 7.4. *Let N and R be full $\mathcal{O}_K$-lattices and let $B : N_K \times R_K \to K$ be bilinear. Suppose integers a_w, zero outside a finite set, satisfy $B(N_w, R_w) \subset \varpi_w^{a_w}\mathcal{O}_{K,w}$ for every finite place. Set $I = \prod_w \mathfrak{p}_w^{-a_w}$ and $N' = IN$. Then $B(N', R) \subset \mathcal{O}_K$, 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

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

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

Proof. Localization gives $I_w = \varpi_w^{-a_w}\mathcal{O}_{K,w}$, so bilinearity makes $B(N'_w, R_w)$ integral at every w. An element of K integral at all finite places belongs to $\mathcal{O}_K$, which proves the first assertion. If $n = \operatorname{rk} N$, then $\det(IN) = I^n \det N$. The normalized arithmetic degree therefore changes by $-n \log N(I)/d_K$; 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 $2\ell_w$. □

These two elementary mechanisms can now be combined for a single step. This formulation separates the algebraic factorization from the local estimate and the complex estimate, so that the same calculation can be reused through the tower.

Lemma 7.5. *Let K be a number field of degree d_K. Let N, R be nonzero hermitian $\mathcal{O}_K$-lattices, let A, V be finite-dimensional K-spaces with compatible archimedean Hilbert norms, and let $C : N_K \to A$ and $T : V \to R_K$ be linear, with T onto. Suppose a bilinear form $P : A \times V \to K$ vanishes on $A \times \ker T$. There is a unique bilinear form $B : N_K \times R_K \to K$ satisfying $B(n, Tv) = P(Cn, v)$. Suppose $B \neq 0$, and let $\lambda_\sigma > 0$ be the least eigenvalue of $T_\sigma T_\sigma^*$. If integers a_w, zero outside a finite set, satisfy $B(N_w, R_w) \subset \varpi_w^{a_w}\mathcal{O}_{K,w}$, put $I = \prod_w \mathfrak{p}_w^{-a_w}$. Then B is an integral map $IN \to R^\vee$, and*

$$\mu_2(IN) - h(B) \ge \mu_2(N) + 2\sum_w \frac{\log Nw}{d_K} a_w - \frac{1}{d_K}\sum_\sigma \log \frac{\|P_\sigma\|^2 \|C_\sigma\|^2}{\lambda_\sigma}. \tag{7.3}$$

Here the norm of a bilinear form is the supremum of its absolute value on pairs of unit vectors, and $R^\vee$ has its dual norm.

Proof. If $Tv = Tv'$, then $v - v' \in \ker T$, so $P(Cn, v) = P(Cn, v')$. This proves that the proposed formula for B is well defined and bilinear. Surjectivity of T proves uniqueness. At an embedding σ, the unique lift of $r \in R_\sigma$ of minimum norm is $v = T_\sigma^*(T_\sigma T_\sigma^*)^{-1} r$. Its squared norm is at most $\lambda_\sigma^{-1}\|r\|^2$. Consequently $|B(n, r)| \le \|P_\sigma\|\|C_\sigma\|\lambda_\sigma^{-1/2}\|n\|\|r\|$, giving the squared-norm bound inside the logarithm. A nonzero K-linear map remains nonzero at every embedding, so the displayed factors are positive. Average the logarithmic bound, noting that $h(B) = d_K^{-1}\sum_\sigma \log \|B_\sigma\|^2$. Lemma 7.4 proves integrality and gives the stated change of slope. Enlarging the finite lattice does not change its archimedean normed vector space, so it does not change this map height. Subtracting the height bound proves (7.3). □

The local hypothesis in this lemma is imposed on the actual row lattice R_w. A valuation bound known only for a smaller row lattice would not imply it. Likewise, the vanishing on $\ker T$ is an algebraic factorization condition; it cannot be inferred merely from a norm estimate for P. These are the two checks that allow the original period to descend to the retained row space before its integrality is used. They explain why a statement about the local theta series alone is not yet a global divisibility statement.

For a sequence of column steps, replace C by their composition. Submultiplicativity gives $\|C_\sigma\|^2 \le \prod_j \|K_{j,\sigma}\|^2$. For the compatible row steps, Lemma 7.2 identifies the least eigenvalue by the product formula, with the common positive factor $\kappa_{R,\sigma}$ included. Combining these two estimates recovers exactly the height expression (7.2). The finite-place part of the same lemma then raises the column slope by the ideal gain. The analytic and arithmetic estimates are thus combined after composition, on the same pairing. No assertion that the individual steps are integral without a shift, or that the intermediate inverse maps preserve integral lattices, is needed. Their denominators must instead be included in the final local valuation estimate.

There is a small but useful distinction concerning rounding. For real numbers $t_1, \dots, t_m$ one has $0 \le \lfloor \sum_j t_j \rfloor - \sum_j \lfloor t_j \rfloor \le m-1$, because the difference is the floor of the sum of their fractional parts. Rounding a proved valuation bound for the composed pairing therefore avoids inserting a separate possible loss at every step. This does not license adding valuations of unrelated matrix entries: a composition involves a sum of products, and the bound must first be shown for every summand with the same normalization. Once such a uniform bound is established, the nonarchimedean inequality for a sum preserves it. The source ideal is indeed formed by rounding the total local shift, as specified in its definition of I_u.

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_K 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 $\sum_{w|p} \ell_w = (\log p / d_K) \sum_{w|p} f_w \le \log p$, because $\sum_{w|p} e_w f_w = d_K$ and $e_w \ge 1$. The rounding loss is therefore at most $2 \sum_{p \le X} \log p$, independently of d_K. 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 gain from enlarging the final column lattice must be compared with the cost of the smaller lattice on which the local estimate was proved. This can be done exactly before any error bounds are inserted. Suppose $N \subset M$ are full lattices of rank r with the same archimedean metrics, and put $t_w = \mathrm{length}_{\mathcal{O}_{K,w}}(M_w/N_w)/(re_w)$ for $w \mid p$. All but finitely many t_w vanish. For a finite-support real shift s_w, let $I = \prod_w \mathfrak{p}_w^{-\lfloor e_w s_w \rfloor}$. The determinant ideal and the ideal-shift formula give

$$\mu_2(IN) = \mu_2(M) + 2\sum_w \ell_w e_w (s_w - t_w) - 2\sum_w \ell_w \{e_w s_w\}, \tag{7.4}$$

where $\{y\} = y - \lfloor y \rfloor$. Indeed, the local quotient has cardinality $(Nw)^{re_w t_w}$, so replacing M by N lowers its slope by $2\sum_w \ell_w e_w t_w$; replacing N by IN raises it by $2\sum_w \ell_w \lfloor e_w s_w \rfloor$. Adding these changes proves the identity. This proof works for positive or negative shifts; the fractional-part term always has the stated nonnegative sign.

In the local construction, the colength bound places t_w below $\tau_w + \iota_w$, while s_w contains the sum of the jet gains and row contributions minus the changes of level. Thus the expression $\sum_j g_{j,w} - \tau_w$ in (7.12) is already a net gain. Charging τ_w once more after applying that inequality would lose the same lattice cost twice. The residual ι_w and the rounding term still have to

be bounded. The comparison with the ramified column lattice has a separate clause at the distinguished column prime, so this model identity must be applied there only after that clause has supplied the relevant inclusion and colength. Equality of the complex vector spaces does not supply those missing integral facts.

The factors e_w also have a precise origin. A coefficient estimate normalized by $v_p(p) = 1$ becomes an integer valuation estimate after multiplication by e_w. A local module length is already measured with $v_w(\varpi_w) = 1$. Dividing the latter by re_w is what allows it to be combined with the former in (7.4). The weight $\ell_w = \log Nw/[K : \mathbb{Q}]$ then converts both to a normalized arithmetic degree. These conversions are independent of the size of the coefficient field and are compatible with scalar extension. They are also the reason a local valuation, a local length, and a logarithmic map height cannot be added before their normalizations have been specified.

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

$$\mathcal{L}_s(X) = \sum_{h \geq 0} p^{-h} X_{s-2h}^{p^{2h}}. \tag{7.5}$$

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 7.6. *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 (7.5): 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

$$\mathcal{F}_p(D) = \frac{p}{p^2 - 1} \sum_s D_s - \max_x \sum_{s,i} x_{s,i}, \quad \sum_i p^i x_{s,i} + \sum_i x_{s-i,i} \leq D_s, \tag{7.6}$$

where $x_{s,i} \geq 0$ and the index i ranges over odd integers $1 \leq i < d$. The feasible set is nonempty and compact, since the zero array is feasible and $x_{s,i} \leq D_s/p^i$. Thus the maximum exists. Summing the constraints shows $(p+1)\sum_{s,i} x_{s,i} \leq \sum_s D_s$, and hence

$$\frac{1}{p^2-1}\sum_s D_s \leq \mathcal{F}_p(D) \leq \frac{p}{p^2-1}\sum_s D_s. \tag{7.7}$$

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.

The optimization can be understood completely on an orbit of length two. The only permitted displacement is $i = 1$, so with $x = x_{0,1}$ and $y = x_{1,1}$ its constraints are $px + y \leq D_0$ and $x + py \leq D_1$. Its maximal objective is

$$\max(x+y) = \min\left(D_0, D_1, \frac{D_0+D_1}{p+1}\right). \tag{7.8}$$

Every feasible pair satisfies these three upper bounds. If $D_0 \geq pD_1$, the pair $(D_1, 0)$ attains the minimum D_1; if $D_1 \geq pD_0$, use $(0, D_0)$. In the remaining case the pair $x = (pD_0 - D_1)/(p^2-1)$ and $y = (pD_1 - D_0)/(p^2-1)$ is nonnegative, saturates both constraints, and attains $(D_0+D_1)/(p+1)$. This proves the formula, including boundary cases.

For equal capacities $D_0 = D_1 = D$, the residual is $\mathcal{F}_p(D, D) = 2D/(p^2-1)$, at the lower end of (7.7). If $D_1 = 0$, no nonzero transport is possible, and $\mathcal{F}_p(D_0, 0) = pD_0/(p^2-1)$, at the upper end. The estimate therefore remembers how degree is distributed between the parity classes. Replacing the constrained maximum by a separate optimization in each variable would erase this information.

For general orbit length, write $M(D)$ for the maximum in (7.6). Scaling a feasible array shows $M(tD) = tM(D)$ for $t \geq 0$. Adding two feasible arrays shows $M(D+D') \geq M(D) + M(D')$. Since the first term in $\mathcal{F}_p$ is linear, the residual is homogeneous and satisfies $\mathcal{F}_p(D+D') \leq \mathcal{F}_p(D) + \mathcal{F}_p(D')$. It need not be increasing in each capacity. For example, with $p = 5$ and orbit length two, the residual is $25/24$ at $(5, 0)$ but only $1/4$ at $(5, 1)$. The extra capacity permits a transport that reduces the residual by more than its own linear contribution. Consequently the stability estimate (7.10) requires its stated structured perturbation of the weights; it does not follow from monotonicity of an arbitrary capacity problem.

There is a useful elementary way to certify a lower bound without solving the optimization. If nonnegative numbers y_s satisfy $p^i y_s + y_{s+i} \geq 1$ for every allowed pair (s, i), multiplying the capacity constraints by y_s and summing gives $\sum_{s,i} x_{s,i} \leq \sum_s y_s D_s$. Hence $\mathcal{F}_p(D) \geq \sum_s (p/(p^2-1) - y_s)D_s$. The constant choice $y_s = 1/(p+1)$ gives the lower bound in (7.7). More refined choices can retain information about the marked positions. This is the elementary comparison mechanism behind capacity bounds; deriving the particular weights and errors needed for the jet theorem remains the specialized local input.

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

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

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 7.7 (`Loc.ThmA`)**.** *The admissible local jet construction supplies an absolute constant C with the following property. Suppose $p \leq a$, $|\nu_s| \leq a$ for every s, and $2\log_p a + 2 \leq N_0 - 1$. 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*

$$\sum_{s \in \mathcal{T}} \left(\mathcal{T}_s - \frac{p}{p^2-1} C_s\right) + \mathcal{F}_p(D) - CdN_0N_1(\log_p a + 1). \tag{7.9}$$

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 (7.9), 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 (7.9) 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 $k \geq 3$, $w_s > 0$, and $0 \leq E_s \leq 4w_s$, 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

$$\mathcal{F}_p(D) \geq \mathcal{F}_p((k-1)w) - \frac{11}{p}\sum_s w_s. \tag{7.10}$$

Together with (7.9), 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

$$\mathcal{B}_s = \sum_{\substack{i \geq 1 \text{ odd} \\ i \neq d_s}} \left((A_s - (\nu_s)_2)/p^i - c_{s+i}\right)_+.$$

Indeed $\mathcal{B}_s \leq \mathcal{T}_s$: 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 (7.10) 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 $\alpha \geq 0$ and numbers $w_i \geq 0$, the geometric identity $\sum_{i \geq 1,\ i \text{ odd}} p^{-i} = p/(p^2-1)$ gives

$$\frac{p}{p^2-1}\alpha - \sum_{i \geq 1,\ i \text{ odd}} \min(\alpha/p^i, w_i) = \sum_{i \geq 1,\ i \text{ odd}} (\alpha/p^i - w_i)_+. \tag{7.11}$$

Indeed $z - \min(z,w) = (z-w)_+$ term by term, and both series converge by comparison with $\alpha \sum p^{-i}$. If $w_i \geq 1$, the right side vanishes whenever $p^i \geq \alpha$. More generally, truncating after odd indices $i \leq J$ 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

$$\sum_{j=1}^{m} \sum_{t \in \mathcal{T}_w} (\mathcal{V}_{j,t,p} - L_p) - mF_p \le \sum_{j=1}^{m} g_{j,w} - \tau_w. \tag{7.12}$$

Here $\mathcal{V}_{j,t,p}$ is the finite odd-power positive-part sum of (7.11) 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 (7.12) 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

$$\log \Omega = \gamma_E + \log(\pi/v) - \operatorname{Re} \frac{L'(1, \chi_{-d})}{L(1, \chi_{-d})}. \tag{7.13}$$

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,

$$\begin{aligned} P_{\mathrm{ell}} &= \gamma_E + \log(2\pi/\sqrt{d}) - \operatorname{Re} \ell_{\chi_{-d}}, \\ 2\widehat{\deg}_{\mathrm{Ar}}(L_\tau) &= P_{\mathrm{ell}} + \tfrac{1}{2} \log d - \log \mathrm{rd}_E = \log \Omega. \end{aligned} \tag{7.14}$$

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 (7.14) are assembled in `CmEigenLineDegreeFormula.lean`. Classical CM period and Kronecker limit formulas underlie the first identity [28]; 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 4.

Three fields play different roles in this normalization. The field E is the CM field of the unitary construction; its root discriminant appears in the eigenline degree. The auxiliary imaginary quadratic field is determined by $-d$ and supplies the elliptic period and the quadratic character χ_{-d}. The coefficient field K is chosen large enough to split the relevant actions and define the algebraic pairings. An embedding τ selecting a CM eigenline is therefore not the same index as an embedding $\sigma : K \hookrightarrow \mathbb{C}$ at which a norm is measured. The arithmetic degree averages the latter norms over all σ, including both members of every complex conjugate pair. Enlarging K to make the constructions compatible does not multiply this normalized degree.

Lemma 7.8. *Let K'/K be a finite extension, and let M be a nonzero hermitian $\mathcal{O}_K$-lattice. Equip $M' = \mathcal{O}_{K'} \otimes_{\mathcal{O}_K} M$ with the norms pulled back along the restrictions of embeddings. Then $\widehat{\deg}_{\mathrm{Ar}}(M') = \widehat{\deg}_{\mathrm{Ar}}(M)$ and $\mu_2(M') = \mu_2(M)$. If a nonzero map is extended with these same metrics on its source and target, its normalized height is unchanged.*

Proof. Take a nonzero rational determinant section v. At a finite place w of K' above v_0 of K, its order after base change is $e(w/v_0)$ times its old order. Also $\log Nw = f(w/v_0) \log Nv_0$ and $\sum_{w|v_0} e(w/v_0) f(w/v_0) = [K':K]$. Thus the finite sum in the degree formula is multiplied

by $[K' : K]$. Each complex embedding of K extends to exactly $[K' : K]$ embeddings of K', counted as embeddings into $\mathbb{C}$, and the prescribed determinant norms are the same on each such extension. The archimedean sum is multiplied by the same factor. Division by $[K' : \mathbb{Q}]$ proves the degree identity; the rank is unchanged. The proof for map heights uses the identical embedding count and the unchanged operator norm at each extension. □

This lemma concerns scalar extension of the same integral lattice and the same metric. Choosing a new integral model after extension can introduce a finite index, which must still be measured. In the CM argument, the stronger assertion needed to compare the particular Lie lattices and their polarizations is supplied by the compatible base-change and tensor statements cited above. Their role is to identify that geometric model with the one to which the elementary normalization applies. The terminal algebra in (7.14) is then exact: $\log(2\pi/\sqrt{d}) + \frac{1}{2}\log d - \log \mathrm{rd}_E = \log(2\pi/\mathrm{rd}_E)$. There is no residual $\frac{1}{2}\log d$ to be charged a second time to the error budget.

For a rank-one line, passage from the Lie line to its dual reverses its degree. More concretely, if v has squared norm n and $v^\vee(v) = 1$, then $v^\vee$ has squared norm n^{-1}. Thus the contribution of this distinguished vector to a normalized squared-norm degree is $-\log n$ for the Lie line and $+\log n$ for the dual. This sign is the reason the vector in the elliptic period formula has been specified. In the jet estimate the period enters as $|c_\tau r_\tau|^2 = \Omega_\tau^{-1}$, so a derivative of total degree N contributes $-N \log \Omega_\tau$ to the logarithm of its squared norm. A change to a differential convention must change this identification as well; it is not a harmless alteration of notation.

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 7.9. *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

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

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 $|c_i r_i|^{2\beta_i} = \Omega_i^{-\beta_i}$. □

A normalized monomial makes the leading terms visible without any asymptotic representation theory. For an integer $N \ge 1$ and $r > 0$, put $G_N(z) = (er^2/N)^{N/2}z^N$. The maximum of $t^N e^{-r^2t^2/2}$ for $t \ge 0$ occurs at $t^2 = N/r^2$ and equals $(N/(er^2))^{N/2}$. Hence $|G_N(z)| \le e^{r^2|z|^2/2}$, with equality in the radial estimate at that radius. If $|cr|^2 = \Omega^{-1}$, then

$$|c^N G_N^{(N)}(0)|^2 = (N!)^2\left(\frac{e}{N\Omega}\right)^N. \tag{7.15}$$

Stirling's formula gives its logarithm as $N\log N - N - N\log\Omega + \log(2\pi N) + O(N^{-1})$. Thus both the entropy term and the linear period term already occur in a one-variable example attaining the optimized Cauchy estimate. Their cancellation later in the tower is a comparison of substantial terms, not a consequence of choosing an artificially loose upper bound.

The factorial in this example also identifies the lattice issue in a Taylor expansion. Writing $G(z) = \sum a_N z^N$ gives $G^{(N)}(0) = N!a_N$. At a finite prime, replacing the derivative by the coefficient changes its valuation by $v_p(N!)$; at infinity, it changes the squared norm by $(N!)^2$. One must make both changes if one uses divided powers. The local divisibility statement and the complex block estimate here use the ordinary derivatives, so their factorial contributions are compatible.

A related Hilbert-space calculation explains why coordinate estimates alone do not determine the norm of a representation-valued jet. For the Gaussian measure $(r^2/\pi)e^{-r^2|z|^2}\,dA(z)$ on $\mathbb{C}$, polar coordinates give $\int_{\mathbb{C}} |z|^{2N}(r^2/\pi)e^{-r^2|z|^2}\,dA = N!r^{-2N}$. The angular integral makes distinct monomials orthogonal. Therefore for a polynomial $G = \sum a_N z^N$ its Gaussian squared norm is $\sum |a_N|^2 N!r^{-2N}$, and the functional $G \mapsto G^{(N)}(0)$ has squared dual norm $N!r^{2N}$. The upper bound follows from Cauchy–Schwarz on the coefficient, and equality is attained by z^N. Products of these identities give the multivariable formula. This is a Hilbert norm on a scalar polynomial space. The actual block map uses invariant norms on Schur modules and a specified highest-weight operator. Its passage from these scalar coefficients to the block norm is additional information, precisely the information retained in (7.16).

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 $a \geq 1$ and λ'_τ a dominant weight of length $b \geq a$. 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

$$|G(z)| \leq S \exp\left(\tfrac{1}{2}\sum_\tau r_\tau^2 \|z_\tau\|_{\mathrm{HS}}^2\right), \qquad S \geq 0.$$

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

$$\|J\|^2 \leq S^2 \prod_\tau \left[\Omega_\tau^{-|\nu_\tau|} e(2|\nu_\tau| + 2)^{2ab} \exp\left(\sum_i \Phi(\nu_{\tau,i})\right)\right]. \tag{7.16}$$

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 $a_0, \ldots, a_{n-1}$ in its admissible range, the source uses

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

The second sum is a Weyl denominator. The admissible ordering makes its arguments positive. The estimate used in (7.2) 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, eventually along the constructed family, a stagewise lower bound $L_{6,j}$ holds for $\log \lambda_{j,\pi,\sigma}$ at every selected nonzero row component and every embedding, for the fixed base tuple of local types. The bound is independent of π and σ. Summing in j, taking the minimum, and averaging therefore preserves this lower bound. The source indexes the Hermitian norms by infinite places, so the two conjugate embeddings contribute the same value. The order of the minimum and the embedding average can be tested on a two-by-two example. Take two embeddings with equal weights and two spectral components. Let the sums $\sum_j \log \lambda_{j,\pi,\sigma}$ be $(-M, 0)$ at the first

embedding and $(0,-M)$ at the second, where $M>0$. The average of the minimum is $-M$, whereas the minimum of the two embedding averages is $-M/2$. Replacing the former by the latter would underestimate the inverse-norm cost by $M/2$. In general, for any finite array $u_{\pi,\sigma}$,

$$\frac{1}{[K:\mathbb{Q}]}\sum_{\sigma}\min_{\pi}u_{\pi,\sigma}\le\min_{\pi}\frac{1}{[K:\mathbb{Q}]}\sum_{\sigma}u_{\pi,\sigma},$$

because each pointwise minimum is at most each selected entry before averaging. This inequality has the wrong direction for making the suggested replacement in an upper bound for the inverse height.

There are two legitimate ways to obtain a lower bound for the quantity actually required. One may estimate the minimum separately at every embedding and then average, or prove a lower bound valid simultaneously for every component at every embedding. The stagewise bounds just stated use the latter method: if $\log\lambda_{j,\pi,\sigma}\ge L_{6,j}$ for every π,σ, then the average of the minimum of the stage sum is at least $\sum_j L_{6,j}$. A mean bound over components, even with positive multiplicity weights, would be insufficient. A single very small eigenvalue can govern the inverse operator while having an arbitrarily small weight in such a mean.

The precise uniform input is `IndefData.L6` in `IndefSide.lean`: $N.l6_j\le\log S.\lambda_{j,\mu_0,\pi,v}$ for every j,π,v. The family statement `LambdaLowerBound₂` in `Realised.lean` is proved by `lamLow_of_hypotheses_alt_alt` in `LamLowOfPublished.lean`. Its seven inputs are cusp rapid decay, the Weil–Siegel and convergent Siegel–Weil formulas, local doubling zeta integrals, unramified GL doubling, Bernstein–Zelevinsky finiteness, and the compact-dual formal-degree formula. The construction supplies its remaining size, rationality, and compatibility conditions; the proof treats the base place and the other places separately. In the active assembly, `ev entually_L6_bound_of_lamLow` and `row_rRows_of_parts_alt` put this bound into the row data. The fixed-scale hook identity and the comparison with the numerical h_X are proved in `EigenvalueSumLeClosedOf.lean`. Then `wire_sel` and `wire4F` insert the bound into the division package consumed by `linkDIVt_closedK` in `ShapeWin.lean`.

The correct Stirling expansion is $\log\Gamma(y)=(y-\frac12)\log y-y+\frac12\log(2\pi)+O(y^{-1})$. For $\log\Gamma(y+1)$ the coefficient of $\log y$ is instead $y+\frac12$. 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 7.10. *Let $a_0\ge\cdots\ge a_{n-1}\ge0$ be integers, with $n\ge1$, 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 $a\ge1$, 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.

A single matched block illustrates the cancellation quantitatively. Let $\nu=(\nu_0,\dots,\nu_{n-1})$ be a nonnegative decreasing integer weight, let $A=\nu_0$, and let $J\ne0$ satisfy the one-block version of (7.16), with source rank n and target rank $b\ge n$. Suppose a positive row eigenvalue satisfies $\log\lambda\ge\mathcal{H}_n(\nu)-E$ with $E\ge0$. Subtracting this lower bound from the logarithm of the block

bound and using Lemma 7.10 gives

$$\log\frac{\|J\|^2}{\lambda} \le -|\nu|\log\Omega + 2\log S + E + 1 + n + 2nb\log(2|\nu|+2) + \frac{n(n+1)}{2}\log(A+n). \tag{7.17}$$

The proof is simply that the two sums $\sum_i \Phi(\nu_i)$ cancel, with their discrepancy from the gamma expression bounded by the hook estimate. All remaining weight dependence outside the linear period term is logarithmic, multiplied by the stated rank factors. The condition $J \neq 0$ ensures $S > 0$, so the logarithm is defined.

This matched-block calculation is an illustration of a justified cancellation, not a substitute for matching the actual weights of the tower. When the kernel and row weights are different, one must write their difference before applying an estimate such as (7.17). A discrepancy in every row can add a substantial error after summing over all stages and embeddings. Conversely, once the exact parameter identities match their large parts, the elementary remainder bound shows why a large factorial term need not remain in the final comparison. The construction is designed to preserve precisely this matching while retaining the finite-place divisibility gain. Its striking feature is the compatibility of these two requirements; neither a scalar Gaussian estimate nor a qualitative theta norm identity supplies that compatibility on its own.

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

$$M_u = \sum_{t\in\mathcal{T}} \left[m\{\Phi(a_t) - a_t\log\Omega\} - a_t\log T_u \right]. \tag{7.18}$$

The $a_t \log T_u$ term comes from evaluation at the final vector. The period term is the same Ω as in (7.13), and the finite contribution $F_{j,w}$ is retained until the lattice comparison.

Proposition 7.11 (`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 $B_u \neq 0$, and*

$$\mu_2(R^\vee) + h_u - F_0 \le M_u + C|\mathcal{T}|a\mathcal{L}_a + Cm(B_0 + N_c)x^{C_{\mathrm{abs}}\varepsilon}, \tag{7.19}$$

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

$$\left| h_u + \Big(\sum_t a_t\Big)\log T_u - h_{u'} - \Big(\sum_t a_t\Big)\log T_{u'} \right| \le C(|\mathcal{T}|a\mathcal{L}_a + mB_0).$$

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 (7.12) and (7.19) without referring again to the internal theta models.

8. 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.

The three tasks correspond to inputs of the comparison in Section 4. The covolume estimate, Proposition 8.5, and the interpolation theorem, Theorem 8.6, control the slopes of the row lattices and so enter the slope inequalities of Proposition 4.1. The spectral estimate, Proposition 8.9, gives block bounds for the Gram operators of the evaluation maps; Lemma 8.10 aggregates them into the frame hypotheses (3.8) of Theorem 3.5, which yields the transfer inequality (4.4). As in Section 7, the named propositions are source inputs, and the lemmas proved here are the reductions that connect them.

Arithmetic Hilbert–Samuel theory and arithmetic bigness are classical inputs [24]. 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

$$\mu_2(L_{\mathrm{core}}) \geq -M_j - C_2E_0, \qquad \mu_2(L) \geq \mu_2(L_{\mathrm{core}}) - C_1E_0. \tag{8.1}$$

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 $\operatorname{lcv}(M_j^{\mathrm{row}})/D \leq -\mu_2(L_g)$, 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

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

Here $b \geq 0$ and $P\kappa > 0$ is the scale of the auxiliary weight; the source size conditions also give $|M_j| \leq P\kappa$. 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.

Here is an elementary form of the passage from spanning to mean slope. It clarifies why the three inputs in the displayed comparison are used together.

Lemma 8.1. *Let A, B be nonzero hermitian lattices over a number field, of ranks $a \le r$. Suppose B_K is spanned by integral vectors whose norm at every embedding is at most e^P, where $P \ge 0$. Then $\mu_{2,\min}(B) \ge -2P$. If an integral injection $T : A \to B$ has $\|T_\sigma\| \le e^b$ at every embedding, then*

$$\mu_2(B) \ge \frac{a}{r}(\mu_2(A) - 2b) - 2\left(1 - \frac{a}{r}\right)P. \tag{8.2}$$

In particular, if $P > 0$, $b \ge 0$, $r/a \le 1 + b/P$, and $\mu_2(A) \le MP$ for some $M \ge 0$, then $\mu_2(B) \ge \mu_2(A) - (M+4)b$.

Proof. Let Q be any nonzero torsion-free quotient of B, of rank q, with its quotient metrics. The images of the given vectors span Q_K, so choose q of them that are linearly independent. Their wedge is a nonzero integral section of $\det Q$. Its finite orders are nonnegative, and Hadamard's inequality bounds its norm by e^{qP} at every embedding. The arithmetic-degree formula therefore gives $\widehat{\deg}_{\mathrm{Ar}}(Q) \ge -qP$. Taking the minimum over quotients proves the first assertion.

Let S be the saturation of $T(A)$ in B. The determinant bound for T gives $\mu_2(T(A)) \ge \mu_2(A) - 2b$, and saturation can only increase degree. The quotient B/S has rank $r - a$ and, if nonzero, mean slope at least $-2P$ by the first assertion. Degree additivity in the sequence $0 \to S \to B \to B/S \to 0$ proves (8.2). For the last assertion, put $\theta = a/r$. The rank hypothesis gives $1 - \theta \le b/P$. The displayed lower bound is at least $\mu_2(A) - (1-\theta)(\mu_2(A) + 2P) - 2b$, which is at least $\mu_2(A) - (M+4)b$ under the remaining hypotheses. □

Apply the lemma to the combined map $s \mapsto (\phi_h(s))_h$ from Λ_- to the orthogonal direct sum of the copies of Λ. The trivial common kernel makes this map injective, and the sum-of-squares inequality is exactly its operator-norm bound. The same short vectors placed in each summand span the direct sum, whose mean slope equals that of Λ. The rank comparison then bounds the relative dimension of the complement. To see the resulting loss directly, write $W = P\kappa$ and let θ be the ratio of the source rank to the rank of this direct sum. Thus $1 - \theta \le b/W$. The component and flat-line estimates below give $\mu_2(\Lambda_-) \ge -M_j - E$ for some $E \ge 0$ on the error scale. Formula (8.2) then gives

$$\mu_2(\Lambda) \ge \theta(-M_j - E - 2b) - 2(1-\theta)W \ge -M_j - E - 5b.$$

Indeed, $(1-\theta)|M_j| \le b$, $\theta E \le E$, $2\theta b \le 2b$, and $2(1-\theta)W \le 2b$. The separate descent estimate loses at most another $2b$, giving $\mu_2(L_{\mathrm{core}}) \ge -M_j - E - 7b$. This argument uses the lower bound for $\mu_2(\Lambda_-)$ and the stated bound for M_j; it requires no upper bound for the mean slope of Λ_-.

This argument explains the need for spanning at every embedding with one common height bound. Finding many short vectors at a single embedding would not bound their arithmetic degree, and finding many integral vectors without verifying their rank would not bound the quotient slope. It also explains why the rank ratio must be quantitative: a complement of relative dimension θ can cost $2\theta P$, which is substantial when the auxiliary weight P grows. The source construction arranges that the product of this relative dimension and the weight is on the prescribed error scale.

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

$$d_Z^{\mathrm{ar}} \ge b_T d_Z \left[T h_Z - 3(q+1)|h_Z| - A\log(2 + |h_Z|) - R\right]. \tag{8.3}$$

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 (8.3). They also explain why changes of component, degree, and tensor power must be tracked together.

Lemma 8.2. *Let $\mathcal{Z}$ be a finite nonempty set. Suppose $h_Z, s_Z \in \mathbb{R}$, $c_s, c_c, X \geq 0$, and $0 \leq \rho, z < 1$ satisfy $h_Z + c_s \geq 0$, $(1 - \rho)\max(0, s_Z + c_s) \leq h_Z + c_s + c_c$, and $(\#\mathcal{Z})^{-1}\sum_Z s_Z \geq -X$. Suppose also $d_0 \leq d_Z \leq d_0/(1 - \rho)$ for some $d_0 > 0$, and $b_T d_Z \leq r_Z \leq b_T d_Z/(1 - z)$ with $b_T > 0$. Then*

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

Proof. Put $\theta = (1 - z)(1 - \rho)$. The rank bounds imply $r_Z/\sum_W r_W \geq \theta/\#\mathcal{Z}$. Since $h_Z + c_s \geq 0$, the weighted average of these nonnegative numbers is at least θ times their unweighted average. The second hypothesis gives $(\#\mathcal{Z})^{-1}\sum_Z(h_Z + c_s) \geq (1 - \rho)(c_s - X) - c_c$, because $\max(0, s_Z + c_s) \geq s_Z + c_s$. 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 $0 < \theta \leq 1$ and $1 - (1 - z)(1 - \rho)^2 \leq z + 2\rho$ 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 (8.3) is controlled using $|h_Z| \leq h_Z + 2c_s$, 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 7. The multiplication maps and descent then give the first inequality in (8.1); 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.

One can make the remaining numerical part of the component estimate explicit. The following consequence uses only the displayed rank and height inequalities; the geometric construction is responsible for placing its terms on the common error scale.

Lemma 8.3. *Assume the hypotheses of Lemma 8.2 and the degree bound (8.3). Let $\varepsilon > 0$ and suppose $\alpha = T - 3(q + 1) - \varepsilon \geq 0$. Put*

$$B = \big(6(q + 1) + \varepsilon\big)c_s + A\log(2 + c_s + A/\varepsilon) + R.$$

Then

$$\frac{\sum_Z d_Z^{\mathrm{ar}}}{\sum_Z r_Z} \geq -(1 - z)\alpha(X + c_c) - \alpha\big((1 - z)(z + 2\rho) + z\big)c_s - B. \tag{8.4}$$

Proof. Write $y = h_Z + c_s \geq 0$. Then $|h_Z| \leq h_Z + 2c_s = y + c_s$. Concavity of the logarithm, evaluated at $2 + c_s + A/\varepsilon$, gives $A\log(2 + c_s + y) \leq \varepsilon y + A\log(2 + c_s + A/\varepsilon)$: its tangent has slope $A/(2 + c_s + A/\varepsilon) \leq \varepsilon$, and its constant term is at most the displayed constant. The bracket in

(8.3) is therefore at least $\alpha h_Z - B$. Put $w_Z = b_T d_Z / r_Z$, so $1 - z \le w_Z \le 1$. Because $h_Z \ge -c_s$, one has $w_Z h_Z \ge (1-z)h_Z - z c_s$. Also $B \ge 0$, so $-w_Z B \ge -B$. Dividing the component bound by r_Z yields $d_Z^{\mathrm{ar}}/r_Z \ge (1-z)\alpha h_Z - \alpha z c_s - B$. Average with weights $r_Z / \sum_W r_W$ and use Lemma 8.2. This is (8.4). □

The size hypotheses already displayed give a concrete consequence. Suppose $E_0 \ge 1$, $C \ge 0$, $T \ge 3q+4$, and each of TX, Tc_c, $c_s[6q+10+T(z+2\rho)]$, and $A\log(\max(A,1)) + R$ is at most CE_0. Then

$$\frac{\sum_Z d_Z^{\mathrm{ar}}}{\sum_Z r_Z} \ge -(8C+2)E_0. \tag{8.5}$$

To verify this, use Lemma 8.3 with $\varepsilon = 1$. The first term of its loss is at most $2CE_0$, and the second is at most $2CE_0$, since $\alpha \le T$ and $(1-z)(z+2\rho) + z \le 2(z+2\rho)$. For the logarithmic term, concavity gives $A\log(2 + c_s + A) \le A\log(2+A) + c_s$. If $A \ge 1$, use $A \le A\log A + 1$ and $\log(2+A) \le \log A + \log 3$; if $A < 1$, use $A\log(2+A) \le \log 3$. In both cases $A\log(2+A) \le 3A\log(\max(A,1)) + 2$. Consequently the quantity B in the lemma is at most $(6q+8)c_s + 3\{A\log(\max(A,1)) + R\} + 2 \le (4C+2)E_0$. Adding the three estimates proves (8.5).

This calculation accounts for the apparent extra dependence on the shift c_s inside the logarithm. Its cost is at most one further copy of c_s, already covered by the displayed height-variation bound. It also shows why the error scale is taken at least one: the elementary logarithmic estimate has an absolute constant term. All dependence on the number of components has disappeared after normalization by their total rank. The component construction must still supply the rank shares and the height inequalities uniformly, but their subsequent summation is now a complete numerical argument. The source component-sum comparison, recorded in `CovolXOfCovol`, has the normalization $2(\sum_Z d_Z^{\mathrm{ar}})/(\sum_Z r_Z) \le \mu_{\mathrm{base}}$, where μ_{base} is the squared-norm slope of the untwisted section lattice. Consequently (8.5) gives $\mu_{\mathrm{base}} \ge -(16C+4)E_0$. The factor two applies to the whole error term and is independent of the component weights. The compatible flat-line degree comparison then introduces the main period term, while the multiplication maps and integral descent carry the estimate to the core lattice. Those maps and degree identities are part of the geometric construction; the averaging, logarithmic absorption, and normalization in the present passage have all been made explicit.

The proof also separates averaging from tensor growth. The large positive coefficient T is useful only after the possibly negative component heights have been averaged with admissible weights. Controlling the discrepancy between rank weights and component weights uses the lower shift c_s and the rank defects z, ρ. It would be incorrect to replace h_Z by a common positive number on each component, or to assume that increasing T improves a negative component height. What grows is the controlled average, with its explicit variation and logarithmic losses. Those are the quantities that survive descent to the type lattice.

Lemma 8.4. *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 $g \ge 1$ satisfies $ge(L) \subset L$. Put $L_g = L \cap \operatorname{im} e$ and $L_b = L \cap \ker e$. If $0 \le \delta \le 1/12$ and $Q \ge 0$ satisfy*

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

then $L_g \ne 0$ and $\mu_2(L_g) \ge \mu_2(L) - 12\delta Q$.

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 $r_b \le r/4$, we have $r_g > 0$ and $r_b/r_g \le 3\delta/(1-3\delta) \le 4\delta$. 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

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

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

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`.

A simple lattice example shows why an integral projector bound is needed even when the discarded rank is small. Take $L = \mathbb{Z}^r$ with its standard Euclidean metric, where $r \geq 4$, and let e be the orthogonal projection onto the hyperplane perpendicular to $v = (M, 1, 0, \dots, 0)$, with a positive integer M. Then $g = M^2 + 1$ clears the denominators of e, since $e = I - vv^{\mathrm{t}}/(M^2+1)$. The retained lattice has basis $(1, -M, 0, \dots, 0), e_3, \dots, e_r$, whose Gram determinant is $M^2 + 1$. Hence $\mu_2(L_g) = -\log(M^2+1)/(r-1)$, whereas $\mu_2(L) = 0$ and the discarded rank is just one. For a fixed large r, the discarded fraction is small but the slope loss is unbounded as M grows. Here $\delta = 1/(3r)$ satisfies the rank hypothesis of Lemma 8.4. Its denominator hypothesis detects exactly the quantity that is growing. The example concerns an ordinary rational Euclidean lattice; it illustrates the integral obstruction independently of the automorphic setting.

The saturation condition is equally concrete. The integral points in the retained rational subspace form $L \cap \operatorname{im} e$, whereas the orthogonal projection $e(L)$ can be a larger fractional lattice. Confusing these lattices can reverse an index contribution. The proof of Lemma 8.4 measures their discrepancy through $L/(L_g \oplus L_b)$ and the denominator g. In the row construction, projections and images are specified separately for the same reason: their complex spaces may be isometric while their integral covolumes differ. The exponentially small discarded rank becomes useful only after that discrepancy has been bounded.

Proposition 8.5 (`lcv_le`)**.** *Assume (8.1), 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

$$\frac{\operatorname{lcv}(M_j^{\mathrm{row}})}{D} \leq M_j + (C_1 + C_2 + 1)E_0. \tag{8.6}$$

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 8.4 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 (8.6). 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 8.6 (`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

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

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

$$\int_M |F|^2 e^{-\varphi_2-(n+j)\varphi}\, dV_z \le \frac{280\pi^n r^{2n}}{(n-1)!} \sum_{p\in Z} |c_p|^2 e^{-\varphi_2(p)-j\varphi(p)}. \tag{8.7}$$

This is the precise analytic input used here. Its statement is `JacobianInterpolationStatement`; `OtO` 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 [25]. 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 $0 \le \rho \le 1$. 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

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

then shows that the shrinking shell $\varepsilon^2/2 \le \rho \le \varepsilon^2$ 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 $\rho \le 1$ gives $X_\varepsilon \ge 1$ and $\eta_\varepsilon \ge 1$. 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 $\varepsilon > 0$,

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

Choose $\varepsilon < 1/100$ and $(1+\varepsilon^2)^n < 1 + 1/200$. The inequality $(a+b)^2 \le (101/100)a^2 + 101b^2$ 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 (8.7) 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 $\|g\|^2 e^{-\varphi} \le r^2$ is global.

Lemma 8.7. *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 $\|f\|^2 \le C\|c\|^2$, where $C > 0$. Then $EE^* \ge C^{-1}\,\mathrm{id}$ and $\det(EE^*) \ge C^{-q}$.*

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 $\|S\| \le \sqrt{C}$. For $c \ne 0$, $\|c\|^2 = \langle ESc, c\rangle = \langle Sc, E^*c\rangle$ 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 8.6, using the normalized evaluations $F(p)/J_g(p)$ and the point weights in (8.7). 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.

To make the weighted evaluation determinant explicit, enumerate the zeros as $p_1, \dots, p_q$ and set $w_i = e^{-\varphi_2(p_i) - j\varphi(p_i)} > 0$ and $C = 280\pi^n r^{2n}/(n-1)!$. Let $E(F) = (F(p_i))_i$ be the unnormalized evaluation map and let D be the diagonal matrix with entries $\sqrt{w_i}/J_g(p_i)$. Then the normalized map DE sends a function with values $F(p_i) = c_i J_g(p_i)$ to $(\sqrt{w_i} c_i)_i$. Theorem 8.6 and Lemma 8.7 give $(DE)(DE)^* \geq C^{-1} I$. Taking determinants and using $|\det D|^2 = \prod_i w_i/|J_g(p_i)|^2$ yields

$$\det(EE^*) \geq C^{-q} \prod_{i=1}^{q} |J_g(p_i)|^2 e^{\varphi_2(p_i) + j\varphi(p_i)}. \tag{8.8}$$

This is a determinant on the finite-dimensional evaluation target; the Bergman space itself may be infinite dimensional. The positive operator EE^* is nevertheless a finite matrix, and the inverse on the orthogonal complement of $\ker E$ is the minimum-norm interpolation operator.

For the determinant conclusion, separate coordinate interpolants would already suffice without a dimension loss. Indeed, let $\mathcal{E}$ denote any bounded surjective evaluation map to $\mathbb{C}^q$, and suppose each standard basis vector has a preimage of squared norm at most C. Its minimum-norm right inverse is $S_0 = \mathcal{E}^*(\mathcal{E}\mathcal{E}^*)^{-1}$. The squared norm of $S_0 e_i$ is the ith diagonal entry of $(\mathcal{E}\mathcal{E}^*)^{-1}$, so every such entry is at most C. Hadamard's determinant inequality therefore gives $\det((\mathcal{E}\mathcal{E}^*)^{-1}) \leq C^q$, or $\det(\mathcal{E}\mathcal{E}^*) \geq C^{-q}$. Applying this argument to $\mathcal{E} = DE$ recovers the same determinant bound as above.

Simultaneous interpolation gives the stronger operator inequality $\mathcal{E}\mathcal{E}^* \geq C^{-1} I$, which bounds every singular value separately. Coordinatewise interpolation does not imply that inequality with the same constant. For example, if $(\mathcal{E}\mathcal{E}^*)^{-1} = \left(\begin{smallmatrix} 1 & 1/2 \\ 1/2 & 1 \end{smallmatrix}\right)$, both coordinate minimum norms have square 1, but the smallest eigenvalue of $\mathcal{E}\mathcal{E}^*$ is $2/3$. Theorem 8.6 supplies the simultaneous estimate; its determinant consequence can also be understood through this weaker coordinatewise hypothesis.

Here is the exact arithmetic identity into which such a determinant estimate is inserted.

Lemma 8.8. *Let A, B be hermitian $\mathcal{O}_K$-lattices of the same positive rank r, where $d_K = [K : \mathbb{Q}]$, and let $T : A \to B$ be an integral map inducing an isomorphism over K. Set $I = \#(B/T(A))$. Then*

$$\mu_2(B) = \mu_2(A) + \frac{2\log I}{d_K r} - \frac{1}{d_K r} \sum_{\sigma} \log\det(T_\sigma T_\sigma^*). \tag{8.9}$$

The same formula holds for a surjection on rational vector spaces of unequal source and target ranks after replacing A by $A/\ker T$ with its quotient lattice and quotient metrics.

Proof. Put $M = T(A)$ with the induced metric from B. The isomorphism $A \simeq M$ has determinant norm $|\det T_\sigma|$ when determinants are taken in orthonormal coordinates. Comparing degree formulas gives $\widehat{\deg}_{\mathrm{Ar}}(M) = \widehat{\deg}_{\mathrm{Ar}}(A) - d_K^{-1} \sum_\sigma \log|\det T_\sigma|$. For the inclusion $M \subset B$, localization over each discrete valuation ring gives a list of elementary divisors. If their total length at w is $\ell_w(B/M)$, then the determinant ideal has that order and the local quotient has cardinality $(Nw)^{\ell_w(B/M)}$. Consequently $\widehat{\deg}_{\mathrm{Ar}}(B) - \widehat{\deg}_{\mathrm{Ar}}(M) = d_K^{-1}\log I$, where $\log I = \sum_w \ell_w(B/M)\log Nw$. Finally $\det(T_\sigma T_\sigma^*) = |\det T_\sigma|^2$. Multiply by $2/r$ to obtain (8.9). For the last assertion, $\ker T$ is saturated because the target is torsion free. The quotient is a projective lattice of rank r and its induced rational map to B is an isomorphism, so the same proof applies. At infinity the quotient identifies isometrically with the orthogonal complement of the kernel. □

The signs in this identity are useful to check. For multiplication by an integer $n \neq 0$ on $\mathbb{Z}^r$ with standard norms, the index is $|n|^r$ and $\det(TT^*) = |n|^{2r}$; the two terms cancel, as they must because the source and target lattices have not changed. If instead the target is defined to

be the image lattice $T(A)$, then the index term is zero and its slope changes by the negative logarithm of the determinant norm. If the metrics are fixed and the lattice is enlarged, only the positive index term remains. A lower bound for an analytic Gram determinant therefore has no independent arithmetic meaning until the direction of the map and the target lattice have been specified.

For the evaluation construction, the analytic map is first restricted to the finite-dimensional algebraic section space or its appropriate quotient. The source must identify that space, its integral section lattice, the evaluation lattice, and the metrics with which the analytic estimate is compatible. It must also bound any finite index introduced by descent, projection, or change of level. Once those identifications are made, (8.9) converts the Gram calculation into the required degree comparison, with all finite and infinite contributions visible. An estimate obtained for the full Bergman space does not automatically hold for an arbitrary smaller space of algebraic sections: the interpolants must belong to the space used in the comparison, or a controlled restriction argument must supply that passage. This is part of the geometric construction cited above, not a consequence of the finite-dimensional determinant identity.

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

$$\mu_\sharp \le \mu_0 + 3\delta_e(\mu_0 + h_0 + 28\lambda) + 28\lambda. \tag{8.10}$$

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 7.3 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.

The common row coordinates in Proposition 7.3 are essential here. In abstract Hilbert-space notation, suppose $A_u : H_u \to E_0$ are maps, U_u are unitary on H_u, V is one unitary on E_0, and $a_u \ne 0$. For $\widetilde{A}_u = a_u V A_u U_u$, direct multiplication gives

$$\widetilde{A}_u^* \widetilde{A}_v = \overline{a_u} a_v U_u^* A_u^* A_v U_v, \qquad \left\| \widetilde{A}_u^* \widetilde{A}_v \right\|_{\mathrm{HS}}^2 = |a_u|^2 |a_v|^2 \|A_u^* A_v\|_{\mathrm{HS}}^2. \tag{8.11}$$

Thus the scalar factors are removed by the prescribed positive normalizations, and the unitary factors leave the block estimates unchanged. The normalized total frame operator is conjugate by the single V to the original one, so its eigenvalues and trace variance are unchanged as well.

If there were separate row unitaries V_u, the cross block would instead contain $V_u^* V_v$ between A_u^* and A_v. Individual norm identities would survive, but the correlations would change. For example, two one-dimensional maps onto the two coordinate lines in $\mathbb{C}^2$ form an exact frame; applying a row unitary to only the second map can make both images equal to the first line. This is the same failure of coverage exhibited in Section 3. The common-row assertion in the Bessel comparison is what permits the spectral theorem to be used for the entire family of arithmetic maps. It is stronger than knowing that each period can individually be written in some unitary coordinates.

Proposition 8.9 (`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 $\eta_0, \eta_1, e_0 \geq 0$ such that

$$\begin{aligned} \|K_{uv}\|_{\mathrm{HS}}^2 &\leq (1+\eta_0) d_u d_v / D_* \qquad (u \neq v), \\ \|K_{uu} - I_{H_u}\|_{\mathrm{HS}}^2 &\leq e_0 d_u, \qquad\qquad D \leq (1+\eta_1) D_*, \\ \eta_0, \eta_1, \sqrt{e_0} &\leq \exp(-x^{\varepsilon/8}). \end{aligned} \tag{8.12}$$

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 (8.12) 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 8.10. *Assume the first three inequalities in (8.12). For nonnegative numbers p_u with $\sum_u p_u = 1$, put $T_u = (D p_u/d_u)^{1/2} A_u$ and $Q = \sum_u T_u T_u^*$. Then $|\operatorname{Tr} Q - D| \leq D\sqrt{e_0}$, and*

$$\frac{\|Q - I_{E_0}\|_{\mathrm{HS}}^2}{D} \leq (1+\eta_0)(1+\eta_1) - 1 + 2\sqrt{e_0} + D(1+\sqrt{e_0})^2 \sum_u \frac{p_u^2}{d_u}. \tag{8.13}$$

For q equally weighted columns with $d_u \geq D_*/(C_0\Gamma^2)$, 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 $|\operatorname{Tr} E_u| \leq \sqrt{d_u}\|E_u\|_{\mathrm{HS}} \leq d_u\sqrt{e_0}$. Since $\operatorname{Tr} Q = D\sum_u p_u \operatorname{Tr} K_{uu}/d_u$, this proves the trace bound. Also $\|K_{uu}\|_{\mathrm{HS}}^2 = d_u + 2\operatorname{Tr} E_u + \|E_u\|_{\mathrm{HS}}^2 \leq d_u(1+\sqrt{e_0})^2$. Expanding the square of Q and using cyclicity of the trace gives

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

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 $D/D_* \leq 1 + \eta_1$. For equal weights, $\sum_u p_u^2/d_u \leq C_0\Gamma^2/(qD_*)$, giving the final assertion. □

Among all probability weights on columns of given positive dimensions, the rank weights minimize the diagonal term in this estimate. Indeed, Cauchy–Schwarz gives

$$1 = \left(\sum_u \frac{p_u}{\sqrt{d_u}} \sqrt{d_u}\right)^2 \leq \left(\sum_u \frac{p_u^2}{d_u}\right)\left(\sum_u d_u\right). \tag{8.14}$$

Equality holds exactly when p_u is proportional to d_u, namely $p_u = d_u/R$ with $R = \sum_u d_u$. Thus the minimum is $1/R$. This is also the weighting required for pooling individual arithmetic slopes: every slope occupies one of the R source dimensions. The arithmetic and spectral reasons for choosing these weights are therefore compatible.

Equal weights are useful when the dimensions are comparable, as they are in the specified family. They can be inefficient without that hypothesis. For two columns of dimensions d and 1, equal weights give $\sum p_u^2/d_u = (1 + d^{-1})/4$, whereas rank weights give $1/(d+1)$. The ratio grows like $d/4$. A large column and a one-dimensional column should not have equal influence on a

calculation that counts individual source directions. The bounds in Proposition 8.9 keep this discrepancy uniformly controlled, but the rank-weighted version is the direct one for the slope application.

There is an elementary necessary dimension condition as well. Since the image of $Q_0 = \sum_u A_u A_u^*$ is contained in the sum of the images of the columns, its rank is at most R. The trace-variance condition from Proposition 3.3 therefore implies $R \geq D/(1+\eta_F)$ by the positive-eigenvalue calculation in Section 3. When each column has dimension comparable to D/Γ^2, at least a constant multiple of Γ^2 columns is needed even for approximate coverage of almost all target directions. The stronger cardinality bound below is a sufficient condition for the much smaller error required by the arithmetic comparison. It is obtained from the specific block estimates and is not asserted to be optimal among all possible frames.

To match Proposition 3.3, 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

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

For $e_0 \leq 1/4$, the hypotheses of Proposition 3.3 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 $\eta_F \geq 0$. If there are q columns, the dimension lower bound gives $D/R \leq C_0(1+\eta_1)\Gamma^2/q$. 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 $0 \leq \delta_{\mathrm{row}} \leq \exp(-x^{e_z/7}) < 1$, where $e_z > 0$ is fixed for the family. Thus (3.10) 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`.

The passage to the retained target is a dimension calculation with no unstated norm estimate. If $\pi : \widetilde{\mathcal{H}} \to F_\sigma$ is any surjective linear map and $S \subset \widetilde{\mathcal{H}}$, then the inverse image of $\pi(S)$ is $S + \ker \pi$. Hence

$$\operatorname{codim}_{F_\sigma} \pi(S) = \operatorname{codim}_{\widetilde{\mathcal{H}}} S - \dim\bigl(\ker \pi/(S \cap \ker \pi)\bigr) \leq \operatorname{codim}_{\widetilde{\mathcal{H}}} S. \tag{8.15}$$

Dividing by $D_{\mathrm{rows}} \geq (1-\delta_{\mathrm{row}})D_{\mathrm{all}}$ introduces the factor $(1-\delta_{\mathrm{row}})^{-1}$ and no other loss. For the high-slope subspaces of the columns, take $S = \sum_u \widetilde{B}_u(N_u)$. Compatibility of the lifts gives $\pi(S) = \sum_u B_u(N_u)$, which lies in the desired arithmetic filtration piece by Lemma 3.2. The frame estimate on the full spectral target therefore bounds the missing dimension in that filtration piece of the retained target.

The calculation does not require the projection to have arithmetic height zero, or even require the chosen auxiliary coordinates to agree with the arithmetic metrics. No map-height estimate is transported through this projection. The heights used by the slope theorem are already the heights of the integral maps into $R^\vee$ with their specified metrics. Only dimensions are transported from the larger spectral model. This distinction is what allows full spectral coverage and integral retention to be handled by separate estimates without making an unjustified claim that projection preserves the arithmetic lattice.

In particular, if the desired normalized error is $O(x^{-C_F})$, a sufficient cardinality is $q \geq C_0 \Gamma^2 x^{C_F}$, up to a fixed constant. The exponential terms in (8.12) 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 (8.13) 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 7 assumes $p \geq 5$. 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 [6]; the classification of p-divisible groups by admissible prismatic Dieudonné crystals is due to Anschütz–Le Bras [3]. 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

$$v_{i,\pi} \leq s_{i,\pi} + \mathcal{B}_{23} + \mathcal{V}(p) + h_E(p) + \epsilon(p), \qquad \mathcal{B}_{23} = Cfmb \log x + Cx^C L_2^2. \tag{8.16}$$

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 (8.16) 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

$$\eta_j \leq D_j^0 \left[C\{fmbL^2 + (\log x)^2 N_c\} + \sum_p (\log p) \sum_{0<e<t} \operatorname{length} \operatorname{Sha}_{p,e} \right], \tag{8.17}$$

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 $\eta_j \leq CD_0 x^{\varepsilon}(B_0 + N_c)$. This is how the small primes enter the common error budget. One must not insert the factor 2 from squared-norm slopes into (8.17); that factor enters only at the later arithmetic-degree conversion.

The conversion from local defects to a global degree is the finite part of Lemma 8.8. If two full lattices $M \subset N$ differ at finitely many places, then

$$\mu_2(N) - \mu_2(M) = \frac{2}{[K : \mathbb{Q}] \operatorname{rk} N} \sum_w \operatorname{length}_{\mathcal{O}_{K,w}} (N_w / M_w) \log Nw. \tag{8.18}$$

For a short exact sequence of finite local modules, lengths add; therefore a chain of integral comparisons adds its local defects. All terms are nonnegative when they measure an inclusion in

the stated direction. A comparison given only after inverting p loses these lengths entirely, even though they are the terms that control the integral slope loss. This is why rational comparison isomorphisms cannot replace the integral input at 2 and 3.

The series in (8.17) is slightly more general than the finite support in (8.18), because it packages defects throughout a family. Its terms are nonnegative, so finite partial sums increase. The construction may sum local bounds first and then pass to the limit once the cohomological estimate proves that these partial sums are uniformly bounded. Before that estimate, an expression involving the whole prime series is only a possibly infinite upper bound. No cancellation between prime defects is available to establish convergence. The subsequent cohomological argument supplies this missing global control, while the present section explains how that control is used in the lattice comparison.

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 8.5, the base-to-sharp loss is absorbed using (8.10), and spectral block estimates are converted to the required approximate frame by Lemma 8.10. Together with the cited construction inputs, they supply the geometric and spectral inequalities used in the final arithmetic comparison.

9. Cohomological bounds on a common window

The integral comparison in Section 4 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.

The output of the section is Theorem 9.10. It is consumed in one place, the bound for the integral comparison defect (the property `LinkN` in Proposition 11.2), where the admissible size is $\sqrt{\exp(x^{\varepsilon})}$. By (9.23) this is enormous compared with any fixed power of the conductor of the host, so bounds polynomial in the conductor and in the twist cutoff suffice. The difficulty lies in the uniformity and in a bounded number of exceptional twists, which can be controlled only on a window. The minus estimate (Proposition 9.2) and the fixed-layer bound (Proposition 9.4) hold at every scale; the window enters through the gap selection of Lemma 9.9.

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

$$\begin{aligned} \mathrm{Sha}_{\ell,e}(E) &= \ker\left(H^1(G_{E,S_\ell}, \mathbb{Q}_\ell/\mathbb{Z}_\ell(-e)) \longrightarrow \prod_{v\mid\ell} H^1(E_v, \mathbb{Q}_\ell/\mathbb{Z}_\ell(-e)) \right), \\ v_{\ell,e}(E) &= \mathrm{length}_{\mathbb{Z}_\ell} \mathrm{Sha}_{\ell,e}(E). \end{aligned} \tag{9.1}$$

The twist is negative: g acts on the coefficient module by $\chi_\ell(g)^{-e}$. The localization product in (9.1) 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

$$\sum_{\ell \in S} \log \ell \sum_{1 \leq e < t} v_{\ell,e}(E) \leq B \qquad \text{for every finite set } S \text{ of rational primes.} \tag{9.2}$$

This formulation avoids an unnecessary convergence convention. Since the summands are nonnegative, (9.2) 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.

The integrality of these lengths gives a little more than convergence. A finite $\mathbb{Z}_\ell$-module M has $|M| = \ell^{\operatorname{length} M}$: a composition series has successive quotients $\mathbb{F}_\ell$, and cardinalities multiply in short exact sequences. Conversely, a discrete ℓ-primary module of finite length is finite by the same argument. Consequently, for fixed E and t, condition (9.2) permits only finitely many nonzero groups among the $\mathrm{Sha}_{\ell,e}(E)$, $1 \leq e < t$. Indeed, any m nonzero pairs (ℓ, e) contribute at least $m \log 2$ to a finite prime sum. Thus their number is at most $\lfloor B / \log 2 \rfloor$. This observation does not produce a fixed finite set of primes for all the varying fields; it explains why a bound for every finite prime sum has its intended arithmetic meaning without requiring an interchange of infinite sums.

For a CM field, complex conjugation acts on (9.1). 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 (9.1) 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 (9.1). 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 4. The congruence-to-extension principle and generalized matrix algebras are classical [27, 5]; 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 extension interpretation can be made explicit in rank two. Let A be a commutative coefficient ring, let $\alpha, \beta : G \to A^\times$ be continuous characters, and consider an extension of the free rank-one module with character β by that with character α. After choosing a lift of a basis of the quotient, its matrices have upper-right entry $b(g)$. Matrix multiplication says that $b(gh) = \alpha(g)b(h) + b(g)\beta(h)$. Therefore $z(g) = b(g)\beta(g)^{-1}$ satisfies the cocycle identity for the character $\alpha\beta^{-1}$. Replacing the lift by itself plus u times the basis of the submodule replaces $z(g)$

by $z(g) + u(\alpha(g)\beta(g)^{-1} - 1)$. The extension class is accordingly an element of $H^1(G, A(\alpha\beta^{-1}))$, and it vanishes exactly when the extension splits. The same calculation after restriction to a decomposition group gives the local splitting criterion.

This calculation determines the sign of the twist, as well as its magnitude. If the subcharacter divided by the quotient character is χ_ℓ^{-e}, the coefficient module is the negative twist used in (9.1). Reversing the extension reverses that ratio. For finite coefficient quotients one performs this calculation at each quotient and then uses the compatible coefficient maps. Passing to $\mathbb{Q}_\ell/\mathbb{Z}_\ell(-e)$, and imposing unramifiedness outside ℓ and local splitting above ℓ, gives precisely the kernel in (9.1). Neither the choice of an upper-triangular notation nor a convention for Hodge weights can replace this ratio calculation.

A simple order illustrates the integral issue. If $\mathcal{O}$ is a discrete valuation ring with uniformizer ϖ and fraction field K, the subring $A_r = \{(x, y) \in \mathcal{O}^2 : x \equiv y \pmod{\varpi^r}\}$ has $A_r \otimes K = K^2$ and $\operatorname{length}_{\mathcal{O}}(\mathcal{O}^2/A_r) = r$. The quotient is identified with $\mathcal{O}/\varpi^r$ by $(x, y) \mapsto x - y$. The two rational idempotents separate the two characteristic-zero eigensystems, but they do not preserve A_r when $r > 0$. Hence even a completely split characteristic-zero algebra can have an arbitrarily large integral splitting defect. Semisimplicity of the associated rational representation alone places no bound on that defect.

The length step following a genuine congruence-to-extension comparison is elementary. Suppose a finite module D has a finite filtration whose successive quotients inject into specified finite modules M_i. Additivity and monotonicity of length give $\operatorname{length} D \le \sum_i \operatorname{length} M_i$. If some quotients are only killed by ϖ^{a_i} and generated by b_i elements, their contribution is at most $a_i b_i$. This is the form in which extension groups and projector errors can control a lattice index. The mathematical content of the comparison is the construction of the filtration and the injections, together with their local conditions and multiplicities. The rank-two cocycle calculation explains the targets of those maps; it does not construct them for an arbitrary Hecke algebra. Keeping these two steps distinct is necessary because the later bound is a sum over many character differences and many residue characteristics.

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 $h \ge 2$, write

$$E_h = F_\psi\, \mathbb{Q}(\zeta_{2^{h+2}})^+\, \mathbb{Q}(\sqrt{-q})\, \mathbb{Q}(\sqrt{-d}), \qquad N_h = \operatorname{lcm}(N_\psi, 2^{h+2})qd. \tag{9.3}$$

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 9.1. *Let M be a nonzero integer and let $a_\ell \ge 0$ satisfy $a_\ell \le \operatorname{ord}_\ell(M)$ 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 $a_\ell \log \ell \le \sum_{\alpha\in\mathcal{A}} -\log|\alpha|_\ell$, 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 $r_1 \ne r_2$. □

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.

There are two useful elementary models for the denominator bookkeeping. First, let M be a free $\mathcal{O}$-module of rank r carrying an action of a finite abelian group A, and suppose the fraction field of $\mathcal{O}$ contains the values of all its characters. Write $M_\vartheta = M \cap e_\vartheta(M \otimes K)$, where $e_\vartheta = |A|^{-1} \sum_{g \in A} \vartheta(g)^{-1} g$. Character orthogonality gives a direct sum over K. Since character values are units of $\mathcal{O}$, each $|A| e_\vartheta$ preserves M. Thus

$$|A|M \;\subseteq\; \bigoplus_{\vartheta} M_\vartheta \;\subseteq\; M, \qquad \operatorname{length}_{\mathcal{O}} \left(M / \bigoplus_{\vartheta} M_\vartheta \right) \le r \operatorname{ord}_{\varpi}(|A|). \tag{9.4}$$

For the inequality, the quotient is a quotient of $M/|A|M$. In particular, the character decomposition is integral when the residue characteristic does not divide $|A|$. This proof explains the special role of primes dividing the degree of an abelian field. It gives a denominator bound for a given lattice; it does not assert that the corresponding arithmetic eigenspaces satisfy the special-value estimate.

Second, if L is a number field and $\alpha \in \mathcal{O}_L$ is nonzero, then $N_{L/\mathbb{Q}}(\alpha)$ is a nonzero integer and its logarithm is $\sum_{\sigma : L \hookrightarrow \mathbb{C}} \log |\sigma(\alpha)|$. Bounds $|\sigma(\alpha)| \le A_\sigma$ therefore bound the logarithm of its absolute norm by $\sum_\sigma \log A_\sigma$. If α is only algebraic and $D\alpha$ is integral for an integer $D \ge 1$, the same argument applied to $D\alpha$ adds $[L : \mathbb{Q}] \log D$. The added denominator cost must remain in the estimate. In a matrix calculation, Smith normal form gives the corresponding local rule: an injective map between free $\mathcal{O}$-modules of the same rank, represented by a matrix T with nonzero determinant, has cokernel length $\operatorname{ord}_{\varpi}(\det T)$. Indeed, invertible row and column operations diagonalize T with diagonal entries units times ϖ^{a_i}, and both quantities equal $\sum_i a_i$. These norm and determinant identities are the elementary reason that an integral nonzero automorphic comparison can control a sum of local lengths.

Proposition 9.2 (Minus estimate; `thmBcFree_of`)**.** *The cyclotomic-character and integral annihilator inputs of the source imply the existence of an absolute constant $C \ge 0$ such that, for every field (9.3), every $h \ge 2$, and every integer $t \ge 2$, the odd-prime minus contribution is finite and bounded by*

$$3 n_E t^2 \log(tF) + C \sum_{\substack{\ell \mid o \\ \ell \neq 2}} \ell n_E^2 t^2 \log(t\ell F)^2. \tag{9.5}$$

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 9.1, applied for each $1 \le e < t$, bounds the prime sum. The source's explicit denominator and factorial estimate yields the first term in (9.5) 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 (9.3) 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 (9.5). 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 $n_E \ge 1$, $tF \ge 2$ 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 [20, 10]. Those references explain the provenance of the input; they do not, without its local conditions and denominator calculation, establish (9.5). The contribution relevant here is the uniform arrangement of this theory for the changing fields (9.3) 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 9.3. *If A is a finite abelian group, then $\sum_{H \leq A,\ A/H \text{ cyclic}} \varphi([A : H]) = |A|$. 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 $30N\,|\,\operatorname{disc} K_*|$, the input bounds the relevant finite prime sum by

$$4c_5\varphi([A:H])|\operatorname{disc} K_*|^5 N^{18} e^7(\log r_1 + \log r_2). \tag{9.6}$$

The precise character and local hypotheses are part of that input. Lemma 9.3 absorbs the orbit factors, and Lemma 9.1 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 (9.6) 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 [19]; 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 [16, 32]; 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 9.4 (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*

$$\log|H^2(G_{L,S_\ell}, \mathbb{Z}_\ell(e+1))| \leq \exp((2e[L:\mathbb{Q}])^{C_K} \log(2|\operatorname{disc} L|)). \tag{9.7}$$

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 (9.1), 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 $N = 4e + 6 = 2(k+1) + 4$.

We give the passage from this rank assertion to the required finite count. It is useful to see exactly where the two adjacent homology degrees enter. For a group G and a commutative ring A, let $C_m(G, A)$ be the free A-module on symbols $[g_1|\cdots|g_m]$. Its boundary is

$$\partial[g_1|\cdots|g_m] = [g_2|\cdots|g_m] + \sum_{i=1}^{m-1}(-1)^i[g_1|\cdots|g_ig_{i+1}|\cdots|g_m] + (-1)^m[g_1|\cdots|g_{m-1}]. \tag{9.8}$$

Each term obtained by applying two face maps occurs twice with opposite signs, so $\partial^2 = 0$. Applying a group homomorphism to every entry commutes with ∂. These formulas provide the bar complex for group homology with trivial coefficients, including its integral and reduced-coefficient versions.

For later use, conjugating every entry has an explicit chain homotopy to the identity. In the convention of the source put

$$P_h[g_1|\cdots|g_m] = \sum_{i=0}^{m}(-1)^i[g_1|\cdots|g_i|h|h^{-1}g_{i+1}h|\cdots|h^{-1}g_mh]. \tag{9.9}$$

Then $\partial P_h + P_h\partial = (\mathrm{conj}_{h^{-1}})_* - 1$. To check this identity, expand both sides using (9.8). A face strictly before or strictly after the inserted entry is cancelled by the matching summand of $P_h\partial$. The two faces adjoining the inserted entry cancel between consecutive values of i. Only the first face at $i = 0$ and the last face at $i = m$ remain; they are respectively the conjugated chain and minus the original chain. This also proves naturality under a group homomorphism, since each displayed entry is expressed by multiplication and inversion. Thus the conjugation manipulations used to join arithmetic cycles have a literal chain-level meaning; they are not identifications made only after forgetting integral coefficients.

Lemma 9.5. *Let $C_\bullet$ be a chain complex of free abelian groups and $Q \geq 1$. There is a natural exact sequence*

$$0 \longrightarrow H_m(C_\bullet)/QH_m(C_\bullet) \longrightarrow H_m(C_\bullet/QC_\bullet) \longrightarrow H_{m-1}(C_\bullet)[Q] \longrightarrow 0. \tag{9.10}$$

In particular, a reduced cycle with zero image on the right is the reduction of an integral cycle, after alteration by a reduced boundary; at the level of cycles, zero Bockstein permits an integral cycle lifting the reduced cycle itself.

Proof. Lift a reduced cycle $\bar{z}$ to $z \in C_m$. Its boundary is Qy for some $y \in C_{m-1}$. Since $Q\partial y = \partial^2 z = 0$ and C_{m-2} is torsion-free, y is a cycle. Associate the class $[y]$ to $\bar{z}$. A different lift changes y by an integral boundary, and adding a reduced boundary does not change its class. Also $Q[y] = [\partial z] = 0$. Every class killed by Q arises this way: if $Q[y] = 0$, choose z with $\partial z = Qy$.

If $[y] = 0$, write $y = \partial w$ and replace z by $z - Qw$. The replacement is an integral cycle with precisely the original reduction. This proves exactness at the middle term. Finally, if an integral cycle x reduces to a boundary, write $x = \partial u + Qv$. Applying ∂ and using torsion-freeness shows that v is a cycle. Thus $[x] = Q[v]$, proving injectivity of the first arrow and exactness there. All constructions commute with chain maps. □

The elementary rank count is as follows. Suppose an abelian group H has finite torsion subgroup of order T, a subgroup J has rank at most r, and $f : H \to B$ is a homomorphism with $QB = 0$. Every finite subset of $f(J)$ has cardinality at most $Q^r T$. Choose a preimage of each element of the finite subset and let J_0 be the subgroup they generate. The structure theorem gives $J_0 \simeq \mathbb{Z}^{r_0} \oplus T_0$, with $r_0 \le r$ and $|T_0| \le T$, because J_0 is a subgroup of H. The restriction of f factors through J_0/QJ_0, whose cardinality is at most $Q^r T$. The reference to the ambient torsion subgroup is essential: the torsion of an image of a free abelian group need not be trivial.

Now let $Q = \ell^\nu$, and let $\Phi_N : C_{k+1}(\mathrm{GL}_N(\mathcal{O}_L), \mathbb{Z}/Q) \to B$ be linear maps to a $\mathbb{Z}/Q$-module. Assume that they vanish on boundaries, agree on stabilized cycles, and satisfy $\Phi_{dN}(\Psi_d z) = d\Phi_N(z)$ on cycles. Reducing integral cycles and applying Φ_N gives homomorphisms ϕ_N on integral homology with the same properties. For the filtered classes introduced above, induction gives

$$\phi_{\ell^j N}(z_j) = \ell^j \prod_{i=1}^{j} (1 - \ell^i)\, \phi_N(x). \tag{9.11}$$

Indeed the next difference multiplies the previous value by $\ell - \ell^{j+2}$. The product in (9.11) is a unit modulo Q. Consequently every value $\ell^k \phi_N(x)$ is the value of $\phi_{\ell^k N}$ on an integer multiple of z_k. This is the exact reason for the power ℓ^k in the finite count: multiplication by it moves the relevant values into the filtered subgroup to which the odd-rank bound applies.

To count a finite set of values $\ell^k \Phi_N(z)$, first stabilize all its cycles to a common size $N \ge 2(k+1)+4$. The Bockstein map sends these cycles to $H_k(\mathrm{GL}_N(\mathcal{O}_L), \mathbb{Z})[Q]$. Within one Bockstein fiber, the difference of two cycles has zero Bockstein and therefore lifts to an integral cycle by Lemma 9.5. Formula (9.11) and the rank count bound the number of different values in that fiber by Q^{r_2} times the torsion order in $H_{k+1}(\mathrm{GL}_{\ell^k N}(\mathcal{O}_L), \mathbb{Z})$.

Here one needs integral homological stability, not only rational stability. The source proves that stabilization from size N_0 to any larger size induces an isomorphism in degree j when $N_0 \ge 2j + 4$. In particular, the torsion orders in degrees k and $k+1$, at all sizes just used, are respectively $T_{2(k+1)+4,k}(L)$ and $T_{2(k+1)+4,k+1}(L)$. The former bounds the number of Bockstein fibers and the latter bounds the torsion contribution within each fiber. Multiplying the bounds, and taking $k = 2e$, proves that every finite set of the required values has cardinality at most

$$(\ell^\nu)^{r_2} T_{4e+6,2e+1}(L) T_{4e+6,2e}(L), \qquad 2e + 1 < r_2. \tag{9.12}$$

This proof remains valid when B is an infinite product of finite coefficient modules: only the chosen finite set of values is counted. More explicitly, for homomorphisms $\beta : Y \to T$ and $f : Y \to B$ with T finite, subtract one representative in each β-fiber. If every finite subset of $f(\ker \beta)$ has at most C elements, a finite subset of $f(Y)$ has at most $C|T|$ elements. This is the finite-fiber argument used above, and requires no finiteness assertion about Y or B.

The proof just given is the assembly in `SUpperOf.lean`. Its Bockstein, filtered-value, and stability steps are respectively `BockAbs.lean`, `PhiZSeq.lean`, and `TorsStab.lean`. The rank bound and integral homological stability are the substantive homological inputs to this reduction. A quantitative estimate for the two torsion orders is a further input, used only after the class-character count has removed its layer-dependent factor.

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

$$|X_\nu| \leq T_{4e+6,2e+1}(L)T_{4e+6,2e}(L)\ell^{(B_e+v_\ell(e))g+c+v_\ell(e)}. \tag{9.13}$$

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. We now explain how the cycle comparison, the detection argument, and the class-group count combine to give it. The powers and the degree of the cycle will be kept explicit, since changing one of them changes the input being used.

Put $Q = \ell^\nu$, $\kappa = 1+\ell^c$, $M = 2e+1$, and $a = \kappa^e$. A divisor below is a finitely supported integer combination of the finite primes of F. The twisted principal-divisor relation attached to $\alpha \in F^\times$ and a divisor $\mathfrak{d}$ is

$$\mathrm{ord}_P(\alpha) = \mathfrak{d}(\gamma P) - a\mathfrak{d}(P) \quad \text{for every finite prime } P. \tag{9.14}$$

The inverse in the character condition defining X_ν is consistent with this formula: the coefficient of P in $\gamma^{-1}\mathfrak{d}$ is $\mathfrak{d}(\gamma P)$. Define the weighted norm and the interpolation factor by

$$\mathcal{N}_a(\alpha) = \prod_{j=0}^{[F:L]-1} \gamma^j(\alpha)^{a^j}, \qquad D_e = \prod_{\substack{0 \leq j < M \\ j \neq e}} (\kappa^e - \kappa^j). \tag{9.15}$$

These expressions have different functions. The weighted norm records the equivariance of a divisor relation, while D_e selects one weight from the finite set of powers κ^j. It is not legitimate to discard D_e as a unit: it is usually highly divisible by ℓ.

Here is the precise cycle-specialization input. For every relation (9.14), the source constructs an integer N, a cycle $y_\alpha \in C_{2e+1}(\mathrm{GL}_N(\mathcal{O}_L), \mathbb{Z}/Q)$, and a nonzero $b_\alpha \in \mathcal{O}_F$. Suppose $r : \mathcal{O}_F \to k$ maps to a finite field, its restriction to $\mathcal{O}_L$ is surjective, and $r(b_\alpha) \neq 0$. For every character $v : k^\times \to \mathbb{Z}/Q$ and every compatible additive evaluation family θ, one has

$$\theta_{k^N}(r_* y_\alpha) = D_e\, v(\overline{\mathcal{N}_a(\alpha)}). \tag{9.16}$$

The bar on the right means reduction of the fraction after a denominator prime to the chosen kernel of r has been selected; existence of this reduction is part of the assertion. An additive family here assigns a boundary-vanishing functional on degree-$2e+1$ chains of $\mathrm{GL}(V)$ to each finite-dimensional k-vector space V, is unchanged by linear isomorphism, and is additive on direct sums of representations evaluated on cycles. Its normalization is that a prism around the scalar character $u \mapsto r(\zeta)^{xu}$ evaluates to $x^e v(\det g)$, where g is the prism element. The source constructs such families for finite fields containing a primitive Qth root. Additivity makes their pullbacks to $\mathrm{GL}_N(\mathcal{O}_L)$ primitive in exactly the sense required for (9.12).

Several concrete constructions enter (9.16). The cyclic group $C_Q = \mathbb{Z}/Q$ has the degree-$2e$ cycle $c_e = \sum_{j_1,\dots,j_e \in \mathbb{Z}/Q} [j_1|1|\cdots|j_e|1]$, with coefficients in $\mathbb{Z}/Q$. Applying (9.8), the interior terms cancel by translating one of the j_i; the remaining end contribution is a multiple of Q and vanishes. Thus $\partial c_e = 0$. The maps $d_t : C_Q \to C_Q^2$, $u \mapsto (u, tu)$, for $0 \leq t \leq 2e+1$, supply $2e+2$ images of this cycle.

Their homology classes satisfy a relation with at least one coefficient a unit modulo Q. One way to see the numerical reason is to use the periodic free resolution of a cyclic group, whose maps alternate between multiplication by a generator minus one and its norm element. Tensoring two such resolutions gives a resolution for C_Q^2; after tensoring with the trivial module $\mathbb{Z}/Q$, all differentials vanish and the degree-$2e$ term has $2e+1$ summands. Thus its homology is generated by $2e+1$ elements. Lift the $2e+2$ specified classes to a free abelian group on those generators. Rational linear dependence gives an integer relation, and division by the common power of ℓ in its coefficients leaves a relation with a coefficient prime to ℓ. Mapping the relation back gives a boundary relation between the $(d_t)_* c_e$. The periodic resolution used here is exact because the

kernel of multiplication by the generator minus one consists of multiples of the norm, and the kernel of the norm consists of elements whose coefficients sum to zero. These equal the two successive images. Its tensor product remains a resolution because its augmented complex is split exact as a complex of abelian groups. This explains the relation used by the source without requiring any assumption about the arithmetic homology group being counted.

A second elementary construction isolates the index carrying that unit coefficient. We include its proof because it is a useful explicit part of the cycle argument.

Lemma 9.6. *Let $\nu \geq c \geq 1$, let $M < \ell^c$, and choose $t_0 \in \{0, \dots, M\}$. Write $P_M(X) = \prod_{j=0}^{M-1}(X - \kappa^j) = \sum_{i=0}^{M} p_i X^i$, with $\kappa = 1 + \ell^c$. There is a finitely supported integer function η on $(\mathbb{Z}/\ell^\nu)^2$ such that its pushforward by $(x, y) \mapsto x + ty$ is zero for every $t \neq t_0$ in this interval, and its pushforward for $t = t_0$ is $\sum_{i=0}^{M} p_i \delta_{\kappa^i}$.*

Proof. For an additive finite group A, write $\mathbb{Z}[A]$ for its integral group ring, with basis δ_x and multiplication $\delta_x \delta_y = \delta_{x+y}$. Put $w = \ell^c$ and let E act on $\mathbb{Z}[T]$ by $E(f)(T) = f(T^{1+w})$. If $(T^w - 1)^j$ divides f, then $(T^w - 1)^{j+1}$ divides $E(f) - (1+w)^j f$. To prove this, write $f = (T^w - 1)^j g$. The quotient $E(T^w - 1)/(T^w - 1)$ is the geometric sum $\sum_{i=0}^{w} T^{wi}$, congruent to $1 + w$ modulo $T^w - 1$, and $E(g) \equiv g$ modulo the same polynomial. Subtraction gives the claimed extra factor. Apply this successively for $j = 0, \dots, M - 1$ to T. The commuting operators $E - \kappa^j$ show that $(T^w - 1)^M$ divides $\sum_i p_i T^{\kappa^i}$.

Evaluate T at δ_1 in $\mathbb{Z}[\mathbb{Z}/Q]$, where $Q = \ell^\nu$. The resulting element $\lambda = \sum_i p_i \delta_{\kappa^i}$ is divisible by $(\delta_w - 1)^M$. For each $t \neq t_0$, the nonzero integer $t_0 - t$ has absolute value below ℓ^c, so its ℓ-part divides ℓ^c. The congruence $(t_0 - t)b = w \pmod{Q}$ is therefore soluble. A geometric sum shows that $\delta_{t_0 - t} - 1$ divides $\delta_w - 1$. Consequently $\lambda = (\prod_{t \neq t_0}(\delta_{t_0 - t} - 1))u$ for some $u \in \mathbb{Z}[\mathbb{Z}/Q]$.

Embed this group ring into $\mathbb{Z}[(\mathbb{Z}/Q)^2]$ along the first coordinate, and denote the embedding by ι. Take

$$\eta = \prod_{t \neq t_0} (\delta_{(0,1)} - \delta_{(t,0)})\, \iota \left(\delta_{-\sum_{t \neq t_0} t\, u} \right). \tag{9.17}$$

Pushforward for an index $t \neq t_0$ kills its corresponding factor. Pushforward for t_0 replaces that factor by $\delta_t(\delta_{t_0 - t} - 1)$; the product of the δ_t is cancelled by the displayed translation. The resulting expression is λ, as required. □

This is the finite-difference construction in `ExpStep.lean`. It works even when some differences $t - t_0$ are divisible by ℓ; replacing those differences by units would invalidate the proof. The strict inequality $M < \ell^c$ is precisely what permits their divisibility by ℓ^c in the cyclic coefficient ring.

The arithmetic step then realizes the integer coefficients of η as exponents of fractional ideal lattices. The divisor relation (9.14) gives the determinant needed to carry one such lattice to another. The source uses the classification of projective modules over a Dedekind domain to construct the carrying maps with that determinant, rather than choosing arbitrary rational matrices. The boundary relation between the $(d_t)_* c_e$ and the prism identity (9.9) join their prisms into a cycle in a lattice stabilizer. A finite projective lattice becomes a direct summand of a finite free lattice, so adjoining a complement gives a homomorphism from that stabilizer to some $\mathrm{GL}_N(\mathcal{O}_L)$. Additivity of the evaluation family removes the trivial complement. After excluding the finitely many primes in one nonzero denominator, reduction splits the relevant arithmetic representation into its $\mathrm{Gal}(F/L)$ and character blocks. Evaluating each block gives x^e times its determinant character. The distribution kills every unwanted block, while the chosen block contributes $P_M(X)/(X - \kappa^e)$ evaluated at $X = \kappa^e$, which is D_e. The Galois blocks supply the powers in $\mathcal{N}_a$.

This paragraph describes a specific chain and lattice construction whose full carrying-map and reduction identities remain part of the cycle-specialization input. Its assembly is `CJoin.lean`, invoked by `SClassOf.lean`. The elementary homological relation and Lemma 9.6 explain two of its mechanisms. They do not replace the integral lattice realization or the finite-field evaluation

theorem. In particular, a rational cycle without the asserted integral realization would not yield (9.16) at every prime required for detection.

The cost of the interpolation factor admits a simple bound. For $i \neq j$ in $\{0, \dots, M-1\}$, one has $v_\ell(\kappa^i - \kappa^j) = c + v_\ell(i-j)$. Here $c \geq 2$ when $\ell = 2$. To verify the formula, write $|i-j| = \ell^r u$ with $\ell \nmid u$. The binomial expansion first gives $v_\ell((1+\ell^c)^u - 1) = c$, and taking an ℓth power increases the valuation by one. For odd ℓ the intermediate binomial terms have strictly greater valuation than the first term; for $\ell = 2$ the same conclusion uses $c \geq 2$. Iterating proves the formula. Since $|i-j| < M$, the valuation is below $c + M$. There are $M-1$ factors in D_e, so D_e divides $\ell^{M(c+M)}$ in $\mathbb{Z}_\ell$, and also modulo Q. This deliberately loose bound is sufficient and is uniform in ν.

Lemma 9.7. *Let F be a number field, let $Q \geq 1$ be an integer with $\mu_Q \subset F$, and let $\beta \in F^\times$. If, outside a fixed finite set of finite primes of F, the reduction of β is a Qth power at every prime of absolute residue degree one, then β is a Qth power in F.*

Proof. Let $K = F(\beta^{1/Q})$. Since $\mu_Q \subset F$, the extension K/F is Galois and its automorphisms send a chosen root to a Qth root of unity times that root. Enlarge the excluded set to contain primes dividing Q, the numerator or denominator of β, and the ramified primes of K/F. At any remaining prime where β is a Qth power in the residue field, $X^Q - \beta$ has a root in that field. Its roots are distinct, and all Qth roots of unity are already in the residue field, so it splits completely there. Frobenius is therefore trivial in $\mathrm{Gal}(K/F)$.

Chebotarev supplies primes with Frobenius in any specified conjugacy class [9], and these can be chosen of absolute residue degree one. For the latter qualification, prime ideals of degree at least two have Dirichlet density zero: their norm is at least the square of the underlying rational prime, so the associated series stays bounded as its exponent tends to one. Removing them and a finite set from a positive-density Chebotarev set leaves a nonempty set. Every element of $\mathrm{Gal}(K/F)$ must therefore be trivial. Hence $K = F$, proving the assertion. □

This is the mathematical detection argument in `SDetectCut.lean`. It uses Chebotarev, rather than a claim that arbitrary local power conditions automatically imply a global power. Absolute residue degree one also explains why the reduction map from $\mathcal{O}_L$ is surjective: its image contains the prime field, which is the entire residue field in that case. Thus these primes belong to the precise family of reductions allowed in (9.16).

We can now prove the finite separation bound used by the count. Let S be a finite set of elements $\alpha \in F^\times$ admitting relations (9.14), with the following separation property: if $\mathcal{N}_a(\alpha)^{\ell^{Be}} = \mathcal{N}_a(\alpha')^{\ell^{Be}} \delta^Q$ for some $\delta \in F^\times$, then $\alpha = \alpha'$. For every $\alpha \in S$, select its cycle and its nonzero denominator. Exclude the primes dividing $\ell \prod_{\alpha \in S} b_\alpha$, and take all remaining finite-field reductions allowed above. The reduced root of unity still has order Q. Since the multiplicative group of a finite field is cyclic, there is a character $v : k^\times \to \mathbb{Z}/Q$ whose kernel is precisely $(k^\times)^Q$. Choose the evaluation family for this character at each reduction and take the product of all the resulting primitive functionals. This product is permitted by (9.12), even though the set of reductions can be infinite.

Send α to the tuple of values $\ell^{2e}\theta(r_* y_\alpha)$. If two tuples are equal, (9.16) gives $\ell^{2e} D_e(v(\overline{\mathcal{N}_a(\alpha)}) - v(\overline{\mathcal{N}_a(\alpha')})) = 0$ at each reduction. Multiply by a coefficient expressing the divisibility $D_e \mid \ell^{M(c+M)}$. It follows that the reduction of $(\mathcal{N}_a(\alpha)/\mathcal{N}_a(\alpha'))^{\ell^{Be}}$ lies in the kernel of v and hence is a Qth power. Lemma 9.7 makes it a global Qth power. The separation property gives $\alpha = \alpha'$, so the tuple map is injective. Formula (9.12) now gives

$$|S| \leq Q^{r_2} T_{4e+6,2e+1}(L) T_{4e+6,2e}(L). \tag{9.18}$$

The equality $B_e = 2e + M(c+M)$ has a transparent origin: $2e$ is the homological filtering loss, and the second term clears the interpolation factor. This argument is the specialization and detection assembly in `SFewOfCut.lean`.

The remaining class-group count has a different direction. Its exact assertion is that a bound B for every separated finite set of the kind just considered implies

$$|X_\nu|\, Q^{r_2} \leq B\, \ell^{(B_e + v_\ell(e))g + c + v_\ell(e)}. \tag{9.19}$$

The finiteness of X_ν is included. The comparison is assembled in `SelmerOrderCounts.lean`. It uses the divisor sequence, units, the class group, and the kernel of the weighted norm. The bound $g = 2n + \lfloor \log_2 h_L \rfloor + 2$ controls a required mod-ℓ generator count. For a finite abelian group C, the elementary relation $|\operatorname{Hom}(C, \mathbb{Z}/Q)| = |C/QC|$ follows by writing its ℓ-primary part as a sum of cyclic groups, each of which contributes $\ell^{\min(\nu,r)}$ on both sides. Thus the equivariant characters count a specific twisted quotient of the class group; they are not arbitrary characters of an unrelated finite group. The unit and norm comparison is what supplies the factor Q^{r_2} on the left of (9.19).

The finite-group estimate within the class count is worth isolating. If a finite abelian ℓ-group W satisfies $|W[\ell]| \le \ell^g$, then $|W[\ell^r]| \le \ell^{rg}$ for every $r \ge 0$. The structure theorem writes $W = \bigoplus_{i=1}^{t} \mathbb{Z}/\ell^{a_i}$ with $a_i \ge 1$. The first hypothesis says $t \le g$, and the group killed by ℓ^r has order $\ell^{\sum_i \min(r,a_i)} \le \ell^{rg}$. Consequently, if a homomorphism from W has kernel killed by ℓ^r and image of order at most B, then $|W| \le B\ell^{rg}$. More generally, if the kernel is killed by $m\ell^r$ with m prime to ℓ, the same conclusion holds, since multiplication by m is an automorphism of W.

This explains the exponent involving the generator bound in (9.19). In the source's divisor-and-unit comparison, the weighted-norm map has a kernel whose ℓ-part is killed after a loss $\ell^{B_e+v_\ell(e)}$, while a separate unit comparison supplies the remaining $\ell^{c+v_\ell(e)}$ factor. The valuation $v_\ell(e)$ is forced by the cyclotomic layer relation. Indeed, the binomial valuation calculation already used above gives $v_\ell(a-1) = c + v_\ell(e)$ and $v_\ell(a^{\ell^{\nu-c}} - 1) = \nu + v_\ell(e)$. The hypotheses on c at the dyadic prime are again necessary for these formulas. The prime-to-ℓ part of the latter integer acts invertibly on an ℓ-primary finite group; the finite-group estimate controls the remaining kernel. Identifying this group and proving the claimed kernel annihilation are the arithmetic content of the source comparison. The elementary estimate shows how those exact assertions, once supplied, lead to the displayed powers, and why neither an unspecified finite kernel nor an unspecified finite generator count would suffice.

Substitute the value of B from (9.18) into (9.19). Since $Q^{r_2} > 0$, it cancels and gives exactly (9.13). This cancellation is the reason for the filtered rank r_2 in the homological argument and for retaining the unit contribution in the class count. Bounding the whole class group crudely at each cyclotomic level would retain an uncontrolled dependence on ν and would not give this conclusion. The interfaces for the cycle construction, detection, finite separation, and class count are specified together in `LemmaKOfSteps.lean`; their proofs are inserted, together with the odd-homology input, in `LemmaKOfLitN.lean`.

The passage through all cyclic layers gives an integer $s \ge 0$ such that $v_{\ell,e}(L) = s$ and ℓ^s is at most the right-hand side of (9.13). This includes finiteness of the length. The finiteness assertion deserves emphasis. A discrete ℓ-primary group A is the union of its subgroups $A[\ell^\nu]$. If their cardinalities are uniformly bounded, these cardinalities are nondecreasing integers and hence eventually constant; the nested subgroups then stabilize and A is finite. This elementary observation explains why a layer-independent bound can imply finiteness, whereas a separate finite bound at each layer cannot. The arithmetic passage used here is more than that observation: it must compare the equivariant class characters with the global cohomology kernel, preserving the local conditions above ℓ. The source's comparison quantifies over every cyclic layer with the prescribed generator and root, and turns their uniform bound into the equality $v_{\ell,e}(L) = s \in \mathbb{N}$ and the inequality $\ell^s \le B$. It is not an assertion that the groups X_ν are literally the subgroups $\mathrm{Sha}_{\ell,e}(L)[\ell^\nu]$ with identical transition maps. That arithmetic comparison is the precise input recorded as `SSha`.

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 $c' \ge 1$ and $c' \ge 2$ if $\ell = 2$. They satisfy $2e+1 < \ell^{c'}$, $2e+1 < r_2(L')$, $n' = [L' : \mathbb{Q}] = dn$, $d \le (n+1)(4e+4)$, $\ell^{c'} \le 2n'$, and $|\operatorname{disc} L'| \le |\operatorname{disc} L|^d (\ell^{c'})^{n'}$. Moreover $v_{\ell,e}(L) \le v_{\ell,e}(L')$. 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 $j \le 2e+1$, $\log T_{4e+6,j}(L) \le \exp((2en)^C \log(2|\operatorname{disc} L|))$. 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, $v_{\ell,e}(L)\log\ell \le \exp((2en)^C \log(2|\operatorname{disc} L|))$. The numerical reduction is `NumThree.lean`.

For completeness, the form of this numerical absorption can be checked without tracking the very large absolute exponent. Put $D = |\operatorname{disc} L|$ and $D' = |\operatorname{disc} L'|$. The displayed extension bounds give $n' \le 16en^2$, $d \le 16en$, and $\log(2D') \le \log 2 + d\log D + n'\log(2n')$. Since $e, n \ge 1$ and $\log(2D) \ge \log 2$, there is an absolute C_1 for which this is at most $(2en)^{C_1}\log(2D)$. Likewise $2en'$ is at most a fixed power of $2en$. Thus the torsion estimate evaluated over L' is bounded by the same form of exponential over L, after increasing its exponent. One cannot simply replace D' by D, since the additional cyclotomic roots ramify; the factor $(\ell^{c'})^{n'}$ is what makes this passage valid.

The remaining factor in (9.13) is much smaller. From $\ell^{c'} \le 2n'$ one obtains $c'\log\ell \le \log(2n')$, while $v_\ell(e) \le \log_2 e$. The class-number input gives a polynomial bound for $\log h_{L'}$ in n' and $\log(2D')$. Even a coarse version suffices: Minkowski's ideal-class bound gives a representative of norm at most $2^{n'}\sqrt{D'}$ in every ideal class. The number of ideals of norm m is bounded by the coefficient of m^{-s} in $\zeta(s)^{n'}$, because at each rational prime there are at most n' prime ideals and every residue degree is at least one. Summing those coefficients up to Y counts tuples of n' positive integers with product at most Y, and is at most $\lfloor Y\rfloor^{n'}$. Taking $Y = 2^{n'}\sqrt{D'}$ gives $\log h_{L'} \le (n')^2\log 2 + (n'/2)\log D'$. This is far from an optimal class-number estimate, but shows directly that the logarithm of the additional power of ℓ is polynomial in the parameters already controlled.

Finally, if $u \ge 2$, $Y \ge \log 2$, and a nonnegative quantity is bounded by $2\exp(u^C Y) + u^a Y$, it is bounded by $\exp(u^{C'}Y)$ for a sufficiently large absolute C'. The preceding class-number estimate gives exactly this linear dependence on Y after increasing powers of u. Indeed, enlarge C' so that $u^{C'}Y$ dominates $4u^{\max(C,a,1)}Y$; then $2e^A + A \le e^{4A}$ for $A \ge 1$ absorbs the sum. Applying this with $u = 2en$ and $Y = \log(2D)$ accounts for both torsion logarithms and the polynomial remainder. Thus the argument only requires the existence of one absolute exponent. Nothing in the window selection depends on its numerical optimization.

The twist in the final duality step can be checked directly. The Tate dual of $\mathbb{Z}_\ell(e+1)$ is $\operatorname{Hom}_{\mathbb{Z}_\ell}(\mathbb{Z}_\ell(e+1), \mathbb{Q}_\ell/\mathbb{Z}_\ell)(1) = \mathbb{Q}_\ell/\mathbb{Z}_\ell(-e)$. Thus the negative-twist localization kernel estimated by the class count is the one paired with the positive-twist second cohomology in Proposition 9.4. Changing the latter coefficient to $\mathbb{Z}_\ell(e)$ would shift the former twist and would give a different assertion. The shift by one comes from Tate duality, independently of the earlier convention for the Galois character attached to an automorphic weight.

The invariant terms in the duality sequence have a direct finiteness explanation as well. If a group acts on $\mathbb{Q}_\ell/\mathbb{Z}_\ell$ through a character η and some element g has $\eta(g) \ne 1$, every invariant is killed by $\eta(g) - 1$. Writing this number as a unit times ℓ^r shows that the invariant group is contained in the finite group killed by ℓ^r, of length r. For $\eta = \chi_\ell^{-e}$ over a finite extension of $\mathbb{Q}_\ell$, the cyclotomic character has open image, so such an element exists when $e \ge 1$. The quantitative bound additionally measures how far into the cyclotomic tower the field already extends. The source's local calculation bounds that initial depth by the local degree and then uses the binomial valuation formula for the eth power. Summing over the places above ℓ gives the degree-dependent error in the next display. This is why finiteness of the localization kernel alone is not presented as a bound for all of H^2: the local invariant terms must first be included and controlled.

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

$$\operatorname{length}_{\mathbb{Z}_\ell} H^2(G_{L,S_\ell}, \mathbb{Z}_\ell(e+1)) \le v_{\ell,e}(L) + n(2 + \lfloor \log_\ell n\rfloor + e). \tag{9.20}$$

Since $\mu_\ell \subset L$ for odd ℓ (and $\mu_4 \subset L$ when $\ell = 2$), one has $\ell \le n+1$. Multiplying (9.20) by $\log\ell$ and increasing the absolute exponent therefore proves (9.7). 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 [30]. 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 [13]. 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.

There is also a field comparison in the dyadic input. The base field L_0 contains μ_4; for each host layer E_h there is a layer L_1/L_0 of height $m \le h + 2$ for which $v_{2,e}(E_h) \le 1 + v_{2,e}(L_1)$ for every $e \ge 1$. Thus the host need not itself be the chosen layer over L_0. The additive one and the bounded shift of height are harmless in the final sum, but the comparison is needed to transfer a tower estimate to the original field. At a fixed odd prime, the relevant base field is selected once from the fixed character data. At the dyadic prime, the assertion instead has the order $\exists R\, \forall (q, d)\, \exists L_0$: the same rank bound works for every admissible prime pair, while the field and the controlling roots can change. This order is expressed by the fixed-prime and two-prime inputs in `SecondBound.lean`.

The closing construction makes the uniformity more explicit. Let g_ψ be the source's mod-2 character-rank bound for the cyclotomic tower over the fixed field $F_\psi(i)$. It obtains this bound from Ferrero–Washington and then takes

$$R_0(\psi) = 4(g_\psi + 6o).$$

The comparisons that first allow ramification at q, d and then adjoin the two quadratic factors give this bound for every admissible pair, using $q \equiv 3 \pmod 8$ and $d \equiv 7 \pmod{16}$. Thus Ferrero–Washington is applied to one fixed field, while the ramification and quadratic-extension comparisons supply the uniform passage to the changing fields. These are the assertions `baseRk`, `towerTwo`, and `twoIn` in `IsAlgClosureRat.lean`.

A finite-dimensional model explains why roots enter these estimates. Let M be a free $\mathbb{Z}_\ell$-module of rank r, and let T be an endomorphism. If $\det(T - a)$ is nonzero, Smith normal form gives $\operatorname{length}_{\mathbb{Z}_\ell} \operatorname{coker}(T - a) = v_\ell(\det(T - a))$. Over a finite extension splitting its characteristic polynomial, let $s_1, \dots, s_r$ be its roots, with multiplicities and with the valuation normalized by $v_\ell(\ell) = 1$. Then this length is $\sum_i v_\ell(a - s_i)$. No diagonalizability assumption is required, because the determinant identity alone gives the product. A bound $|a - s_i|_\ell \ge \ell^{-B}$ therefore bounds the length by rB. If a equals a root, the determinant vanishes and the cokernel can have a free part; the argument deliberately leaves that specialization to a separate estimate.

This model does not identify the actual tower module with an arbitrarily chosen matrix. In the arithmetic application, control maps, torsion submodules, logarithmic changes of the cyclotomic variable, and the comparison with finite layers must all be established before the root-distance estimate is available. The tower input supplies those steps in the form used below. The model explains its shape: the number of potentially close specializations is governed by a rank, and a second estimate is needed at the exceptional ones. It also explains why bounding the rank separately for each varying base field is insufficient. The number of excluded twists would then depend on the auxiliary prime pair and would enter the window exponent J, which is required to have been chosen earlier.

In the source the tower control is expressed by at most R numbers $s_1, \dots, s_{R'}$, $R' \le R$, 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 $|s_i - e|_\ell \ge \ell^{-B}$ 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. The second estimate uses the entire second cohomology group at the bounded layer: for layers L_1/L_0 and L_2/L_0 of heights m and j_0, respectively, it states

$$v_{\ell,e}(L_1) \le \operatorname{length}_{\mathbb{Z}_\ell} H^2(G_{L_2,S_\ell}, \mathbb{Z}_\ell(e + 1)) + R(m - j_0)_+. \tag{9.21}$$

Replacing the H^2 length by $v_{\ell,e}(L_2)$ requires retaining the degree-dependent local H^0 contribution from (9.20). The displayed form applies directly to Proposition 9.4, including at exceptional twists. 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 9.8. *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. □

The nearest-root argument is particularly efficient because it excludes integer twists rather than entire ℓ-adic balls. For each root, only one integer in the finite interval is removed. All other integers are separated from that chosen integer by a nonzero ordinary integer of bounded size, so the ultrametric inequality supplies the required uniform valuation bound. The root may have multiplicity, lie outside $\mathbb{Q}_\ell$, or be itself an integer; none of these cases changes the proof. Coincident choices of the nearest integer only decrease the size of the exceptional set. When $R = 0$ there are no roots and no exceptional twists, and the ordinary bound applies to the entire interval.

For a fixed X, apply Lemma 9.8 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 (9.7). 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 9.9 (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 9.9 gives $X^{1/J} \leq W$, $W^2 \leq X$, and

$$e \in \mathcal{E}, \quad e \leq x^{401}, \quad W \leq x \leq W^2 \quad \Longrightarrow \quad e \leq x^{\varepsilon/(2C_K)}. \tag{9.22}$$

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

The choice of the window is shared by all the arithmetic estimates. The union of the exceptional twist sets is formed before applying Lemma 9.9; selecting a separate interval for each prime would not ensure that their intervals intersect. The union has bounded cardinality because only the fixed odd-prime set and the one dyadic tower remain at this stage. Case I has already handled the unrestricted odd-prime sum by its norm product, and the other ordinary character cases are controlled without removing new twists. Thus the cardinality bound R_ψ is chosen using the fixed character and the uniform tower ranks, before the auxiliary pair is chosen.

Within a selected window, an exceptional twist has exactly the alternative needed by the proof. If it is larger than x^{401}, it does not occur in the cohomological sum at scale x. If it occurs, (9.22) bounds it by the much smaller power $x^{\varepsilon/(2C_K)}$. The gap selection does not make

its cohomology group vanish and does not bound the distance of that twist from its associated root. Instead, it makes the crude fixed-layer estimate cheap enough at every scale where the twist contributes. This is why the second tower estimate, which applies even at a root, is paired with the first estimate away from the roots.

The dependence on X is equally deliberate. Choosing nearest integers only up to the current x^{401} would permit their exceptional set to change as x moves, so one application of the gap lemma would not control the whole interval. Truncating once at $\lfloor X^{401} \rfloor$ fixes the exceptional set for every $x \le X$. Its elements may grow with X, but the gap lemma depends only on how many elements there are, not on their locations. The auxiliary pair is also fixed throughout the window. Hence all field, divisor, and character comparisons used at different values of x refer to that same pair, even though the dyadic height $h(x)$ changes.

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 $x \ge W \ge Y_0$, the conductor in (9.3) therefore satisfies

$$\log N_h \ll_\psi \log x + x^{\varepsilon/5}(1+\log x)^2, \qquad 2^h \asymp x^{2+8\varepsilon}. \tag{9.23}$$

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 (9.22) into (9.7), 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 (9.23) proves for $t \le x^{401}$, 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 9.10 (Arithmetic window; `vwin_of_row`)**.** *Assume the source's arithmetic inputs described above, with its automorphic Galois and local compatibility input described in Section 10. Fix a primitive character ψ, a real $0 < \varepsilon < 1/52$, and any function $h : [1,\infty) \to \mathbb{N}$ satisfying $x^{2+8\varepsilon} \le 2^{h(x)} \le 2x^{2+8\varepsilon}$. Then there are $J \ge 1$ and X_0 such that, for every $X \ge X_0$ and every auxiliary pair q, d admissible at scale $X^{1/J}$, there is a real W with $X^{1/J} \le W$, $W^2 \le X$, and*

$$\mathcal{V}\left(E_{h(x)}, \lfloor x^{401} \rfloor; \sqrt{\exp(x^\varepsilon)}\right) \qquad (W \le x \le W^2). \tag{9.24}$$

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 9.8 and 9.9, 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 9.2 supplies the odd-prime minus contribution, using $F \le N_h$. Formula (9.23) absorbs all of these into the right side of (9.24). Each input supplies finiteness as well as a bound for every finite prime sum, so the conclusion has exactly the meaning specified in (9.2).

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. □

Within the fixed-layer argument, the distinctive mechanism is the match between the filtered homology rank and the unit factor in the class-character count. The bar complex, the Bockstein

sequence, integral homological stability, and Chebotarev each have an established role. The construction-specific task is to connect them by integral cycles whose finite-field evaluations detect the specified weighted norms, with exactly the losses that permit the layer factor to cancel. The resulting bound is then strong enough for the window argument despite being very large. This identifies the substantive claim that requires inspection in the source and separates it from the elementary reductions proved here.

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.

10. 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 [8]. 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.

Three inputs are stated. Proposition 10.1 is the interface consumed by Theorem 9.10 and by the theta comparison. Proposition 10.5 is the descent step used in constructing it. Proposition 10.6 is the local assignment with respect to which its compatibility is expressed; it is deliberately weaker than a full local correspondence, and the discussion preceding it says where each of its laws is available. The lemmas of the section are elementary and are proved here.

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 10.1 (Galois compatibility interface; `Row0b`)**.** *The automorphic Galois input used in Theorem 9.10 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

$$H_\tau = \{(a-1)/2 - z : z \in \mathrm{Inf}_{\iota\tau}(\pi)\}, \qquad H_{\tau c} = \{w - h : h \in H_\tau\}, \tag{10.1}$$

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 9. A proof of the stronger unrestricted local–global correspondence is not needed to apply this interface.

The compatibility conditions have witnesses. The source constructs local prime and period data P, AI together with a `LocTied` identification with its concrete crystalline and de Rham period rings in `ExistsLocalPrimeAinf.lean`. For each $v \mid \ell$ and valuation ring A of $\overline{E}$ above v, it also constructs an embedding $j : \overline{E} \hookrightarrow \overline{\mathbb{Q}}_\ell$ with

$$x \in A \quad \Longleftrightarrow \quad |j(x)|_\ell \le 1,$$

and, after choosing a sufficiently large finite coefficient field, a compatible local bridge. Its local field is generated over $\mathbb{Q}_\ell$ by $j(E)$, and its topological Galois-group identification satisfies $e(g)(j(x)) = j(g(x))$. The valuation identity is precisely the condition `Induces`. This construction is `exists_locBridge` in `FiniteDimensionalAdjoinRange.lean`. Consequently the universal local conclusions above can be instantiated at every place above ℓ; existence of these realizations is a separate part of the source argument.

The local conclusions have two different levels of precision. In the abelian branch, equality of the Weil–Deligne parameters identifies both the Weil action and monodromy. In the tame branch, the stated conclusion identifies the inertia characteristic polynomials and separately proves that monodromy vanishes. Inertia polynomials by themselves would not give the latter assertion: an unramified Weil representation can support either zero monodromy or, when its Frobenius eigenvalues have the required ratio, nonzero monodromy. Thus the hypothesis $N_{y_v} = 0$ and the conclusion $N_{x_v} = 0$ perform actual work. The large-residue-field and multiplicity conditions are used to obtain this more restricted comparison; they should remain attached to it when the theorem is applied to an extension class.

The Hodge statement likewise contains multiplicities, not just a set of possible integers. For example, if $u \neq v$, two semisimple operators with eigenvalue multisets $\{u, u, v\}$ and $\{u, v, v\}$ have the same dimension and are both annihilated by $(T-u)(T-v)$, but their characteristic polynomials differ. An annihilating polynomial therefore cannot supply the missing multiplicity assertion. The geometric construction initially permits extra summands, and the subsequent extraction must account for their Frobenius eigenvalues and Hodge jumps with multiplicity. Formula (10.1) is the final normalized answer after this removal, including the half-integral shift that converts the chosen infinitesimal-character convention into integral Hodge–Tate weights.

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

$$2\lambda'_k(v) = \begin{cases} 2\lambda_k(v) - d - m_W(v) + n + m_X(v), & k \le n, \\ n + m_X(v), & n < k \le d. \end{cases} \tag{10.2}$$

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 $\langle \pi(f * f^*)v, v\rangle = \|\pi(f)^* v\|^2 \ge 0$. Choosing f so that $\pi(f)^* v \ne 0$ 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 [17, 26]; 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 10.2. *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. □

The equality of virtual classes in this lemma is accessible from characteristic polynomials, but several hypotheses are needed. We make that passage explicit. For a representation R put $P_R(g,T) = \det(T - R(g))$. A direct sum gives a product of these polynomials, and all such polynomials are monic and nonzero. If the theta comparisons supply

$$P_{R_1}(g,T)P_{A_2}(g,T) = P_{R_2}(g,T)P_{A_1}(g,T) \tag{10.3}$$

on a dense set of group elements, continuity of the matrix entries gives this identity everywhere. Indeed, each coefficient is a continuous function, and equal continuous functions into a Hausdorff field agree on the closure of the set on which they agree. In the Galois application the dense set consists of the conjugacy classes of unramified Frobenius elements supplied by Chebotarev [9]. The same algebraic normalization must be used for the automorphic polynomials and all the character factors before this density argument can be applied.

Lemma 10.3. *Over an algebraically closed field of characteristic zero, two finite-dimensional semisimple representations of a group with the same trace at every group element are isomorphic.*

Proof. Let V and W be the representations and let A be the linear span of the diagonal image of the group in $\mathrm{End}(V \oplus W)$. It is a finite-dimensional algebra: products of group elements are again group elements, so their span is closed under multiplication. Its invariant subspaces are exactly the invariant subspaces for the group. Consequently $V \oplus W$ is a faithful semisimple A-module, and the radical of A, which annihilates every simple module, is zero. The finite-dimensional semisimple algebra theorem therefore writes A as a product of matrix algebras over the coefficient field. Let S_i be their simple modules and e_i their central identity idempotents. If m_i, n_i are the multiplicities of S_i in V, W, then the traces of e_i are respectively $m_i \dim S_i$ and $n_i \dim S_i$. Equality of traces on group elements extends linearly to A, so these numbers are equal. Characteristic zero permits division by $\dim S_i$, giving $m_i = n_i$ for every i. Thus $V \simeq W$. □

Apply the lemma to $R_1 \oplus A_2$ and $R_2 \oplus A_1$, using the coefficient of next-to-highest degree in (10.3). It gives equality of their classes, and Lemma 10.2 produces V. In fact $R_i \simeq V \oplus A_i$ for both i. Thus V is an actual direct summand of either realization, has dimension $\dim R_i - \dim A_i$, and satisfies $P_{R_i} = P_V P_{A_i}$ at every group element. Polynomial cancellation recovers P_V even if one of its eigenvalues equals an eigenvalue of A_i. The disjointness needed in Lemma 10.2 is between the two lists of additional irreducible constituents; it is not a claim that the desired representation has no eigenvalue in common with either list at a particular place.

For character summands there is a constructive version of the same argument. Pick a character α occurring in A_1. The isomorphism $R_1 \oplus A_2 \simeq R_2 \oplus A_1$ forces an α-eigenline in its left side. Since α does not occur in A_2, that eigenline lies in R_1. Semisimplicity gives an invariant complement, so one can remove the line and cancel its polynomial factor. Repeat with the remaining characters, including their multiplicities. At the end, the complement is V. This successive removal is the implementation in `PPeelAll.lean`. The source proves the dense polynomial identity in `PDenseId.lean`. These are algebraic extraction steps once the theta and geometric comparisons have produced the two realizations.

The disjointness of the two extra character lists is itself an arithmetic statement. The source chooses the lifting characters so that the two lists of algebraic Hecke characters are disjoint. It then uses the injectivity of their Galois realization. If two such Galois characters were equal, their unramified Frobenius values would agree away from finitely many places; the corresponding Hecke characters would therefore agree by the character uniqueness theorem. This would contradict their chosen separation. This step is recorded in `PCharInj.lean`. Merely assigning different labels to the character lines would not give the necessary disjointness. Nor would an inequality between total dimensions detect a negative multiplicity in a virtual difference; the positivity proof has to be constituent by constituent.

The transport of local properties can be separated from this existence argument. If a global representation splits as $R = V \oplus A$, it still splits after restriction to a decomposition group. The summands need not be semisimple as local representations for this statement to hold: the global invariant projector restricts to a local invariant projector. Unramifiedness passes immediately to V, since an inertia element acting as the identity on R acts as the identity on its invariant subspace. If a polynomial annihilates an operator on R, it annihilates that operator on V. These assertions require an actual invariant subspace or projector; a formal subtraction of traces is not sufficient.

For de Rham and crystalline properties, the elementary dimension argument is also useful. Let D be an additive period functor for which $\dim D(U) \le \dim U$, with dimensions taken over the appropriate period-invariant and coefficient fields. If equality holds for $R = V \oplus A$, then additivity gives equality for each summand, because the two nonnegative deficits add to zero. Applied to D_{dR} or D_{cris}, this proves closure under direct summands. Potential semistability passes similarly after choosing a finite local extension over which R is semistable. The existence and comparison properties of these period functors are the established ℓ-adic Hodge input; the dimension argument is the subsequent linear-algebra step. It must be applied within one compatible local realization, so that its local field, period data, and comparison map are the ones quantified in Proposition 10.1.

Filtrations and Frobenius operators are compatible with these direct sums. Hence their Hodge jumps add as multisets, and their Frobenius characteristic polynomials multiply. If the geometric input gives the Hodge multiset $J \sqcup J_A$ and the character realization gives exactly J_A for A, cancellation yields J for V. In the same way, if the Weil–Deligne parameter of R is a direct sum of that of V and a known character parameter, its inertia polynomial can be divided by the known monic linear factor. Zero monodromy passes to the summand because the monodromy operator respects the direct sum. Compatibility of Weil–Deligne formation with the period comparison is part of the local input; once it is supplied, these operations are ordinary direct-sum identities.

The source carries out this local subtraction in `PClSub.lean`. Its separate assertions include de Rham descent to a summand, potential semistability, subtraction of Hodge jumps, and cancellation of Frobenius factors. The last cancellation divides by a polynomial, not by its value at a chosen eigenvalue. This distinction avoids a false nonvanishing requirement at an eigenvalue shared by two summands. The finite-field or tame multiplicity information used later must similarly be carried along the parameter comparison rather than inferred from the numerical rank alone.

There is a final character twist in the theta extraction. In the notation of the source's construction, the common rank-n part initially has Satake eigenvalues multiplied by the inverse of the value of a character κ. At an embedding its Hodge multiset is shifted by m, while the Galois character attached to κ has Hodge weight $-m$. Tensoring the extracted representation with that character multiplies every Frobenius eigenvalue by κ and shifts every Hodge weight by $-m$. The two changes cancel their respective preliminary normalizations. This is checked explicitly in `PAtPlace.lean`. Keeping the Satake multiplication and the Hodge addition in the same calculation prevents an unintended change of the cyclotomic-twist convention when the final representation is used in Section 9.

Finally, the auxiliary geometric representation can depend on the place being tested. This does not weaken the final conclusion once uniqueness has been proved. Fix one extraction V. At

any other chosen place v, construct the geometric realization with the local assertion needed at v and extract V_v. Both have the same unramified Frobenius polynomials away from a finite set, by the common automorphic data. Chebotarev and Lemma 10.3 give $V_v \simeq V$. Transporting the local assertion along this isomorphism establishes it for V. The representation may therefore be chosen once, before v is chosen. What must be available for every v is a suitable auxiliary realization and its comparison with the common Frobenius data; one does not need a single geometric realization satisfying every local auxiliary requirement at once.

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 10.4. *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, $\operatorname{Ind}_H^G s$ 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. □

Two elementary features of this argument explain the semisimplicity requirements in patching. First, restriction of a finite-dimensional semisimple representation to a normal subgroup of finite index is semisimple. It suffices to start with an irreducible representation W of the larger group. Choose a simple submodule U for the subgroup, which exists by taking a nonzero invariant subspace of minimal dimension. Each translate of U is again simple, and the sum of its finitely many translates by coset representatives is a nonzero invariant subspace for the larger group. It is therefore all of W. A sum of simple submodules is semisimple, proving the assertion. Second, a representation with trivial restriction to the subgroup can still be nontrivial on the quotient. For index two the nontrivial quotient character is exactly the ambiguity between the two extensions of a fixed constituent. Patching must remove this ambiguity by comparisons with additional quadratic subgroups.

In the fixed-constituent case of Lemma 10.4, these are the only possibilities even when a constituent occurs several times. On its isotypic component every extension has the form of the chosen intertwiner tensored with an operator J on the multiplicity space. The relation for the square of the nontrivial coset gives $J^2 = 1$. Characteristic zero makes J diagonalizable with eigenvalues 1 and -1, so its conjugacy class is specified by the two multiplicities m_+, m_-. In the two-element orbit case, the nontrivial coset interchanges the two isotypic components, and a change of bases makes the extension the stated induced representation with its prescribed multiplicity. This proves the finiteness of the extension list without an assumption that the original group or its image is finite.

Proposition 10.5 (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 10.4 gives an extension of every r_i. Fix $i = 0$ and let $R_1, \dots, R_s$ 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 10.4. 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 10.5 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 [29].

The representation supplied by patching is unique up to isomorphism for the given compatible family. Suppose ρ' has the same eventual restrictions. Apply the Zariski-image observation in the proof of Proposition 10.5 to $\rho \oplus \rho'$. Outside a finite set of quadratic characters, the kernel has the same Zariski image as the whole group. Choose one also outside the finite exceptional sets for the two restriction comparisons. An intertwiner between their restrictions commutes with that kernel image, and the equations defining an intertwiner are polynomial equations in the image matrices. They therefore hold on its Zariski closure and on the image of the whole group.

Thus $\rho \simeq \rho'$. The argument uses a single member of an infinite family after all finite exclusions have been made, so it does not require choosing compatible matrix bases in the different r_i.

This makes the distinction between the global and local choices precise. The global representation is produced once from the family. Now fix a place v and its decomposition group. Choose a quadratic extension E_i/E in the allowed family in which v splits and whose index is outside the finite exceptional set. For a chosen place w above v, the completion $E_{i,w}$ is E_v. After choosing compatible algebraic closures, its absolute Galois group identifies with the original decomposition group, and r_i there is isomorphic to the local restriction of ρ. Every local assertion invariant under isomorphism transfers along this identification, including the specified inertia polynomials, crystalline condition, Hodge jumps, and the comparison made in the chosen period realization. The extension may depend on v, while the global ρ does not.

The splitting condition cannot be replaced by merely passing to an arbitrary quadratic extension. A nontrivial ramified quadratic character becomes trivial on the absolute Galois group of its quadratic field, so unramifiedness of a restricted representation need not imply unramifiedness downstairs. More generally, a local extension can remove ramification or alter the comparison data. At a split place no local extension occurs, which is why the desired property returns to the original local field. The separate arithmetic input providing infinitely many suitable splitting extensions is therefore logically distinct from the representation-theoretic patching proposition. The source's base-change and local-comparison arguments are responsible for verifying it together with the pairwise trace agreement.

A similar distinction applies to the density argument on the intersections $H_i \cap H_j$. The equality needed there concerns all group elements of the intersection. Initially automorphic base change supplies Frobenius polynomials at almost all unramified places of the common overfield. Continuity and Chebotarev extend the equality of traces to its Galois group, after the restriction maps have been identified. Semisimplicity then gives an isomorphism. Thus the two uses of density have complementary roles: Frobenius density converts automorphic information to group-wide equality, and Zariski density makes restriction to a generic quadratic subgroup retain the intertwiners of a finite list of candidate representations. Neither asserts that one particular quadratic subgroup is topologically dense in the original profinite group.

We finish by specifying the local parameter assignment used in Proposition 10.1. 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 [12, 22]. Langlands segments then prescribe the parameter of a general irreducible admissible representation as the corresponding sum of segment parameters.

The elementary structure of tame inertia helps explain the class singled out by the interface. A Weil representation is continuous with finite image on inertia, so in characteristic zero its tame inertia action is diagonalizable. If γ is a tame generator and Φ a Frobenius element, the relation $\Phi\gamma\Phi^{-1} = \gamma^{q_v}$ permutes its eigencharacters by the q_v-power map, with multiplicities preserved. Under the tame multiplicity condition, an eigenvalue of multiplicity at least two is fixed by this map. Equivalently, any orbit of eigencharacters of length greater than one occurs with multiplicity one. Repeated eigenvalues are allowed on the fixed orbits, including the trivial inertia character. This is the useful combinatorial distinction; it would be lost by replacing the condition with a blanket assertion that every inertia eigenvalue is simple.

For a depth-zero supercuspidal attached to a regular character of the multiplicative group of a residue-field extension, regularity means that its Frobenius conjugates are distinct. The

associated induced Weil parameter restricts to exactly that orbit of tame characters. This explains why the finite-field description is suited to the controlled tame branch. General good representations can also have character twists and segment factors, so this observation alone is not a proof of the local compatibility proposition. The type comparison, the control of monodromy, and the bound $p > 16a^2$ in the prime-field case remain separate ingredients of that proposition. On abelian parameters, the character reciprocity description gives the different, exact comparison used by its other branch.

There is also a distinction between compatibility of an assignment and uniqueness of the full local correspondence. The arithmetic argument tests the assignment on specified character, spherical, and controlled tame situations. An assignment on every local isomorphism class is a convenient way of stating those tests uniformly, provided its stronger laws retain their good-class qualification. Choosing values on the unused wild classes cannot establish an exact determinant identity or arbitrary twist compatibility on those classes. Conversely, proving the restricted laws suffices wherever the hypotheses place the local component in the specified range. The interface therefore states both the available laws and the range in which each one is available.

Proposition 10.6 (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 9.10.

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 5 has the following quantifier form: for each fixed $J \geq 1$, 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 9, 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 $c_\psi \ge 1$ 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 $c \ge 1$ there is a $J \ge 1$ 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 $[T, x]$, $x^{1/J} \le T \le x$,*

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

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 $\kappa = c^{1/(2+8\varepsilon)} \ge 1$. Given a desired lower bound for Z, choose M above that bound, T_0, 1, κ^3, and κ^{400}. Choose $X \ge \max(X_0, M^J)$ and put $T = X^{1/J}$. Then $T \ge M$, and analytic pair selection supplies q, d on $[T, X]$.

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

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

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

The cohomological estimate (11.2) is available at x'. Since $x' \ge W \ge \kappa^{400}$, we have $x^{400} = \kappa^{400}(x')^{400} \le (x')^{401}$, and therefore $\lfloor x^{400} \rfloor \le \lfloor (x')^{401} \rfloor$. Also $x' \le x$ gives $\exp((x')^{\varepsilon}/2) \le \exp(x^{\varepsilon}/2)$. Monotonicity of the finite-prime-sum bound in its twist cutoff and its upper bound proves (11.4). The pair was never changed, and $Z \ge W \ge M$ 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, and κ is fixed by the original character before the final threshold is chosen.

The source declaration `winPairs` is stated in a slightly weaker form: for each x in the window it asserts the existence of a scale, a prime pair, and a dyadic exponent with these properties,

without recording that the pair can be chosen independently of x. The proof above, which follows the source's, shows that one pair serves the whole window. Only the weaker form is used in the final deduction, as explained after Proposition 11.2.

The dyadic exponent is allowed to jump when x changes. This does not disrupt the interval estimate. At each real x the selector chooses an integer satisfying (11.3); the cohomological theorem quantifies over every such input field in the specified class, and the analytic conclusion concerns the same function $H_\psi(x)$ regardless of the selected field. No continuity or differentiability of the family of host fields is needed. The regularity used by the Laplace-transform argument belongs to H_ψ, whose definition is fixed at the beginning of the paper.

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 $P_1, \dots, P_s$ 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.

Two points in this principle carry its content. The order is fixed: the exponent first, then each threshold for that exponent, then their finite maximum, and the existence theorem last; choosing a family separately for each property and then identifying the resulting objects would not be legitimate. And each property must hold for *every* admissible family beyond its threshold: the existence of arbitrarily large candidates satisfying each property separately would not give one candidate satisfying all of them. The source's closure machinery records exactly this form, with thresholds that are functions of the fixed constants and are combined by pointwise maxima, and the existence theorem is invoked within the intersection of the required geometric and arithmetic classes.

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 (4.15)*, 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 4, 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 7 and 8 explain the local, geometric, and frame inputs; Sections 9 and 10 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.

Before giving the final deduction, we record exactly how a family estimate becomes an analytic window estimate. This small observation isolates the last uniformity requirement from all the preceding constructions.

Lemma 11.3 (From a fixed family to distant windows)**.** *Let $\phi : [1,\infty) \to \mathbb{R}$, let $\beta \in \mathbb{R}$, and fix $r > 1$. Suppose that for every sufficiently small $\theta > 0$ there exist a set $W_\theta \subseteq [1,\infty)$, a finite constant $K_\theta \geq 0$, and a threshold X_θ such that*

$$\phi(x) \leq K_\theta x^{\beta+\theta} \quad (x \in W_\theta,\ x \geq X_\theta),$$

and, for every $A \geq 1$, the set W_θ contains some interval $[Y, Y^r]$ with $Y \geq A$. Then for every $\eta > 0$ and every $A \geq 1$ there is a $Y \geq A$ such that $\phi(x) \leq x^{\beta+\eta}$ throughout $[Y, Y^r]$.

Proof. Choose an allowed θ with $0 < \theta < \eta$, and fix the resulting set, constant, and threshold. Enlarge the lower bound to $B = \max\{A, 1, X_\theta, \max(1, K_\theta)^{1/(\eta-\theta)}\}$. Choose $[Y, Y^r] \subseteq W_\theta$ with $Y \geq B$. For every x in this interval, $K_\theta \leq x^{\eta-\theta}$ and the asserted estimate follows by multiplication with the positive number $x^{\beta+\theta}$. □

This lemma permits constants that deteriorate without bound as $\theta \to 0$. It does not require one family or one threshold to work for all exponents simultaneously. To rule out a zero with a fixed positive distance from the critical line, the analytic argument chooses one sufficiently small exponent and then works with that exponent throughout. Likewise, the lemma imposes no regularity on the auxiliary data as functions of x. It only asks that the constants in the resulting estimate be independent of the particular scale in the chosen family.

The order of choices matters. A collection of families, each with a threshold lying beyond its own windows, would supply arbitrarily distant intervals and no estimate on any of them. Proposition 11.2 supplies the whole-window property for one family after its threshold has been fixed, so the application of Lemma 11.3 is not vacuous.

There is a second distinction between the two window statements used in this section. Proposition 11.1 keeps the auxiliary prime pair fixed on each input interval. Proposition 11.2 then constructs a family with scale-independent comparison constants and its own full-window property. The final analytic deduction uses the latter. It does not need to strengthen the family theorem by claiming that its prime pair is constant on every final window. The fixed-pair statement has already served its purpose in constructing compatible input data; the function being estimated is independent of that pair.

Proof of Theorems 1.1–1.3. Fix a primitive character ψ. The Galois construction in Section 10 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 9. The auxiliary-prime theorem of Section 5, 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 6 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 4 therefore supplies x_ε such that

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

Given $A > 0$, choose $Y \geq \max(A, x_\varepsilon, 1)$ 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 (2.14) 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 [23] at 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 5	`pairEx.lean`; the exact pair predicate is in `PairGood.lean`.
Fourier balance and extraction; Section 6	`ChBalance.lean`, `ChCore.lean`, and `ChTop.lean`.
Frame codimension estimate; Proposition 3.3	`FrameLinearAlgebraFirst.lean`.
Sharp mean-slope transfer; Theorem 3.5	`StatementFrameF.lean`, `TheoremFSharp.lean`, and the public wrapper in `ShiftInvariance.lean`.
Theta spaces and integral data; Section 7	`DataSpaces.lean`, `DataKernels.lean`, and `DataArith.lean`; the section gives the more specialized local input links.
Geometric covolume and spectral inputs; Section 8	`TopCovolSlim2Use.lean`; the interpolation and frame interfaces are linked at their statements.
Uniform fixed-layer cohomology; Proposition 9.4	`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 9	`Statement.lean`, `JointWinOfCounts.lean`, and `VwinOfRow.lean`.
Galois row and local parameter; Section 10	`RowSrc.lean`, `RecTExists.lean`.
Exact normalization cancellation; Proposition 4.1	`WeightedMeanData.lean`. The pointwise input is kept distinct from the average formulation.

Mathematical passage	Source entry points
Final exponent calculation; Proposition 4.3	`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`.

A short route through the final source declarations helps explain how to use this table. Start with `Statements.lean` to fix the meaning of the public assertion, then pass through `MainTheorems.lean`. The right-half-plane assertion there is supplied by `MainConclusion.lean`. That file first proves the conclusion from a precise Galois row, using the row both in the arithmetic window theorem and in the geometric comparison. It then substitutes the constructed row. The local parameter existence theorem is supplied as a separate input. This short chain shows which result closes the argument, but its brevity should not be confused with the size of the mathematics imported into it.

For an individual construction, the useful question is what an input contains before asking where it is proved. For example, the frame input contains two trace estimates and a retained-target rank estimate; it is stronger than nonvanishing of one evaluation map. The integral comparison input contains a lattice map and a height bound with specified metrics; it is stronger than an isomorphism of complex representations. The arithmetic window input quantifies over varying dyadic layers and specified host fields; it is stronger than finiteness for a fixed field. The statements in the corresponding sections record these differences explicitly, so that following a source link does not require guessing which strengthening is needed.

Some propositions unpack a principal source theorem. Others, such as Corollary 4.2, combine previously stated inputs; their combining arguments are proved here, without implying that Lean contains a declaration with the same prose title. Explanatory examples likewise illustrate a mechanism without claiming to realize the entire arithmetic construction.

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 7–10 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 on Andrew Sutherland's MIT verification page [31]. The following describes that page as consulted on 23 November 2026, when its own stamp read "last updated 2026-11-22 16:20 UTC". It reports compilation of all 175,187 modules; only the three standard axioms `propext`, `Classical.choice`, and `Quot.sound` for the five public theorems;

replay of every module by `leanchecker` and by `lean4lean`, five modules requiring a raised step limit in the latter; acceptance of the full export by `nanoda` and `con-leche`; and statement comparisons by `con-leche` and `con-ron` confirming the five theorems against the challenge statements. A separate sandboxed `lake build` reproduces the export byte for byte. Three caveats are recorded there. One compiled file, `LocModel.olean`, is produced non-deterministically, so that 175,186 of the 175,187 compiled files agree with the author's manifest and the files have to be compared individually; the page attributes this to a naming race that does not affect the export or kernel checking. The author's original build invoked `lean` directly, not through `lake`. And at that update the kernel stages of the `lake comparator` run and the Lean-kernel replays of the export were still running, with no verdict. These are reported external checks; a new full Lean build was not performed for this revision. The correspondence between this exposition and the source 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] L. Alpöge, *Remark*, announcement note, 20 November 2026. https://alpo.ge/remark.pdf.

[3] 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.

[4] 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.

[5] 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/.

[6] 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.

[7] 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/.

[8] 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.

[9] N. Tschebotareff, *Die Bestimmung der Dichtigkeit einer Menge von Primzahlen, welche zu einer gegebenen Substitutionsklasse gehören*, Math. Ann. **95** (1926), 191–228. https://doi.org/10.1007/BF01206606.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] D. R. Grayson, *Reduction theory using semistability*, Comment. Math. Helv. **59** (1984), 600–634. https://doi.org/10.1007/BF02566369.

[16] 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.

[17] 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.

[18] 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.

[19] 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/.

[20] V. A. Kolyvagin, *Euler systems*, in The Grothendieck Festschrift, Vol. II, Progr. Math. **87**, Birkhäuser, Boston, MA, 1990, pp. 435–483.

[21] E. Landau, *Über einen Satz von Tschebyschef*, Math. Ann. **61** (1905), no. 4, 527–550.

[22] 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.

[23] The mathlib Community, *The Lean mathematical library*, in Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020), ACM, New York, 2020, pp. 367–381. https://doi.org/10.1145/3372885.3373824.

[24] 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.

[25] 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.

[26] 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/.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] A. V. Sutherland, *Verification files for the Lean proof of the Riemann Hypothesis*, Massachusetts Institute of Technology, consulted 23 November 2026 (page stamp 2026-11-22 16:20 UTC). https://math.mit.edu/~drew/grh/.

[32] 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.