Specialist research manuscript

Heights of Iterated Theta Lifts and the Dirichlet Generalized Riemann Hypothesis

A reconstruction of the shortened Alpöge–Claude argument

Second exposition — working draft of 26 November 2026

Abstract. We reconstruct the mathematical structure and many of the constituent proofs of the compact Lean development asserting RH and GRH for Dirichlet L-functions. The analytic criterion uses smoothed prime sums, a quadratic-twist zero-density estimate and a high-moment window argument. The arithmetic construction compares upper and lower bounds for the heights of iterated theta lifts between carefully chosen unitary groups. We give explicit field, weight and level constructions; develop local CM theta-jet divisibility and Hecke repair, quantitative Bessel–Poincaré spanning, determinant exchange and one-lift comparisons; and explain the fit-family theorem via Galois congruences, Selmer modules and integral crystalline losses. The final cancellation is derived explicitly. Specialized arithmetic estimates whose written proofs have not yet been reconstructed are stated as formal-source inputs, with references to the exact declarations; thus the manuscript is not yet an independently complete conventional proof.

Attribution and scope. The underlying proof is attributed by the supplied repository to Levent Alpöge with Claude. Both the readable and compact Lean editions were used, with the latter primary for numerical statements. This manuscript is an expository working reconstruction; it does not claim new authorship of the proof, independent kernel replay, or peer verification. The shortened Lean code on which this exposition is based is credited to Luke Sernau; see his repository, https://github.com/lsernau/grh-lean.

AI disclaimer. The text of this document was written by AI, ChatGPT-6.2 Sol, and has not been checked by a human.

Contents

1 Introduction and mathematical status **4**
1.1 The architecture of the proof 4
1.2 Primary material, conventions, and the remaining challenge 5

2 The analytic problem: a one-sided bound for weighted prime sums **6**
2.1 The particular smoothing used in the proof 6
2.2 Why the prime sum contains information about zeros 8
2.3 An elementary form of Landau's one-sided principle 9
2.4 Quadratic twists avoiding a prescribed half-strip 10
2.5 The one-sided zero-free criterion 13
2.6 A good second scale in a long window 14
2.7 From the target bound to GRH 15

3 The admissible data and coefficient spaces **16**
3.1 The CM extension and the different 16
3.2 The scales, and the reason for so many parameters 17
3.3 Admissibility and construction of the data 18
3.4 Derived weights, level and integral structures 20
3.5 Two kinds of automorphic forms 20
3.6 Precise weights and levels of the source and target 21
3.7 Integral gauges and the precise uniform index 23
3.8 The theta chains and their relation 24
3.9 Norms, squared covolume and index 24
3.10 The common leading term and the prime remainder 25

4 Local CM theta analysis and Schur jets **26**
4.1 The arithmetic local scales 26
4.2 The CM period and the differential coordinate 27
4.3 Integral and logarithmic theta identities 27
4.4 The regularized CM parameter at nonsplit primes 28
4.5 The exceptional exponential factor 29
4.6 Tight frames and theta vectors 31
4.7 Frobenius pairings in tight-frame products 32
4.8 Schur matrices and determinant losses 34
4.9 Application to the prescribed weights 36
4.10 From theta products to the standard integral kernel data 38

5 The definite chain, collisions and Hecke repair **40**
5.1 Coefficient lattices, class spaces and index 40
5.2 The groups and the chain of theta kernels 42
5.3 The geometric seesaw and its denominator control 44
5.4 Ramification and short generators on the definite group 47

5.5 The collision loss and the marked set 48
5.6 Moving the marked Hecke operators through the chain 50
5.7 The Frobenius-twisted elementary-divisor theorem 53
5.8 Application to the actual column lattice 59
5.9 The prime N: moving denominators to a Fourier operator 61
5.10 The simultaneous local divisibility statement 63
5.11 Archimedean growth and the level normalization 64
5.12 Projection, product formula scaling and Theorem A 65

6 Quantitative spanning by projected Bessel–Poincaré series 67
6.1 Statement and finite-dimensional framework 67
6.2 Finite level, central projection, and the transfer bound 67
6.3 Hermitian geometry of the Bessel subgroup 69
6.4 The archimedean Bessel system 71
6.5 Construction and exact unfolding of the Poincaré map 74
6.6 Energy, pairs of congruent terms, and removal of lower cells 76
6.7 The sharp coherent main term 80
6.8 A finite-field positivity argument for singular minors 84
6.9 First moment and bad finite-type loss 85
6.10 From the moment bounds to the covering theorem 86

7 The determinant lower bound 88
7.1 Conventions for norms of conjugate coefficient vectors 89
7.2 Heights with varying finite-place entries 90
7.3 The determinant lemma with index 92
7.4 Special-value data and fit twists 93
7.5 The correspondence of complementary rank and the finite square 95
7.6 The two-grid exchange and the crude rows 97
7.7 The triangle of weighted lifts 100
7.8 Comparison forms for one lift 104
7.9 Compressed Gram determinants 106
7.10 The height saving in a rank-one special value 108
7.11 Combining the special values and closing the lower bound 112

8 Congruences, auxiliary primes, and the fit family 113
8.1 The integral projection problem 113
8.2 Labels and Galois representations of lifted eigensystems 114
8.3 Selmer modules over finite coefficient rings 116
8.4 New level characters and tied cocycle conditions 118
8.5 Shrinking several tied Selmer modules 120
8.6 Finite presentations and interpolation of lengths 121
8.7 Degree-free selection of roots of unity 123
8.8 The global twisting theorem and value families 124

8.9 The Tate characters and triangular Taylor–Wiles places 127
8.10 The first-order Ribet construction 129
8.11 Hecke operators realizing the first-order relations 131
8.12 The denominator bound from the local length budgets 134
8.13 Global numerical summation and proof of the fit theorem 135

9 Local crystalline defects and the integral cup-product estimate 136
9.1 Lattices, crystalline cocycles and imperfectly dual reductions 136
9.2 Why this defect bounds the local discrepancy of congruent cocycles 139
9.3 The Kummer chart, letter decomposition and height of a frame 140
9.4 Exponent lattices and the cost of increasing depth 142
9.5 Cokernel estimates: divisibility by p and by the Eisenstein element 144
9.6 Cup-product vanishing from the three sides 145
9.7 The window estimates: the three ways to obtain safe points 147
9.8 The tame type bound, uniform in the ramification degree 151
9.9 Case analysis for the primes and the explicit local loss 152
9.10 Insertion of the global weights and the sum over all blocks 153

10 The arithmetic inequality and conclusion 155
10.1 Collecting the arithmetic inputs 155
10.2 From A and B to an almost-basis of short forms 156
10.3 Comparison with the lower bound 156
10.4 Extracting the analytic remainder 157
10.5 Choosing the second cutoff and the power saving 158
10.6 GRH 158

11 Dependency audit and remaining proof obligations 159
11.1 Status of the reconstruction 159
11.2 Reference matching is a mathematical task 160

1 Introduction and mathematical status

The compact edition of the Alpöge–Claude formal development [1] is a remarkably large argument for the following assertion.

Theorem 1.1 (Dirichlet GRH; asserted in the formal source)**.** *For every integer $q \geq 1$, every Dirichlet character χ modulo q, and every $s \in \mathbb{C}$ with $\Re(s) > 1/2$,*

$$L(s, \chi) \neq 0,$$

where the pole of a principal character at 1 is excluded. In particular every zero in $0 < \Re(s) < 1$ lies on the line $\Re(s) = 1/2$, and the nontrivial zeros of ζ lie on this line.

The qualifier Dirichlet is important: the paper concerns the L-functions of degree one over $\mathbb{Q}$, not the generalized hypothesis for all automorphic L-functions. The proof uses higher rank unitary groups over auxiliary CM fields, but it uses them to construct and compare algebraic objects, not to assert GRH for their standard L-functions.

This manuscript is the second stage of a proposed series of expositions. Unlike the preceding overview, it is organized as a specialist research article: definitions are fixed once, estimates are given in their operative form, and the derivations of the constituent results are expanded. It is a *working reconstruction of the formal proof*, not a claim that every arithmetic argument has yet been independently translated into a conventional proof. Every assertion called a "formal-source input" is a mathematical statement extracted from, or a stated interface to, the supplied Lean development that still needs an independent written proof or a verified literature replacement. A statement followed by a proof records a mathematical derivation of that statement from the explicitly stated inputs; it does not thereby establish those inputs. In particular the manuscript must not be cited as an independently complete proof of Theorem 1.1 until the remaining formal-source inputs have been discharged. The main outstanding packages are listed at the end. This convention allows a specialist to test an argument at the level of its precise hypotheses, instead of having to infer whether an omitted detail is formal bookkeeping or a substantial arithmetic claim.

1.1 The architecture of the proof

The principal idea is to use the same automorphic object for two opposed quantities. At the archimedean places it carries a metric; at the finite places it carries an integral coefficient structure. These are linked by a product formula. To encode primes, consider jets of a theta kernel for a CM elliptic curve in a specified algebraic normalization. Their divisibility at a nonsplit rational prime depends on how the local formal-group parameter transforms under Frobenius. The valuation bounds contain sums of positive parts indexed by powers of the prime. A choice of the infinite-place weights turns their global sum into the weighted prime remainder.

One cannot simply apply the product formula to an arbitrarily chosen jet. The proof instead uses a family of forms obtained through an iterated theta correspondence. Suitable source forms are transferred by a seesaw identity to the lifts of Poincaré series on an indefinite unitary group. The arithmetic choice of fields, levels and Harish-Chandra parameters serves three tasks at once: the finite-place integrality estimate, a nearly sharp second moment for Poincaré series, and a matching lower bound on the determinant of the lift. The latter requires a family of twists, one at each finite embedding, rather than a single twist valid at every place.

Here are the three central estimates, in the notation that will be fixed in Section 3. Let W be a row space of dimension n, C the source space of dimension c, and $C^{\mathrm{ram}} \subseteq C$ its

ramified part. Let Θ_θ be a composite of m lifts, and $P_s : C^{\mathrm{ram}} \to W$ the projected translates of a Poincaré map, $s \in \mathcal{G}$. The large positive numbers $B = fa(\log b)^3$ and $L = B \log b$ are error budgets, while $H = mk\mathcal{U}$ denotes a signed common main term. The weighted prime remainder is $\mathcal{R}$.

A. There exists a positive rational scaling of the P_s and a spanning set of integral elements of C^{ram} for which all trivial-twist lifts have squared norm at most $\exp(H - m\mathcal{R} + L)$ at every archimedean conjugate. A family from the images of J translates has logarithmic index at most $|J|cB$ simultaneously in the admissible integral structures.

B. Both c and n are positive. There are translates indexed by a finite set J with $|J|c \leq 5n \log(1 + b^2)$ whose images span at least $(1 - 4/b^2)n$ dimensions.

C. Every independent family of at least $(1-4/b^2)n$ rows whose trivial-twist lifts have squared covolume at most e^V at all conjugates and whose logarithmic index is at most d_{ind} satisfies $n(H - 3L) \leq V + 2d_{\mathrm{ind}}$.

Assertions A, B, C are the interfaces `Data.short_forms`, `Data.enough_forms`, and `Data.integral_forms_long` of the compact source. Their proof decompositions occupy the bulk of the paper. Notice that C is uniform in the choice of row family, unlike A, and that B is quantitative; neither of these features can be dropped in the endgame.

To see the algebraic mechanism, choose $q \geq (1 - 4/b^2)n$ independent short rows using A and B. Hadamard's inequality gives $V \leq q(H - m\mathcal{R} + L)$, while the index is at most $5n \log(1 + b^2)B$. Combining with C, and bounding the missing-dimension term by a crude estimate, gives

$$m\mathcal{R} \leq 4L + 4x^{99} + 10 \log(1 + b^2)B \leq fa(\log b)^5.$$

The two instances of H cancel exactly. The number m of lifts is of size $\sqrt{x}$; this is the source of the square-root factor in the final saving. The normalized prime remainder $R_{N_0}(x)$ is related to $\mathcal{R}$ by the exact identity

$$\frac{\mathcal{R}}{a} = |\mathcal{S}|(R_{N_0}(x) - R_{N_0}(x_1)) + |\mathcal{T}|R_1(x_1),$$

where $\mathcal{T}, \mathcal{S}$ are two subsets of infinite places. The compact analytic argument chooses x_1 in a long logarithmic window to give a one-sided upper bound for $R_{N_0}(x_1) - |\mathcal{T}|R_1(x_1)/|\mathcal{S}|$. With $|\mathcal{S}| \geq |\mathcal{T}|/\log b$, $f = \varphi(N_0)|\mathcal{T}|$, and $\log b = x^\varepsilon$, one concludes $R_{N_0}(x) \leq x^{-1/2+8\varepsilon}$ eventually. A Landau argument then excludes zeros to the right of $\frac{1}{2} + 8\varepsilon$. Quadratic twists are varied to prevent cancellation of the singularities of logarithmic derivatives.

1.2 Primary material, conventions, and the remaining challenge

The supplied snapshot `grh-lean-main` contains the two variants `grh/` (readable) and `grh-compact/` (compact), of approximately 300,000 and 76,000 Lean lines respectively. Their high-level interfaces agree but some proofs differ. We take the compact variant as primary and use the readable one when clarifying its definitions or constituent steps. Unless qualified otherwise, paths in the text are relative to `grh-compact/src/`. The root README, attribution and guide identify the underlying proof as due to Levent Alpöge with Claude and the shortening as a collaboration using Claude. The formal source reports kernel checks of all exported results and dependence only on `propext`, `Classical.choice`, and `Quot.sound`; this manuscript makes no claim of a new independent machine replay. The statement of the result imports only Mathlib's definitions, not the auxiliary arithmetic theory of this argument. The exact goal is in `Statements.lean`, and the assembly is `New.grh` in `New/Main.lean`.

The new mathematical difficulty is not measured by Lean line count. In particular the division into `Lib` and `New` is not a proof that all of the former can be replaced by standard citations. Uniformity in growing ranks, large weights and varying auxiliary fields, integral versions of the correspondence, normalization of class sums and precise determinant comparisons must all survive such replacement. We therefore take care to mark specialized inputs instead of silently replacing them by named theorems from the literature. Classical results used in the analytic and parameter-construction sections are identified explicitly and do not presuppose GRH.

Notation. The letter d always denotes the chosen auxiliary rational prime; d_{ind} denotes an index bound. The size b is a scale, not a weight coordinate. We use Φ for the set of embeddings of E above the chosen embedding of K, $\mathcal{T}$ for the tops, and $\mathcal{S}$ for the thick places. We write $A(y,\sigma)$ for the weighted sum in `Data.primeSum`, $A_p(y,\sigma)$ for its normalized p-part, and $J(j)$ for `Data.jetSize`. In some local arguments letters may denote temporary variables, but the main budgets $B, L, H, \mathcal{R}$ retain their meanings. The notation W_j for the auxiliary Hermitian spaces is different from W, the vector space of row forms. The group $\mathcal{G}$ is also denoted $\mathfrak{G}$ in the finite moment estimates; $C_{\mathrm{ram}} = C^{\mathrm{ram}}$.

2 The analytic problem: a one-sided bound for weighted prime sums

Before constructing the automorphic objects, we isolate the precise problem that they solve. The analytic reduction is useful in its own right: it separates the difficult arithmetic estimate from the comparatively short argument about zeros. In particular, the arithmetic construction need not produce a two-sided error estimate for the ordinary prime counting function.

Throughout this section $M \geq 1$ is a modulus, d is a prime with $d \equiv 7 \pmod 8$ and $(d, M) = 1$, and

$$K = \mathbb{Q}(\sqrt{-d}), \qquad \chi_K(\nu) = \Big(\frac{\nu}{d}\Big).$$

The discriminant of K is $-d$; quadratic reciprocity identifies the displayed character with its quadratic Dirichlet character. Write $\varphi(M)$ for Euler's totient. The dependence on d in the notation R_M below is suppressed, but it is essential that the auxiliary prime d can eventually be varied.

2.1 The particular smoothing used in the proof

Let $\Lambda(p^j) = \log p$ for prime powers, and let $\Lambda(\nu) = 0$ otherwise. Put

$$\lambda_{d,M}(\nu) = \mathbf{1}_{\nu \equiv 1 \bmod M} \Lambda(\nu)(1 - \chi_K(\nu)). \tag{1}$$

Thus $\lambda_{d,M}$ is non-negative. A prime split in K contributes zero, an inert prime contributes twice its ordinary von Mangoldt weight, and the ramified prime has factor one. The same rule applies to prime powers; for instance an even power of an inert prime has factor zero. Congruence $\nu \equiv 1 \pmod M$ is imposed on the integer ν, not on its underlying prime. In particular, the weighted sum is not literally a sum over primes alone.

For any Dirichlet character ξ, with its modulus understood, denote by

$$\widetilde{L}(s,\xi) = \begin{cases} L(s,\xi), & \xi \text{ non-principal}, \\ (s-1)L(s,\xi), & \xi \text{ principal} \end{cases}$$

the entire function obtained by removing the possible pole at 1 (the second line is interpreted by continuation). The character $\xi\chi_K$ is taken modulo Md; it is non-principal because d is a new odd prime. Define the meromorphic function

$$G_{d,M}(s) = \frac{1}{\varphi(M)} \left\{ \sum_{\xi \bmod M} \left(\frac{\widetilde{L}'}{\widetilde{L}}(s, \xi\chi_K) - \frac{\widetilde{L}'}{\widetilde{L}}(s, \xi) \right) - 1 \right\}. \tag{2}$$

Here and throughout, a sum over characters includes the principal and imprimitive characters. The constant subtracted in the prime error is

$$c_{d,M} = \operatorname{Re} G_{d,M}(1).$$

The classical non-vanishing on $\operatorname{Re} s \geq 1$ of the pole-removed Dirichlet L-functions ensures that this constant exists. In fact $G_{d,M}(1)$ is real, by conjugation of characters.

The error in the compact Lean development, called `primeError`, is exactly

$$R_M(x) = \sum_{1 \leq \nu \leq x} \lambda_{d,M}(\nu) \left(\frac{1}{\nu} - \frac{1}{x} \right) - \frac{\log x}{\varphi(M)} - c_{d,M} \qquad (x > 0). \tag{3}$$

It is important to include the constant as well as the logarithm in the normalization. Omitting it would prevent the Laplace transform used below from being regular at $s = 1$. In modulus one, the constant has a more specific expression. If γ is Euler's constant, write

$$\kappa_K = \gamma - \operatorname{Re} \frac{L'}{L}(1, \chi_K).$$

Since $(s-1)\zeta(s) = 1 + \gamma(s-1) + O((s-1)^2)$,

$$c_{d,1} = -1 - \kappa_K. \tag{4}$$

The number κ_K, called `chowlaSelbergConst` in the code, recurs in the arithmetic estimates: the appearance of the same constant in the norm of a CM theta period is part of the eventual cancellation.

The choice of smoothing in (3) is particularly convenient for the arithmetic proof. For $y > 0$ set $(y)_+ = \max(y, 0)$, and extend this notation to all real numbers. For an integer a and $0 < z \leq a$ one has the exact identity

$$\frac{1}{a} \sum_{\nu=1}^{a} \lambda_{d,M}(\nu) \left(\frac{a}{\nu} - \frac{a}{z} \right)_+ = R_M(z) + \frac{\log z}{\varphi(M)} + c_{d,M}. \tag{5}$$

Indeed the positive part vanishes for $\nu \geq z$ and gives $a(1/\nu - 1/z)$ for $\nu < z$. No asymptotic replacement is involved. The importance of this identity is that the arithmetic construction bounds valuations of jets expressed using precisely such positive parts, with integer order parameter a and scale z. The apparently unusual smoothed prime sum is therefore dictated by a divisibility calculation, not just by analytic convenience.

Our ultimate target is, for every $0 < \varepsilon < 1/100$ and every suitable d,

$$R_{N_0}(x) \leq x^{-1/2+8\varepsilon} \qquad (x \text{ sufficiently large}), \tag{6}$$

where $18480 \mid N_0$ and $d \equiv N_0/2 - 1 \pmod{N_0}$. Only the upper bound occurs in (6). The threshold for x may depend on the fixed parameters N_0, d and ε. We shall explain why such a bound is already strong enough.

Source. The definitions, the constant in modulus one and (5) are in `New/Endgame/PrimeError.lean`, declarations `lambdaClass`, `logDerivMean`, `primeError`, `re_logDerivMean_one`, `primeError_eq_sum`. The function $\widetilde{L}$ is `entireL` in `Lib/General/NumberTheory/QuadField.lean`.

2.2 Why the prime sum contains information about zeros

Set

$$E_M(v) = R_M(e^v), \qquad \Phi_M(s) = \int_0^\infty E_M(v) e^{(1-s)v}\, dv.$$

The variable s is the same complex variable as in the L-functions. The initial integral converges for $\operatorname{Re} s > 2$: for $v \geq 0$, the elementary bound $|\lambda_{d,M}(\nu)| \leq 2\nu$ gives $|E_M(v)| \leq Ce^v$, with C depending on M, d.

Proposition 2.1 (Transform and residues)**.** *The integral Φ_M continues meromorphically, with the expression*

$$\Phi_M(s) = \frac{G_{d,M}(s)/s - c_{d,M}}{s-1}. \tag{7}$$

It is holomorphic in a neighborhood of the closed half-plane $\operatorname{Re} s \geq 1$. For $0 < \operatorname{Re}\rho < 1$ its residue at ρ is

$$\operatorname{res}_{s=\rho} \Phi_M(s) = \frac{\sum_{\xi \bmod M} \left(\operatorname{ord}_\rho \widetilde{L}(\cdot,\xi) - \operatorname{ord}_\rho \widetilde{L}(\cdot,\xi\chi_K)\right)}{\varphi(M)\rho(1-\rho)}. \tag{8}$$

In $\operatorname{Re} s > 0$ any poles are at most simple.

Proof. By orthogonality of characters and the logarithmic derivative of the Euler product, the Dirichlet series of (1), absolutely convergent for $\operatorname{Re} s > 1$, is

$$\begin{aligned} D_{d,M}(s) &= \sum_{\nu \geq 1} \frac{\lambda_{d,M}(\nu)}{\nu^s} \\ &= \frac{1}{\varphi(M)} \sum_{\xi \bmod M} \left(\frac{L'}{L}(s,\xi\chi_K) - \frac{L'}{L}(s,\xi)\right) \\ &= G_{d,M}(s) + \frac{s}{\varphi(M)(s-1)}. \end{aligned}$$

In the last equality the unique principal character modulo M contributes $1/(s-1)$ besides the pole-removed logarithmic derivative, while no twisted character is principal. The additional 1 in the definition of $G_{d,M}$ has now combined with that pole term.

Absolute convergence justifies integrating the sum in (3) term by term for $\operatorname{Re} s > 2$. For an integer $\nu \geq 1$,

$$\int_{\log \nu}^\infty (\nu^{-1} - e^{-v}) e^{(1-s)v}\, dv = \nu^{-s}\left(\frac{1}{s-1} - \frac{1}{s}\right) = \frac{\nu^{-s}}{s(s-1)}.$$

The transform of v is $1/(s-1)^2$, and that of the constant function one is $1/(s-1)$. Consequently

$$\begin{aligned} \Phi_M(s) &= \frac{D_{d,M}(s)}{s(s-1)} - \frac{1}{\varphi(M)(s-1)^2} - \frac{c_{d,M}}{s-1} \\ &= \frac{G_{d,M}(s)/s - c_{d,M}}{s-1}. \end{aligned}$$

This gives the desired continuation. The apparent singularity at 1 is removable because $c_{d,M} = G_{d,M}(1)$. The classical non-vanishing in $\operatorname{Re} s \geq 1$ rules out all other poles there.

If an entire function F has a zero of order j at ρ, then we can write $F(s) = (s-\rho)^j f(s)$ locally, with $f(\rho) \neq 0$. Hence

$$\frac{F'}{F}(s) = \frac{j}{s-\rho} + \frac{f'}{f}(s).$$

This proves (8), including its sign, and shows that the poles are simple. □

The numerator in (8) is a *difference* of zero multiplicities. Thus a bound proving that Φ_M has no poles in a certain half-plane would not, by itself, show that every $L(s,\xi)$ is zero-free there: a pole might cancel between the two sums. The remedy is to choose d so that the finitely many twisted L-functions have no zero at a hypothetical offending zero, and no real zero to the right of the relevant line.

In the code the properties of E_M and its transform are packaged by the predicate `IsPolar`. The assertion `isPolar_primeError` in `New/Endgame/PrimeError.lean` establishes this predicate with the residue function (8) (zero at 1). This abstraction enables the two subsequent analytic arguments to be proved without referring to primes or automorphic forms.

2.3 An elementary form of Landau's one-sided principle

Here is the precise analytic feature that makes an upper bound sufficient. We give the argument, since it explains the sign required in the later choice of the quadratic twist.

Lemma 2.2 (Landau principle)**.** *Let $E : [0,\infty) \to \mathbb{R}$ be measurable with $|E(v)| \le Ce^v$. Suppose its Laplace transform*

$$\Phi(s) = \int_0^\infty E(v)e^{(1-s)v}\,dv$$

has a meromorphic continuation, holomorphic near $\operatorname{Re} s \ge 1$ and with at most simple poles in $\operatorname{Re} s > 0$. Let $\beta > 0$. Assume that

$$E(v) \le e^{(\beta-1)v} \quad (v \text{ sufficiently large})$$

and that the residue of Φ at each real $\rho > \beta$ is non-negative (or zero when the point is regular). Then Φ is holomorphic in $\operatorname{Re} s > \beta$.

Proof. Choose $A \ge 1$ large enough that

$$h(v) = Ae^{\beta v} - e^v E(v) \ge 0 \qquad (v \ge 0).$$

The eventual upper bound guarantees the inequality at large v, and the growth assumption gives it on the remaining bounded interval after enlarging A. For sufficiently large $\operatorname{Re} s$ we then have

$$\Psi(s) := \int_0^\infty h(v)e^{-sv}\,dv = \frac{A}{s-\beta} - \Phi(s).$$

Let σ_c be the real abscissa of convergence of the integral for Ψ. Because $h \ge 0$, convergence holds in the right half-plane $\operatorname{Re} s > \sigma_c$, where the integral is holomorphic; the displayed meromorphic expression agrees with it there by analytic continuation.

Suppose $\sigma_c > \beta$. If Ψ were to have a simple pole at the real point σ_c, its residue would necessarily be non-negative: for real $s > \sigma_c$, the integral is non-negative, so $(s-\sigma_c)\Psi(s) \ge 0$. But the meromorphic formula gives that the residue equals minus the residue of Φ and is therefore non-positive. This shows that the residue is zero and the singularity removable.

A Laplace transform of a non-negative function cannot be holomorphic at its finite abscissa of convergence. To see this directly, take a real $c > \sigma_c$ sufficiently close to σ_c that the Taylor series of the continued Ψ about c converges at a real $\sigma < \sigma_c$. Differentiation under the integral at c gives

$$(-1)^j\Psi^{(j)}(c) = \int_0^\infty v^j h(v)e^{-cv}\,dv \ge 0.$$

Taylor convergence and monotone convergence therefore give

$$\begin{aligned}\infty &> \sum_{j\ge 0} \frac{(-1)^j \Psi^{(j)}(c)(c-\sigma)^j}{j!} \\ &= \int_0^\infty h(v) e^{-cv} \sum_{j\ge 0} \frac{((c-\sigma)v)^j}{j!}\, dv \\ &= \int_0^\infty h(v) e^{-\sigma v}\, dv,\end{aligned}$$

contrary to the definition of σ_c. Thus $\sigma_c \le \beta$, and the formula for Ψ implies the assertion about Φ on $\operatorname{Re} s > \beta$. □

This argument is based on positivity of an auxiliary integrand, not on a hypothesis that E changes sign or has small absolute value. It is the usual Landau principle for Dirichlet integrals with one sign, in the form formalized by `IsPolar.landau` in `Lib/Analytic/Landau/PolarLandau.lean`.

2.4 Quadratic twists avoiding a prescribed half-strip

We now give a proof of the zero-density ingredient in the form required here. In particular it is not necessary to invoke a zero-density theorem with constants that are uniform in the original modulus M. Fix M throughout this subsection. The following elementary argument follows the compact source `Lib/Analytic/QuadraticTwist/ZeroFreeTwist.lean`. Jutila's classical work on quadratic-character mean values [3] provides the context; the particular assertion needed here is proved below.

Put $g_j(t) = t^j e^{-\pi t^2}$ for $j = 0, 1$, and, for a sequence c of modulus at most 1, put

$$\mathcal{V}_j(c;t) = \sum_{h\ge 1} c(h) g_j(ht), \qquad I_j(c;T) = \int_{1/T}^\infty |\mathcal{V}_j(c;t)| \frac{dt}{\sqrt{t}}, \qquad T > 0.$$

The following properties of the Gaussian are used repeatedly:

$$|g_j(t)| \le 3e^{-2t} \quad (t \ge 0), \tag{9}$$

$$|g_j(t)| \le C_\alpha t^{-\alpha} \quad (t > 0,\ \alpha > 0), \tag{10}$$

$$|\mathcal{V}_j(c;t)| \le 3e^{-t}/t. \tag{11}$$

For the first inequality use $t^j \le 1 + t \le e^t$ and maximize $3t - \pi t^2$; this maximum is $9/(4\pi) < \log 3$. The second follows from boundedness of $t^\alpha g_j(t)$. For the third, summing the first gives $3e^{-2t}/(1-e^{-2t}) = 3/(e^{2t}-1) \le 3e^{-t}/t$, since $2\sinh(t) \ge t$.

Lemma 2.3 (Theta transformation and Mellin integral)**.** *Let ψ be a primitive nonprincipal Dirichlet character of conductor q, and choose $j \in \{0,1\}$ by $\psi(-1) = (-1)^j$. The Mellin transform on $\Re s > 0$ of $\mathcal{V}_j(\psi;t)$ is*

$$M_\psi(s) = \int_0^\infty \mathcal{V}_j(\psi;t) t^{s-1}\, dt = \Gamma_j(s) L(s,\psi), \qquad \Gamma_j(s) = \tfrac{1}{2}\pi^{-(s+j)/2}\Gamma((s+j)/2).$$

Moreover

$$\sqrt{q}\, t\, |\mathcal{V}_j(\psi;t)| = |\mathcal{V}_j(\psi;1/(qt))|, \tag{12}$$

and if $q \le T^2$ then

$$\int_0^\infty |\mathcal{V}_j(\psi;t)| \frac{dt}{\sqrt{t}} \le 2 I_j(\psi;T). \tag{13}$$

Proof. Poisson summation on each progression modulo q applies to the Schwartz function g_j on $\mathbb{R}$, whose Fourier transform (with kernel $e^{-2\pi ixy}$) is $(-i)^j g_j$. Expanding the periodic character in finite Fourier coefficients, the Gauss-sum formula for a primitive character gives

$$\mathcal{V}_j(\psi;t) = \frac{\varepsilon(\psi)}{\sqrt{q}\,t}\,\mathcal{V}_j(\overline{\psi};1/(qt)), \quad |\varepsilon(\psi)| = 1.$$

Indeed the two bilateral sums are twice their positive parts, by parity and $\psi(0) = 0$; the modulus of the primitive Gauss sum is $\sqrt{q}$. As g_j is real, the absolute values for ψ and $\overline{\psi}$ coincide. This proves (12). At zero the transformation and (11) show rapid decay; at infinity (11) applies directly. Termwise Mellin integration on $\Re s > 1$ yields the displayed formula, by substituting $u = ht$ and then $z = \pi u^2$. Both sides are holomorphic on $\Re s > 0$, so the equality extends there. Finally the substitution $u = 1/(qt)$ in the integral over $(0, 1/T)$, using (12), identifies it with the integral over $(T/q, \infty)$ against $du/\sqrt{u}$. Since $T/q \geq 1/T$, the result follows. □

This is the ordinary theta Mellin formula; compare [2, Ch. 10] and the formal declarations `gaussianMellin_eq`, `norm_gaussianTheta_functional_equation`, `thetaTailIntegral_zero_le_two_mul` in `Lib/Analytic/DirichletTheta/MellinTheta.lean`.

Proposition 2.4 (A three-lines zero detector)**.** *For $j = 0, 1$ there is an absolute constant C_j such that, whenever $\eta > 0$, $T \geq 1$, $q \leq T^2$ and ψ is as in Lemma 2.3, the conditions*

$$I_j(\psi;T) + I_j(\mu\psi;T) \leq T^{2\eta}, \qquad L(s,\psi) = 0, \quad \tfrac{1}{2} + 9\eta \leq \Re s \leq 1$$

imply $|\Gamma_j(s)|^2 \leq C_j T^{-2\eta}$.

Proof. Here μ is the Möbius function. Define a truncated Mellin integral

$$G(s) = \int_{1/T}^{\infty} \mathcal{V}_j(\mu\psi;t)\,t^{s-1}\,dt, \qquad F(s) = \Gamma_j(s)^2 - M_\psi(s)G(s).$$

It is holomorphic and bounded on the closed strip $\frac{1}{2} \leq \Re s \leq 2$ (with a bound depending on q, T): this follows from the integral estimates just proved, the cutoff in G, and the Gaussian Mellin formula. We use the three-lines theorem with boundary estimates uniform in the imaginary coordinate. On $\Re s = 1/2$, write $I = I_j(\psi;T)$, $J = I_j(\mu\psi;T)$. Lemma 2.3 implies $|M_\psi(s)| \leq 2I$ and $|G(s)| \leq J$. Thus

$$|F(s)| \leq \sup_{\Re z = 1/2} |\Gamma_j(z)|^2 + 2IJ \ll_j T^{4\eta}.$$

The bound on the gamma factor is immediate from its defining integral. On $\Re s = 2$, absolute convergence of the Euler products gives $L(s,\psi)L(s,\mu\psi) = 1$, where the second factor denotes $\sum_{n\geq 1} \mu(n)\psi(n)n^{-s}$. Therefore

$$F(s) = M_\psi(s) \int_0^{1/T} \mathcal{V}_j(\mu\psi;t)\,t^{s-1}\,dt.$$

By (11), the first factor is bounded in modulus by $3\int_0^\infty e^{-t}\,dt = 3$, and the second by $3/T$. Hence $|F(s)| \leq 9/T$, which in particular is $\ll_j T^{4\eta-1}$. Interpolation across the strip of width $3/2$ gives

$$|F(s)| \ll_j T^{4\eta - (\Re s - 1/2)/(3/2)}.$$

When $L(s,\psi) = 0$, $F(s) = \Gamma_j(s)^2$. For $\Re s \geq 1/2 + 9\eta$, the exponent is at most -2η, as required. □

The formal counterpart is `three_lines_bound_of_LFunction_eq_zero`.

Lemma 2.5 (Smoothed quadratic-character mean value)**.** *For $\eta > 0$, $j = 0, 1$, $X \in \mathbb{N}$, $0 < T$ with $X \le T^2$, $t \ge 1/T$, and $|c(h)| \le 1$, one has*

$$\sum_{1 \le v \le X} \left| \sum_{h \ge 1} c(h) J(-v \mid h) g_j(ht) \right|^2 \ll_\eta T^2 t^{-1-2\eta}.$$

Here $J(v \mid h)$ is the product of the Legendre symbols over odd primes dividing h, each with its multiplicity, and the principal character at 2 when h is even. In particular it is multiplicative in the denominator and bounded in absolute value by 1.

Proof. After expanding the square (absolutely convergent by (11)), it is enough to bound

$$\sum_{n,m \ge 1} |g_j(nt) g_j(mt)| \left| \sum_{v \le X} J(-v \mid nm) \right|.$$

If neither nm nor $2nm$ is a square, the Dirichlet character $v \mapsto J(v \mid nm)$ is nonprincipal modulo nm: its odd-prime squarefree kernel is nontrivial. The Pólya–Vinogradov inequality gives the inner bound $\ll \sqrt{nm}(1 + \log nm)$. Use (10) with exponent $3/2 + \eta$ for both Gaussian factors and the estimate $1 + \log(nm) \ll_\eta (nm)^{\eta/2}$. The sum of the resulting factors $(nm)^{-1-\eta/2}$ is finite. The contribution of this part is $\ll_\eta t^{-3-2\eta} \le T^2 t^{-1-2\eta}$.

If knm is a square for some $k \in \{1, 2\}$, use the trivial bound $X \le T^2$. Write $kn = au^2$, $m = av^2$ with $a \ge 1$ squarefree; each pair occurs in this parametrization for the indicated k. Use (10) with exponent $1/2 + \eta$. Ignoring the squarefree restriction on a and the divisibility requirement on u, the sum is bounded, for each k, by

$$C_\eta T^2 t^{-1-2\eta} k^{1/2+\eta} \sum_{a,u,v \ge 1} (auv)^{-1-2\eta}.$$

This proves the assertion. We have used only the classical Pólya–Vinogradov bound for a (possibly imprimitive) nonprincipal character, not a quadratic large sieve. The finite-character inequality can be found in [2, Ch. 9]. □

Corollary 2.6 (Tail-integral bound)**.** *For $X \le T^2$ and $|c(h)| \le 1$,*

$$\sum_{1 \le v \le X} I_j(c(\cdot) J(-v \mid \cdot); T) \ll_\eta T^{2+\eta}.$$

Proof. By Cauchy–Schwarz and Lemma 2.5, the sum over v of the absolute values inside the integral is $\ll_\eta \sqrt{X} T t^{-1/2-\eta} \le C_\eta T^2 t^{-1/2-\eta}$. Integrating against $dt/\sqrt{t}$ from $1/T$ to infinity gives T^η/η times this constant. □

Proposition 2.7 (Zero density in the family of prime quadratic twists)**.** *Fix a Dirichlet character ξ modulo M, $Y \ge 0$, and $0 < \eta < 1/18$. The number of primes $d \le X$, $d \equiv 7 \pmod 8$, $(d, M) = 1$, for which $\widetilde{L}(s, \xi\chi_K)$ has a zero in $\Re s \ge 1/2 + 9\eta$, $|\Im s| \le Y$, is $O_{M,Y,\eta}(X^{1-\eta/2})$.*

Proof. Let ψ of conductor $M' \mid M$ induce ξ. The twist $\psi\chi_K$ is primitive of conductor $q = M'd$, by coprimality of the conductors. As $d \equiv 7 \pmod 8$, its parity j is independent of d and quadratic reciprocity gives $\chi_K(h) = J(-d \mid h)$, including the factor at 2. The extra Euler

factors for $\xi\chi_K$ are nonzero at $\Re s > 0$, and classical non-vanishing gives $\Re s < 1$ for any indicated zero. Put $T = \sqrt{M'X}$. By the zero detector, if

$$I_j(\psi\chi_K; T) + I_j(\mu\psi\chi_K; T) \le T^{2\eta},$$

then $|\Gamma_j(s)|^2 \ll T^{-2\eta}$ at the putative zero. But Γ_j has a positive lower bound in modulus on the fixed compact rectangle $1/2 \le \Re s \le 1$, $|\Im s| \le Y$; hence for large enough X the putative zero forces the sum of tail integrals to be larger than $T^{2\eta}$. Corollary 2.6, for $c = \psi$ and $c = \mu\psi$, bounds the sum of the tail integrals over $d \le X$ by $O_\eta(T^{2+\eta})$. Dividing by $T^{2\eta}$ proves the announced count. □

The mean-square, tail-integral and density proofs above translate `sum_sq_gaussianTheta_jacobiSym_le`, `sum_thetaTailIntegral_le`, and `zero_density_quadTwist` in `ZeroFreeTwist.lean`.

Proposition 2.8 (Auxiliary zero-free twists)**.** *Fix a modulus M, a unit class $a_0 \pmod M$ all of whose primes are $7 \pmod 8$ and prime to M, a number $\beta > 1/2$ and $Y \ge 0$. There exists a prime $d \equiv a_0 \pmod M$ such that every $\widetilde{L}(s, \xi\chi_K)$, $\xi \bmod M$, is nonzero in $\Re s \ge \beta$, $|\Im s| \le Y$.*

Proof. Choose $0 < \eta < 1/18$ such that $1/2 + 9\eta < \beta$. By Proposition 2.7 the union, over the finitely many characters modulo M, of the bad sets has count $O_{M,Y,\eta}(X^{1-\eta/2})$. Its sum of reciprocals of primes converges by dyadic summation. The sum of reciprocals of all primes in the unit class a_0 diverges, by the classical prime number theorem for a fixed arithmetic progression; hence the union does not contain that class. Choose d outside the union. This is `exists_entireL_quadTwist_ne_zero` in the formal source. □

2.5 The one-sided zero-free criterion

Proposition 2.9 (Criterion used in the proof)**.** *Let $\beta > 1/2$, and let M, a_0 satisfy the conditions of Proposition 2.8. Suppose that for* each *prime $d \equiv a_0 \pmod M$ the corresponding prime error (3) satisfies*

$$R_M(x) \le x^{\beta-1} \qquad (x \text{ sufficiently large}).$$

Then all $\widetilde{L}(s, \xi)$ for $\xi \bmod M$ are zero-free in the half-plane $\operatorname{Re} s > \beta$.

Proof. If $\widetilde{L}(\rho, \xi_0) = 0$ for some ρ to the right of β, apply Proposition 2.8 with $Y = |\operatorname{Im}\rho|$. Choose d so that no twisted $\widetilde{L}(s, \xi\chi_K)$ vanishes anywhere in the indicated half-strip. Then, at ρ, the numerator of (8) is a positive integer: the first sum includes the zero of $\widetilde{L}(\cdot, \xi_0)$, while the second sum is zero.

Furthermore at each real point $\sigma \in (\beta, 1)$ the numerator of the same residue formula is non-negative and $\sigma(1-\sigma) > 0$. Thus the real residues of Φ_M to the right of β are non-negative; those at and to the right of 1 are zero by classical non-vanishing and normalization. The assumed bound for the chosen d gives $E_M(v) \le e^{(\beta-1)v}$ eventually. The Landau principle (Lemma 2.2) says that Φ_M is holomorphic on $\operatorname{Re} s > \beta$, contradicting its non-zero residue at ρ. □

The logical order matters: the same arithmetic inequality has to be proved for each admissible auxiliary prime d in the given class, while the proof of the criterion chooses d *after* assuming that an offending zero exists. The threshold in the arithmetic estimate need not be uniform in d, and the choice of d need not work for all zeroes at once. This flexibility is what

avoids a zero-cancellation problem that an argument involving just one fixed quadratic field would face.

The formal declaration is `entireL_ne_zero_of_primeError_le` in `New/Endgame/OneSidedCriterion.lean`.

2.6 A good second scale in a long window

The compact edition employs a second important analytic argument. The arithmetic construction gives a bound involving two scales x, x_1 rather than one, with a combination of the two prime errors $R_{N_0}(x_1)$ and $R_1(x_1)$. We cannot simply discard those terms: their size relative to the desired saving at x would need justification. Instead, we have the freedom to vary $\log x_1$ in a sufficiently long interval. A high-moment argument produces a good choice of the second scale without requiring a pointwise prime number theorem error estimate there.

Lemma 2.10 (Window lemma)**.** *Let $E, E' : [0,\infty) \to \mathbb{R}$ be measurable with $|E(v)| + |E'(v)| \le Ce^v$. Suppose their Laplace transforms, with kernel $e^{(1-s)v}$, continue holomorphically to a neighborhood of the closed half-plane $\operatorname{Re} s \ge 1$. For every sufficiently large T and every $0 \le X \le e^T$ there exists $u \in [T, T^2]$ such that*

$$E(u) - XE'(u) \le e^{-T}.$$

The threshold for T may depend on the two functions but not on X in this range.

Proof. The closed disc of radius 3 about 4 lies in $\operatorname{Re} s \ge 1$. Choose a fixed $3 < R < 4$ such that both continued transforms are holomorphic on a neighborhood of the closed disc of radius R about 4. Such an R exists by compactness. Differentiation under the original integral near $s = 4$ gives

$$\Phi^{(j)}(4) = \int_0^\infty (-v)^j e^{-3v} E(v)\, dv.$$

The Cauchy estimates for Taylor coefficients yield, for all sufficiently large T, simultaneously for every integer $j \ge 0$,

$$\left| \int_0^\infty v^j e^{-3v} E(v)\, dv \right| \le \frac{j!T}{R^j}.$$

The same holds for E'. We can also require $|E(v)|, |E'(v)| \le Te^v$ for all $v \ge 0$.

Write $k = \lceil T^2/9 \rceil$ and suppose, for a contradiction, that $E(v) - XE'(v) > e^{-T}$ throughout $[T, T^2]$. Consider the function

$$v^k \left(e^{7T - Rv} + e^{2T - T^2 - v} + e^{-3v}(E(v) - XE'(v) - e^{-T}) \right).$$

On the window this is non-negative by hypothesis. Outside the window, the elementary growth bounds give

$$e^{-3v}|E(v) - XE'(v) - e^{-T}| \le e^{2T-2v} \qquad (v \ge 0),$$

for sufficiently large T, uniformly in $0 \le X \le e^T$. For $0 \le v \le T$, the first exponential in the parentheses dominates this upper bound, since $R < 4$. For $v \ge T^2$, the second exponential dominates it. Thus the integrand is non-negative everywhere.

Integrate over $v \ge 0$, use the moment bounds above, and recall $\int_0^\infty v^k e^{-qv} dv = k!/q^{k+1}$ for $q > 0$. The result, after division by $k!$, is

$$0 \le \frac{e^{7T}}{R^{k+1}} + e^{2T-T^2} + \frac{T(1+X)}{R^k} - \frac{e^{-T}}{3^{k+1}}.$$

This is impossible for large T. Divide all terms by the magnitude $e^{-T}/3^{k+1}$ of the negative term. The three positive ratios are respectively

$$e^{8T}(3/R)^{k+1}, \quad e^{3T-T^2}3^{k+1}, \quad 3e^T T(1+X)(3/R)^k.$$

The first and third tend to zero as $T \to \infty$, since k is quadratic in T and $3/R < 1$; the convergence is uniform for the third ratio by $X \le e^T$. The second ratio tends to zero because $(\log 3)/9 < 1$. Their sum is less than one for all sufficiently large T, giving the desired contradiction. □

This proof is an explicit example of what a genuine shortening of the formal development looks like when translated into ordinary mathematics. The readable edition uses a pointwise prime number theorem error bound at its much larger second cutoff. The compact edition needs only regularity of the two transforms in $\operatorname{Re} s \ge 1$ and a good point in a window. It is the ability to choose that point, not an assumed strong prime number theorem, that makes the argument work.

Apply the lemma with $E(v) = R_{N_0}(e^v)$ and $E'(v) = R_1(e^v)$ for the same auxiliary prime d. The hypotheses follow from Proposition 2.1. If $T = 10^6 \log x$ and X is the positive multiplier supplied by the arithmetic data, with $X \le e^T$, the lemma permits the choice $x_1 = e^u$ with

$$\log x_1 \in [T, T^2], \qquad R_{N_0}(x_1) - XR_1(x_1) \le x^{-10^6}. \tag{14}$$

This is a one-sided bound for a particular linear combination at a particular point; it should not be confused with a uniform pointwise estimate for either error. In the final argument the change of x_1 has to preserve the admissibility of the automorphic data. The scales are chosen in the arithmetic construction precisely so that this is possible.

Source. The moment and window lemmas are `IsPolar.eventually_moment_le` and `IsPolar.exists_window_sub_mul_le` in `New/Endgame/OneSidedCriterion.lean`; `New/Main.lean` applies them to the remainder in the final bound `eventually_primeError_le_rpow`.

2.7 From the target bound to GRH

To see precisely the quantifiers needed at the end, suppose that (6) has been established by the arithmetic construction. Let $q \ge 1$ be any modulus, $\chi \bmod q$ a Dirichlet character, and ρ a hypothetical zero with $\operatorname{Re}\rho > 1/2$, with the pole of a principal character excepted. Take $N_0 = 18480q$ and the unit class $a_0 = N_0/2 - 1 \pmod{N_0}$. Since $16 \mid N_0$, $a_0 \equiv 7 \pmod 8$ and $a_0^2 \equiv 1 \pmod{N_0}$. Choose $\varepsilon > 0$ with $\varepsilon < 1/100$ and $\beta = 1/2 + 8\varepsilon < \operatorname{Re}\rho$. The estimate (6) holds for every prime $d \equiv a_0 \pmod{N_0}$ and gives $R_{N_0}(x) \le x^{\beta-1}$ eventually. Proposition 2.9 implies that all pole-removed Dirichlet L-functions modulo N_0 are non-zero at ρ.

Changing the level of χ to N_0 multiplies its Dirichlet L-function by a finite product of omitted Euler factors, each non-zero in $\operatorname{Re} s > 0$. Thus a zero of $\widetilde{L}(\cdot, \chi)$ at ρ persists at the new level. This contradicts the preceding conclusion. We have proved, subject only to the arithmetic target estimate,

$$L(s, \chi) \ne 0 \qquad (\operatorname{Re} s > 1/2),$$

apart from the pole of the principal character at $s = 1$.

For a primitive character the functional equation, whose gamma factors are non-zero and finite in the open critical strip, sends a zero at s to a zero for the dual character at $1 - s$. Thus there can be no zero with $0 < \operatorname{Re} s < 1/2$ either. For an imprimitive character the omitted Euler factors again do not vanish in the open strip. Hence every zero there has real part $1/2$. This includes the Riemann zeta function (modulus one); its standard functional equation and trivial-zero classification give the usual full statement of RH.

The rest of the article is concerned with the provenance of (6): how a problem about the lengths of automorphic forms and the divisibility of their theta lifts yields exactly the exponent needed for this analytic reduction.

3 The admissible data and coefficient spaces

3.1 The CM extension and the different

We first record the precise arithmetic of the extension in the compact development. Fix $N_0 > 0$ with $18480 \mid N_0$, and a prime $d \equiv N_0/2 - 1 \pmod{N_0}$. Equivalently $N_0 \mid 2(d+1)$ and $N_0 \nmid d+1$. In particular $d \equiv 7 \pmod 8$, $d > 64$, and $(d, N_0) = 1$. Put $K = \mathbb{Q}(\sqrt{-d})$, of discriminant $-d$. For an integer $s \geq 1$ define

$$c_s = \lfloor \log_3 s \rfloor + 5, \quad T_s = 3^{c_s} 2^s, \quad N_1 = N_0 T_s, \quad E = K(\zeta_{N_1}), \quad M_\omega = (N_0 d)^{s+6}.$$

Here T_s is a number, not a set of places, and $T_s = |\mathcal{T}|$ below. For each prime $p \mid N_1$ the character $\chi_K(p) = 1$. Indeed for an odd such prime p, $p \mid (d+1)/4$ since all the primes of N_1 divide N_0; hence $-d \equiv 1 \pmod p$. For $p = 2$ the claim follows from $d \equiv 7 \pmod 8$. Consequently every nonsplit prime of K is at least 13, and d is prime to N_1. Since K and $\mathbb{Q}(\zeta_{N_1})$ are linearly disjoint (their discriminants have disjoint prime supports),

$$E \subset \mathbb{Q}(\zeta_{N_1 d}), \qquad [E : K] = \varphi(N_1) = T_s \varphi(N_0), \qquad \mathrm{Gal}(E/K) \simeq (\mathbb{Z}/N_1)^\times.$$

In the inclusion we use the quadratic Gauss sum of conductor d. The field E is totally imaginary and complex conjugation is central in its Galois group over $\mathbb{Q}$, so E is a CM field with real subfield E^+. It is unramified at the finite places: at every residue characteristic at least one of the primitive third and fourth roots of unity in E has order prime to that characteristic, and complex conjugation acts nontrivially on its reduction. Thus complex conjugation cannot belong to an inertia group of $E/\mathbb{Q}$. The roots of unity in E are exactly the N_1-th roots: the cyclotomic field of even order $N_1 d$ has no other roots of unity, and if E contained a primitive d-th root it would have degree over $\mathbb{Q}(\zeta_{N_1})$ at least $d - 1 > 2$. Also $N_1 d \mid M_\omega$, by inspecting the powers of 2, 3 and the remaining primes.

Let Φ be the set of archimedean places, each represented by the embedding extending our fixed embedding of K. We identify Φ with $\mathrm{Gal}(E/K)$ using one chosen embedding of E. Write $[j] \in \Phi$ for the residue class of j modulo N_1 if it is invertible, and for 1 otherwise (this latter convention occurs in the formal prime sums but contributes nothing when $\chi_K(p) = 1$). Put $\mathcal{T} = \{\tau \in \Phi : \tau \equiv 1 \pmod{N_0}\}$. Then

$$|\mathcal{T}| = T_s, \qquad f = |\Phi| = \varphi(N_0) T_s.$$

For primes $p \nmid N_1 d$ the Frobenius in $\mathrm{Gal}(E/\mathbb{Q})$ has components $(\chi_K(p)p, \chi_K(p))$ in the identification $\Phi \times \{\pm 1\}$ in which the second factor is complex conjugation. This follows because the actual automorphism sends ζ_{N_1} to its p-th power and acts on K by $\chi_K(p)$. In particular a prime of E above p has residual degree the order of this pair. For inert p in K, the sign correction $-p^j$ at odd iterates is the origin of the place permutations in the local prime sums (Section 2). For these assertions see `New/Data/Different.lean` and `New/Data/WeilDataOfE.lean`.

The coprime discriminants imply $\mathcal{O}_E = \mathcal{O}_K[\zeta_{N_1}]$: the integral tensor product is maximal after localizing at each prime, since at least one factor is unramified there. Thus the relative different over K is generated by $P'(\zeta)$ for $P(X) = \Phi_{N_1}(X)$, the cyclotomic polynomial, and the absolute different is generated by $\sqrt{-d}\, P'(\zeta)$. The standard cyclotomic integer/different

facts are recalled in [5]. The reciprocity relation $P(X) = X^D P(1/X)$, with $D = \deg P$ even, gives

$$P'(\zeta^{-1}) = -\zeta^{2-D} P'(\zeta).$$

Consequently multiplication by $i\zeta^{1-D/2}$ makes $P'(\zeta)$ real. We choose a purely imaginary generator ϖ of the absolute different by multiplying the resulting real element by $\sqrt{-d}$. All conjugates of the root-of-unity multiplier $i\zeta^{1-D/2}$ have modulus 1. In particular

$$|\tau(\varpi)| = \sqrt{d}|P'(\tau\zeta)|, \qquad \mathrm{rd}_E^f = \prod_{\tau\in\Phi} |\tau(\varpi)|, \tag{15}$$

where $\mathrm{rd}_E = |\operatorname{disc} E|^{1/[E:\mathbb{Q}]}$. The second equality follows by taking the square root of the absolute norm of the different. The identity

$$\Phi_{N_1}(X) = \Phi_{N_0}(X^{T_s})$$

follows either by listing roots or by successively applying $\Phi_{pN}(X) = \Phi_N(X^p)$ for $p \mid N$. Differentiating at ζ and using $|\tau\zeta| = 1$ shows that every $|P'(\tau\zeta)|$ is T_s times one of the fixed positive numbers $|\Phi'_{N_0}(\zeta_0)|$ for primitive N_0-th roots ζ_0. Hence $\mathrm{rd}_E \asymp_{N_0,d} T_s$. More specifically along the data below, for all sufficiently large x,

$$x^{10} \le |\tau(\varpi)|, \mathrm{rd}_E \le x^{12}. \tag{16}$$

This is the numerical content of `hasLogBound_varpi` and `hasLogBound_rootDisc` in `New/Seesaw/IndexTypes.lean`. The ensuing Weil data also choose a splitting character and compatible finite ideles; we retain exactly those choices of the compact source, since different choices of theta normalizations can change the period and volume identities used below.

3.2 The scales, and the reason for so many parameters

The principal scale is x; the goal is to estimate $R_{N_0}(x)$. The proof uses another cutoff x_1 and a large order scale b, with

$$(\log b)^{100} \le x, \qquad x^{10^6} \le x_1, \qquad x_1^{10^6} \le b, \qquad (\log x_1)^6 \le \log b.$$

These inequalities are intentionally very far apart. They allow polynomial losses in the ranks or field degree to be negligible compared with the principal weights. For each fixed $0 < \varepsilon < 1/100$, the construction starts with $\log b = x^\varepsilon$ and $\log x_1 = x^{\varepsilon/6}$. The second cutoff will later be lowered within an admissible window. In the compact edition it is a real scale; no rounding of x is required.

Parameter	Meaning	Size or relation
m	Number of lifts	$\sqrt{x} \le m \le 2\sqrt{x}$
r	Rank of initial target group	$x^5 \le r \le 3x^5$
k	Rank of initial source group	$k = r + 1 - m$
t	Growth of row ranks per stage	$t = 5r$
h_2	Half-gap in the weights	$h_2 = r^8$
N_1	Cyclotomic order	$N_0 \mid N_1$
$\lvert\mathcal{T}\rvert$	Number of tops	$\asymp x^{11}$ up to log factors
$\lvert\mathcal{S}\rvert$	Number of thick places	$\ge \lvert\mathcal{T}\rvert / \log b$
a	Order of distinguished derivatives	$xb \le a < xb + N_1$
$g = g_0 g_1$	Product of two level primes	$b^{20} \le g \le b^{22}$
N	Extra level prime for source	$m \log b \le \log N \le 30m \log b$

The exact powers 10^6, 100 and 99 which occur in the source are generous margins; one cannot replace them without checking the many subsidiary inequalities, but they are not the source of the exponent $1/2$ in GRH. The significant power is the square-root length of the chain. The choice of m as a power of 2 and r as a power of 3 is compatible with the displayed intervals and serves the discrete combinatorial constructions in the proof.

The effective weights at the infinite places come in two sizes:

$$w_{\mathrm{nom}}(\sigma) = \begin{cases} a/x, & \sigma \in \mathcal{S}, \\ a/x_1, & \sigma \notin \mathcal{S}. \end{cases}$$

The integral weights actually used are rounded upwards to multiples of N_1 and shifted: if $w^-(\sigma)$ is the rounding of $w_{\mathrm{nom}}(\sigma)$ then $w_j(\sigma) = w^-(\sigma) + 2jh_2 + h_2$. Hence neighboring weights have a large gap. This separation permits quantitative archimedean and local estimates while the naive weights still record the two prime cutoffs x and x_1. It is important not to confuse the large weights with the logarithmic scale $\log b$.

3.3 Admissibility and construction of the data

It is useful to give the full existence proof rather than appeal to the entirety of the formal parameter-search module. The following is the portion of the definition of `Data` in `New/Data/AdmissibleData.lean` that is needed to begin the construction. Write $Y = \log b$, $U_1 = \log x_1$. There is one fixed positive constant c_* depending only on N_0, d (the source's `Parameters.c`, which includes the uniform CM generator constant), and the scales satisfy

$$10^{12} c_* \le \log Y, \quad 100 \log Y \le \log x, \quad 10^6 \log x \le U_1, \quad 10^6 U_1 \le Y, \quad 6 \log U_1 \le \log Y. \tag{17}$$

In addition to the rank intervals in the preceding table, require $m = 2^u$ with $u \ge 1$, $s = 22u$, $r = 3^v$ with $v \ge 1$, and $s \le 32 \log x$. The thick set lies in the coset -1 modulo $N_0 3^{c_s}$ and has the following property. If $\phi : \Phi \to (\mathbb{Z}/2^{s+1})^\times$ is reduction and $0 < i < j \le 324s$, then

$$\sigma, \tau, \rho \in \mathcal{S}, \quad \phi(\tau/\sigma)^j = \phi(\rho/\sigma)^i \implies \phi(\tau/\sigma) = 1. \tag{18}$$

The two distinct primes g_i lie in $[b^{10}, b^{11}]$ and are 1 modulo M_ω^2; there is $\pi_g \in \mathcal{O}_K$ with $\pi_g \bar{\pi}_g = g_0 g_1$. Finally the data include prime existence in the windows

$$\forall U \in [100Y, 2^s] \quad \exists p \text{ prime}, \quad p \equiv 1 \pmod{M_\omega^2}, \quad U \le \log p \le U + \sqrt{Y}. \tag{19}$$

The source defines the entire record before constructing any automorphic forms. All our subsequent spaces depend on this common record.

Lemma 3.1 (Behrend construction in a projected cyclic group)**.** *Let G and H be abelian groups, $\phi : G \to H$ a homomorphism, $u, z \in G$, and suppose the order of $\phi(u)$ is D. Let L_0, h, Q be positive integers with $(L_0 Q)^h \le D$ and $L_0 \ge 2$, $Q \ge 2$. There is a subset S of $z\langle u \rangle$ of size at least $Q^h/(hQ^2)$ such that, for $s_0, s_1, s_2 \in S$ and $0 < i < j \le L_0$, the relation $\phi(s_1/s_0)^j = \phi(s_2/s_0)^i$ implies $s_1 = s_0$.*

Proof. Consider the Q^h integer vectors $v = (v_0, \dots, v_{h-1})$ with $0 \le v_a < Q$. The squared Euclidean lengths have at most $h(Q-1)^2 + 1 \le hQ^2$ values; choose a sphere containing at least the asserted number of vectors. Encode a vector by $e(v) = \sum_{a=0}^{h-1} v_a (L_0 Q)^a$ and take $s(v) = z u^{e(v)}$. The map $v \mapsto \phi(s(v))$ is injective since $0 \le e(v) < D$ and the encoding is injective. The assumed multiplicative relation becomes

$$j e(v_1) \equiv i e(v_2) + (j - i) e(v_0) \pmod{D}.$$

Both sides are integers with digits at most $L_0(Q-1) < L_0Q$ in base L_0Q, hence lie in $[0,(L_0Q)^h) \subset [0,D)$. The equality therefore holds digit by digit and $v_1 = (i/j)v_2 + (1-i/j)v_0$. On a Euclidean sphere a strictly interior convex combination has strictly smaller norm unless the two endpoints are equal. Thus $v_0 = v_1 = v_2$. □

This is the combinatorial ingredient of `exists_three_term_free` in `New/Data/DataExist.lean`; it is a simultaneous Behrend-type construction (compare [4]).

Proposition 3.2 (Existence and flexibility of admissible data)**.** *For fixed N_0, d, $0 < \varepsilon < 1/100$, and every sufficiently large x, one can construct admissible data $\mathcal{D}$ satisfying*

$$Y = x^{\varepsilon}, \qquad U_1 = x^{\varepsilon/6} \geq (10^6 \log x)^2, \qquad |\mathcal{S}| \geq \frac{T_s}{Y}. \tag{20}$$

The same record is admissible if U_1 is replaced by any $u \in [10^6 \log x, (10^6 \log x)^2]$, interpreting every derived quantity with the new U_1.

Proof. Take m and r to be the least powers of 2 and 3 which are at least $\sqrt{x}$ and x^5, respectively, and choose s as above. Then $2^s = m^{22}$, $s \leq 32 \log x$ eventually, and

$$81s < 3^{c_s} \leq 243s, \qquad T_s = 3^{c_s} m^{22} \asymp x^{11} \log x.$$

In particular $x^{11} \leq T_s \leq x^{11} Y^{1/5}$ for sufficiently large x. The choice of Y, U_1 in (20) satisfies (17): the two relevant upper comparisons hold as $x^{\varepsilon/6} = o(x^{\varepsilon})$ and $6 \log U_1 = \log Y$; the remaining lower bounds follow from $\varepsilon < 1/100$. This proves the size comparisons used in (16).

To choose $\mathcal{S}$, apply Lemma 3.1 with $z = [-1]$, $u = [1 + M']$, $M' = N_0 3^{c_s}$ and $L_0 = 324s$. Put $v_0 = v_2(N_0) \geq 4$. Since $v_2(M') = v_0$ and the order of $1 + M'$ modulo 2^{s+1} is $D = 2^{s+1-v_0}$ when $s + 1 > v_0$, every power lies in the required coset after multiplication by z. The order formula follows inductively from $v_2((1+M')^{2^h} - 1) = v_0 + h$, by factoring successive differences of squares. Fix an integer $h > 44/\varepsilon$ and choose $Q = \lfloor D^{1/h}/L_0 \rfloor$, which is at least 2 for large x. Then

$$|\mathcal{S}| \geq \frac{Q^{h-2}}{h} \gg_{\varepsilon, N_0} \frac{2^{s(1-2/h)}}{(324s)^{h-2}}.$$

For the desired lower bound, compare with $T_s/Y \asymp s 2^s x^{-\varepsilon}$. Since $2^s \asymp x^{11}$, the ratio of the displayed lower bound to T_s/Y is bounded below by a positive constant times $x^{\varepsilon - 22/h}/(\log x)^{h-1}$, which tends to infinity. The sparsity property is immediate from the lemma.

It remains to find the primes. Here a standard theorem replaces a substantial segment of the formal background: the Siegel–Walfisz theorem, in the usual form uniform for moduli at most a fixed power of the logarithm of the argument ([2, Cor. 11.19]). We use exactly the following consequence. For each fixed $A > 0$, if $Q_0 \leq Y^A$, $U \geq 10Y$ and $Y \to \infty$, then there is a prime $p \equiv 1 \pmod{Q_0}$ with $U \leq \log p \leq U + \sqrt{Y}$, uniformly in U. Indeed, with $v = e^{U+\sqrt{Y}}$, $u = e^U$, the two asymptotics for the numbers of primes in the class 1 mod Q_0 have main terms $\mathrm{Li}(v)/\varphi(Q_0)$ and $\mathrm{Li}(u)/\varphi(Q_0)$. Their ratio tends to infinity uniformly, while their relative errors are $o(1)$, by choosing an exponent in Siegel–Walfisz larger than $A + 2$. This applies because $Q_0 = M_\omega^2 = \exp(O_{N_0,d}(s)) \leq Y^A$ for a fixed A depending on ε, N_0, d. It establishes (19), and also the existence of distinct primes $p_i \equiv 1 \pmod{M_\omega^2}$ in disjoint windows in $[10Y, 11Y]$, with $i = 1, \ldots, h_K + 1$ where h_K denotes the class number of K. For example use the windows $[10Y + 2i\sqrt{Y}, 10Y + (2i+1)\sqrt{Y}]$. All p_i split in K, since $p_i \equiv 1 \pmod{d}$. Choose a prime ideal $\mathfrak{p}_i$ of K above each. By pigeonhole there exist distinct i, j with $[\mathfrak{p}_i] = [\mathfrak{p}_j]$ in the ideal class group. Then $\mathfrak{p}_i \bar{\mathfrak{p}}_j$ is principal; an integral generator π_g satisfies $\pi_g \bar{\pi}_g = p_i p_j$. Put $g_0 = p_i, g_1 = p_j$. This completes the construction of the record.

Lastly, any replacement value of U_1 in the asserted interval is positive, no larger than $x^{\varepsilon/6}$ and at least $10^6 \log x$. It preserves all inequalities in (17), while the other fields of the record do not depend on U_1. This proves flexibility. In the code this is used by `eventually_prime Error_le_rpow` to replace the scale in the structure before constructing any derived weights or levels. □

This supplies a conventional proof of `exists_data` using a classical uniform prime-distribution theorem, rather than reproducing the source's weaker prime-window lemma `exists_prime _modEq_one_log_mem`, in `Lib/Analytic/Primes/AuxiliaryPrimes.lean`. The latter avoids invoking a quantitative prime number theorem for progressions. It is not the same analytic use as the pointwise prime number theorem at the second cutoff: the latter is avoided altogether by the window lemma.

3.4 Derived weights, level and integral structures

In the actual integral weights write $a = N_1\lceil xb/N_1\rceil$, $a_1 = N_1\lceil (a/x)/N_1\rceil$ and $a_2 = N_1\lceil (a/x_1)/N_1\rceil$ (denoted b_1, b_2 in the formal source and in the local chapters). Thus $w_0(\sigma) - h_2$ is a_1 or a_2, with $w_j(\sigma) = w_0(\sigma) + 2jh_2$. For $0 \le j \le 2m$ we have

$$w_{\mathrm{nom}}(\sigma) \le w_j(\sigma) \le w_{\mathrm{nom}}(\sigma) + x^{42}, \quad w_j(\sigma) \le 2b, \quad \log w_j(\sigma) \le \log w_{\mathrm{nom}}(\sigma) + 1/x. \tag{21}$$

For the first upper bound use $N_1 + (4m+1)r^8 = O(x^{81/2}) + O(x^{11}\log x) \le x^{42}$ eventually. Next $w_{\mathrm{nom}} \le a/x \le b + N_1/x$, and $b \gg x^A$ for every fixed A. Finally $w_{\mathrm{nom}} \ge a/x_1 \ge xb/x_1 \gg x^{43}$ by the scale assumptions, and $\log(1 + x^{42}/w_{\mathrm{nom}}) \le x^{42}/w_{\mathrm{nom}} \le 1/x$. These translate `w_le_nominalWeight_add`, `w_le_two_mul_b` and `log_w_le` in `New/Data/Weights.lean`.

Write $w_* = f^{-1}\sum_\sigma \log w_{\mathrm{nom}}(\sigma)$ and $G_* = \log(g\,\mathrm{rd}_E/(2\pi))$, where $g = g_0 g_1$. The extra prime N is selected as in `Data.N`, using the logarithmic target

$$\lambda_N = m(w_* + G_*) - G_* - \tfrac{15}{2}\log b, \quad \lambda_N \le \log N \le \lambda_N + \sqrt{\log b}, \quad N \equiv 1 \pmod{M_\omega^2}. \tag{22}$$

The large ranges give $w_* = (1 + o(1))Y$ and $G_* \in [(20 - o(1))Y, (22 + o(1))Y]$, uniformly after replacement of U_1. Hence eventually $mY \le \lambda_N \le 29mY$, $100Y \le mY$ and $29mY \le 2^s = m^{22}$; this uses $Y = x^\varepsilon$. We may therefore use (19), giving $mY \le \log N \le 30mY$. Since $4 \mid N_1 \mid M_\omega$, Fermat's two-square theorem furnishes $\pi_N = u + iv \in \mathcal{O}_E$ with $\pi_N\bar\pi_N = N$. Every prime above N has residual degree one, and the support of (π_N) contains exactly one member of each conjugate pair (because N is split and unramified and the exponents sum to 1). Likewise the chosen places above g_0, g_1 are the prime divisors of $\bar\pi_g$; they are conjugate-disjoint and there are f above each. These observations justify the finite groups and the determinant normalizations used later.

3.5 Two kinds of automorphic forms

If V is a nondegenerate Hermitian space over E with respect to the involution of E/E^+, $U(V)$ denotes its unitary group over the totally real field E^+. Automorphic forms on such a group are functions (with the specified transformation law at infinity) on an adelic quotient. For a totally definite Hermitian space, the group is compact at infinity, and spaces of algebraic modular forms at a fixed level can be described by values in an algebraic representation on finitely many double cosets. In particular their coefficient integrality and Hermitian pairings can be studied by actual lattice calculations. For the indefinite groups, one instead uses cusp forms of prescribed discrete-series type at infinity. The source imposes suitable finite level and controls their behavior in the cusp.

The proof uses both types. Here are the main spaces; indices on Hermitian spaces should not be confused with indices of vectors.

- The source Hermitian spaces X_j are definite of rank $k + j(t+1)$, $0 \le j \le m$. Let C denote the space of algebraic modular forms on $U(X_0)$ of weight $(a, 0, \dots, 0)$ at tops and zero at the other infinite places, with the prescribed level at g and N. Let $C^{\mathrm{ram}} \subset C$ be the subspace of forms with vanishing averages under the prescribed local scalings at the places of N. These are the "ramified columns."
- The row spaces are forms on Y_j of rank $r + jt$. The starting Y_0 is indefinite, and Y_m is definite. Let W be the selected subspace of forms on $U(Y_0)$, at its prescribed archimedean type and at level g. (We use the letter W only for this vector space; the source also uses W_j for auxiliary rank-r Hermitian spaces.) The spaces selected by a finite-group projector are called good forms.
- For the determinant computation there are also single-lift auxiliary pairs of Hermitian spaces W_j, V_j of ranks $r, r + t$. These are used to replace the norm calculation for a whole chain by the calculations associated to individual lifts and their comparisons.

At the chosen finite places v above g, the principal-level form spaces carry an action of the finite group

$$\mathcal{G} = \prod_v \mathrm{GL}_r(\mathbb{F}_v).$$

The good projector e_0 is the sum of specified idempotents in $\mathbb{C}[\mathcal{G}]$. The good labels are defined through squarefree polynomials over the residue fields, with exclusions at roots of unity and primitive-polynomial conditions. This is more than a harmless way to cut down the dimension: it is designed to guarantee local representation types favorable both to the filling argument and to the Galois-theoretic analysis of congruences. The precise conditions are discussed in the subsequent sections. We write $n = \dim W$ and $c = \dim C$; neither of these is a Hermitian rank.

3.6 Precise weights and levels of the source and target

For a weight at a fixed infinite place we use column multiplicities $\kappa_\tau(h)$, meaning the number of Young-diagram columns of height h. Write $h_0 = 2h_2$ for the weight gap to distinguish it from the dummy height h, and put $n_j^X = k + j(t+1)$. The compact source uses the same multiplicity function at every stage, restricted to heights strictly below the rank of the stage:

$$\kappa_\tau(h) = \begin{cases} a\mathbf{1}_{\tau\in\mathcal{T}}, & h = 1, \\ w_1(\tau) - \lfloor (k+t+1)/2 \rfloor, & h = k, \\ h_0 - (t+1), & h > k,\ h \equiv k \pmod{t+1}, \\ 0, & \text{otherwise.} \end{cases}$$

The determinant exponents are $e_0(\tau) = 0$ and $e_j(\tau) = \lfloor n_{j-1}^X/2 \rfloor - w_j(\tau)$ for $j \ge 1$. These are `Data.` followed by the Greek letter κ, and `Data.lastEntry` in `New/Curve/LiveSum.lean` (the source uses Greek letters in the identifiers). At X_0 the only column heights below k are the height-one columns supported on the tops, and this recovers the source weight described above. At the endpoint, the branching to Y_m retains

$$u(h) = \begin{cases} h_2 - \lfloor (t+1)/2 \rfloor, & h \ge k,\ h \equiv k \pmod{t+1}, \\ 0, & \text{otherwise,} \end{cases}$$

columns of their former height; its new multiplicity is $u(h) + \kappa_\tau(h+1) - u(h+1)$. Its determinant normalization comes from `restrictToY` in `New/Columns/ColumnDefs.lean`. This weight convention matters: it is not permissible to use the ordinary weight w_{nom} in the representation instead of the rounded weights.

Put $F = E^+$. For a standard definite unitary group of rank q', the coefficient space $V_{\kappa,e}$ is the product-place tensor of the complex Weyl modules with multiplicities $\kappa_\tau(h)$, tensored with the determinant of exponent $e(\tau)$. The basis in the code is its integral Weyl-coordinate basis, and its Gram form has an integral, positive-definite matrix G_κ; the determinant twist is unitary at infinity. The action of $\gamma \in U(q')(F)$ uses $\tau(\overline{\gamma})$ in the algebraic weight representation. For an open compact $K' \subset U(q')(\mathbb{A}_{F,f})$, the algebraic forms of fixed level are functions v on the finite adelic group with values in $V_{\kappa,e}$ satisfying

$$v(\gamma g k') = \rho_{\kappa,e}(\gamma) v(g), \qquad \gamma \in U(q')(F),\ k' \in K'.$$

There are finitely many adelic classes. Choosing a representative g of each class gives the direct sum of invariant coefficient spaces under the groups $\Gamma_g = U(q')(F) \cap gK'g^{-1}$. This follows at once from the equivariance condition. It is the model `AlgForm` of `Lib/GLn/SchurM odule/WeightRepBasic.lean`. The class-sum Hermitian form is

$$\langle v, w \rangle_{K'} = \sum_{[g]} \frac{v(g)^* G_\kappa w(g)}{|\Gamma_g|}.$$

When the level K' has Haar volume one, this class sum is the integral of the invariant scalar pairing over the finite adelic quotient. For the correspondence with highest-weight matrix coefficients of the automorphic form at infinity, the source uses the factor $\dim V_{\kappa,e}$ from Schur orthogonality and the factor $[K_0 : K']$ from its finite measure convention. The precise comparison is `integral_mul_eq_classIntegral`, `Space.algFormEquiv` in `Lib/Adelic/Cus pForms/AlgebraicForms.lean`.

Let K_g denote the level which is principal at the chosen places above g and maximal integral elsewhere. For X_j the column level is K_g intersected with the condition that the last row of the matrix at places dividing π_N is $(0, \dots, 0, 1)$ modulo that prime. For each $w \mid \pi_N$, a scalar finite adelic unit z_w represents a generator of $\mathbb{F}_N^\times$ in the compatible standard split model. The ramified part of C is exactly

$$C^{\mathrm{ram}} = \mathrm{im} \prod_{w \mid \pi_N} (1 - \rho(z_w)).$$

The scalar operators commute and are of finite order on this level, so this is equivalently the intersection of the kernels of their cyclic averaging projections. It follows that for each fixed w, every ramified column has the form $(1 - \rho(z_w))v$; this elementary observation is used in the ramified kernel integrals. The codimension bound is established in Section 6 via the precise local level and class count. Source: `New/Columns/ChainValuation.lean`, `Data.ramifiedColumns` and `exists_eq_sub_centralGen`.

For the row representation, number the basis of Y_0 by $i = 0, \dots, r-1$, using the numbering compatible with the Bessel flag in the source. At τ the (undoubled) Harish-Chandra parameter is

$$\varsigma_\tau(i) = \begin{cases} k/2 - 1 - i, & i + 1 < k, \\ h_2 - w_{i+2-k}(\tau), & i + 1 \geq k. \end{cases}$$

Here k is even. The Harish-Chandra coordinates of all the row spaces, together with the signs specifying their real form, are chosen so that the parameter gap between coordinates of opposite signs is at least $2h_2$ in the doubled coordinate convention of the code. In particular

the rank-r auxiliary space W_j is specified by the first $k+j-1$ positive and remaining $m-j$ negative coordinates, while Y_j has rank $r+jt$ and $\lfloor (m-j)/2 \rfloor$ negative signs at each place. On Y_j the doubled new blocks have coordinates $t-1-2w_l(\tau)-2b'$ ($0 \le b' < t$, $1 \le l \le j$); all doubled coordinates, including the base coordinates, are reduced by one if j is odd. The starting Y_0 is indefinite and the endpoint Y_m definite. The form of the row spaces and their adelic frames is part of the construction: the code gives `Data.exists_spaceY` in `New/Trace/GoodIdempotent.lean`, using the explicit glued Gram matrix and an integral frame. We regard the existence of this frame with its specified congruences as a formal-source input, and retain the flags and identifications of that construction (see Section 5).

The row form space before the good projection consists of automorphic functions of the indicated discrete-series type, right K_g-fixed. More precisely the source uses the subspace fixed by convolution with the reproducing kernel of the prescribed real type and demands rapid decay; its finite-dimensionality is the assertion `finite_levelSpace` for the cusp kernel. The fixed-level convolution representation of K_0/K_g at the split chosen places factors through

$$\mathcal{G} = \prod_{i=0,1} \prod_{w \mid \bar{\pi}_g,\ w \mid g_i} \mathrm{GL}_r(\mathbb{F}_{g_i}).$$

The good projector e_0 is a central idempotent of its group algebra, and $W = e_0 V$ with V the row form space at the principal level. The conditions and dimension estimates for e_0 are handled in the spanning chapter. We will write $C^{\mathrm{ram}}_{\mathrm{int}}$ for the forms in C^{ram} integral at every embedding and use $n = \dim_{\mathbb{C}} W$, $c = \dim_{\mathbb{C}} C$ throughout.

3.7 Integral gauges and the precise uniform index

For each rational prime p and field isomorphism $\iota : \mathbb{C} \xrightarrow{\sim} \overline{\mathbb{Q}}_p$, let $\mathscr{O}_\iota$ be the valuation ring pulled back by ι. If v is an algebraic form, the gauge of its value $v(g)$, with respect to the lattice at g, is

$$\|v(g)\|_{\iota,g} = \left\| \rho_{\kappa,e}\left((g^{-1})_{\iota\tau \circ \mathrm{conj}} \right)_{\tau \in \Phi} \iota(v(g)) \right\|_{\sup,p}.$$

The component matrix is obtained from the finite adele by extension to $\overline{\mathbb{Q}}_p$ via $\iota\tau \circ \mathrm{conj}$; coefficients of the Weyl matrix are integers, and the sup norm is taken in its integral coordinate lattice. This is `weylGauge` in `Lib/GLn/SchurModule/WeightRepBasic.lean`. Integral means every such gauge is at most one. Because the representation action respects the transported lattice, one may test on the finite class representatives. Notice that integrality is not defined merely by integrality of the complex coordinates at $g = 1$. This is significant when local Hecke modifications change the lattice or the branching map transports a different lattice at N.

For a family $y = (y_i)_{i=1}^q$ of row forms, the index inequality $\mathrm{IndexLe}(y, d_{\mathrm{ind}})$ uses real numbers e_p such that $\sum_{p \in S} e_p \le d_{\mathrm{ind}}$ for every finite set S of primes. For each p, ι and every allowed m-fold twist θ of conductor prime to p, there must be a $q \times q$ matrix U over $\mathscr{O}_\iota$ satisfying

$$|\iota(\det U)|_p \ge e^{-e_p}, \qquad \Theta_\theta\Big(\sum_j U_{ij} y_j \Big) \text{ integral at } \iota \quad (1 \le i \le q).$$

The matrix is not required to be the same for different twists, but the budgets e_p are fixed before them. This is precisely `Data.IndexLe` in `New/Endgame/ShortAndEnough.lean` (there the field isomorphisms are supplied as instances). These conventions explain the factor $2d_{\mathrm{ind}}$ in the squared-determinant estimate. There is a stronger local assertion on column lattices, before applying the projected Poincaré map: the corresponding submodule is independent of the twist. The short-form argument uses this stronger assertion to choose change-of-family matrices with a global index bound.

3.8 The theta chains and their relation

Let θ be a weight-zero automorphic character of $U(1)$, with tame conductor at a set of places disjoint from its complex conjugate and prime to gN. The successive powers through θ^m must retain the required ramification at that conductor (the trivial character is also allowed). A theta lift is a linear transformation defined by pairing an automorphic form on one group with a theta kernel associated to a Weil representation on a dual pair. Schematic notation for a lift is

$$(\mathcal{L}F)(h) = \int_{[U(V)]} F(u)\,\overline{\vartheta(u,h)}\,du,$$

where the precise conjugations, kernels, measures and characters depend on the convention. The source fixes these choices by its definitions of the spaces and lifts; this schematic integral is a description of the idea, not a replacement for those definitions. For definite groups, such integrals reduce to sums over finitely many classes.

Denote by Θ_θ the composite lift along the row chain beginning in W. The column chain begins in C, runs along the definite X_j, and afterwards restricts from X_m to Y_m. The ranks have been chosen so that

$$\dim_E X_m = k + m(t+1) = r + mt + 1 = \dim_E Y_m + 1.$$

The one-dimensional complement is used in the branching and level normalization. The m extra rank increments on the two sides make the iterated seesaw identity possible. Informally the relationship is

$$\begin{array}{ccccc} C^{\mathrm{ram}} & \longrightarrow & \text{definite chain} & \longrightarrow & X_m\text{-forms} \\ \downarrow P & & & & \downarrow \text{ restriction} \\ \text{initial row forms} & \longrightarrow & \text{row chain} & \longrightarrow & Y_m\text{-forms.} \end{array}$$

Here P is a Poincaré series constructed from a source form by Bessel induction. The actual identity involves a finite combination of definite chains, nonzero normalization factors and finite-place restrictions, not a literal equality between arbitrary individual theta kernels. The projected translates P_s in the upper bound, for $s \in \mathcal{G}$, are obtained from the Poincaré map by translating and projecting; we include the fixed normalizing factors from `projectedPoincare` in this notation.

The two rows of the diagram serve different purposes. The definite chain gives access to coefficient lattices and the sizes and prime divisibilities of derivatives of theta functions. The indefinite row chain gives a large automorphic space, a spanning theorem for Poincaré series, and a determinant calculation. Neither side alone yields the main inequality.

3.9 Norms, squared covolume and index

For each field automorphism σ of $\mathbb{C}$, the target definite space has a conjugated coefficient pairing, with squared norm denoted $\|\cdot\|_\sigma^2$. We use the pairing obtained by conjugating the coefficient data as specified in `Lib/Adelic/CuspForms/LevelForms.lean`, not the absolute value of an arbitrary conjugate of a real analytic integral. For independent vectors $y_1, \dots, y_q \in W$ define

$$\mathsf{G}_\sigma(y) = \big(\langle \Theta_1 y_i, \Theta_1 y_j\rangle_\sigma\big)_{i,j}, \qquad \mathrm{covol}_\sigma^2(y) = |\det \mathsf{G}_\sigma(y)|.$$

The determinant may be zero for some family before an injectivity statement is available; the upper estimate does not presuppose it is positive. We call V a logarithmic upper bound if

$\mathrm{covol}_\sigma^2(y) \le e^V$ for all σ. By the Gram determinant inequality,

$$\mathrm{covol}_\sigma^2(y) \le \prod_{i=1}^{q} \|\Theta_1 y_i\|_\sigma^2.$$

Thus bounds on individual lifts imply a bound on the determinant of their span. The source's definition uses the notation `covolumeSq`.

The paired finite-place structure and the index bound were fixed in Section 3.7.

3.10 The common leading term and the prime remainder

Here we introduce the quantities that occur in all three theorems. Let $\lambda_{d,1}(\nu) = \Lambda(\nu)(1-\chi_K(\nu))$ be the weight without a congruence restriction, and let s_ν be the place represented by $-\nu$ modulo N_1 when ν is invertible there (with the source's convention of the identity place otherwise). Define the finite prime sum at $\sigma \in \Phi$ by

$$A(y,\sigma) = \sum_{1\le\nu\le a} \lambda_{d,1}(\nu) \left(\frac{y}{\nu} - w_{\mathrm{nom}}(s_\nu\sigma)\right)_+ .$$

Write $\kappa_K = \gamma - \Re(L'(1,\chi_K)/L(1,\chi_K))$ for the Chowla–Selberg constant used in the proof, and

$$\Omega = \frac{2\pi}{\mathrm{rd}_E} e^{\kappa_K}, \qquad J(j) = \log(j!) - j\log(\Omega/g), \qquad \overline{w} = \frac{1}{f}\sum_{\sigma\in\Phi} \log w_{\mathrm{nom}}(\sigma).$$

The quantity $J(j)$ is the archimedean cost of a derivative of order j in the normalization of the proof. Its dependence on the period of the CM curve is the key to matching archimedean terms with the weighted prime sum. An expected main term for $A(a,\sigma)$, averaged appropriately, is

$$A_0 = a(\log a - 1 - \kappa_K - \overline{w}).$$

The exact quantities in the inequalities are

$$\begin{aligned}
\mathcal{R} &= \sum_{\sigma\in\mathcal{T}} (A(a,\sigma) - A_0),\\
\mathcal{U} &= \sum_{\sigma\in\Phi} (J(w_0(\sigma)) - A(w_{\mathrm{nom}}(\sigma),\sigma)),\\
H &= mk\mathcal{U}, \qquad B = fa(\log b)^3, \qquad L = B\log b = fa(\log b)^4.
\end{aligned}$$

These definitions are `Data.primeSum`, `Data.jetSize`, `Data.dirCost`, `Data.mainHeight`, `Data.mainAsk`, and `Data.remainder`. The quantity $\mathcal{U}$ represents the net cost of an ordinary differentiation direction, obtained by subtracting the finite-place divisibility gain from the archimedean derivative size. Both the short-form argument and the determinant argument are calibrated to this quantity. It is not necessary to determine its precise sign or asymptotic value.

The prime remainder has the following exact interpretation, which will be important in the final deduction:

$$\mathcal{R} = a\{|\mathcal{S}|(R_{N_0}(x) - R_{N_0}(x_1)) + |\mathcal{T}|R_1(x_1)\}. \tag{23}$$

This is `Data.remainder_eq`. We explain its mechanism in Section 10: the two weights replace the finite prime sums by R_{N_0} at two scales and a remainder R_1 without a congruence condition. The role of the analytic choice of x_1 is to control the unwanted second-cutoff expression. Notice that the normalized error $R_M(z)$ of the preceding section is not the same as the unnormalized, large remainder $\mathcal{R}$.

4 Local CM theta analysis and Schur jets

The contribution of this section is the local lower bound for valuations of the matrix coefficients of a theta kernel. We isolate it from the later transfer of Hecke operators: it is a result in commutative algebra about differential operators, after a particular CM theta function has been put into normal form. The numerical gain which it produces is the weighted prime-power sum. The subsequent section explains how to use the local estimates for actual lifts and repair the obstruction at primes for which two long rows collide.

4.1 The arithmetic local scales

Write v_p for the valuation normalized by $v_p(p) = 1$, and $|z|_p = p^{-v_p(z)}$. At a specified embedding ι into the algebraically closed p-adic field, put $\mathfrak{o} = \{z : |z|_p \le 1\}$ and $\mathfrak{m} = \{z : |z|_p < 1\}$. Let

$$q_p = \begin{cases} p^2, & \chi_K(p) = -1, \\ p, & \chi_K(p) \neq -1, \end{cases} \qquad \rho_p = \frac{1 + \chi_K(p)^2}{2(q_p - 1)}, \qquad \omega_p = \frac{1 - \chi_K(p)}{2}.$$

The values of ρ_p at split, inert and ramified primes are $1/(p-1)$, $1/(p^2-1)$ and $1/(2(p-1))$, respectively. These are the definitions of `nonsplitWeight` in `Lib/General/NumberTheory/QuadFieldSplitting.lean`. The congruence conditions on d ensure that a nonsplit prime is at least 13 (`Parameters.thirteen_le_of_chiK_ne_one`, in `New/Data/Different.lean`). We call an exponent e live at a nonsplit prime if $e \ge 1$ and either e is odd or $p = d$. Put $\mathfrak{p}_e(\sigma) = [-p^e]\sigma$, using the convention of the arithmetic data for the action on places. The normal-form statement retains also $e = 0$ at $p = d$; that term does not appear in the prime sum. Define

$$A_p(y, \sigma) = \omega_p \sum_{e \text{ live}} \left(y/p^e - w_{\mathrm{nom}}(\mathfrak{p}_e(\sigma))\right)_+, \qquad 0 \le y \le a \tag{24}$$

and set $A_p = 0$ at split primes. This is `Data.localAsk` in `New/Curve/LiveSum.lean`. The factor ω_p accounts for the ramified prime as well as inert primes. Since $w_{\mathrm{nom}} \ge a/x_1$, the sum is finite.

Lemma 4.1. *For every place σ and $0 \le y \le a$,*

$$A(y, \sigma) = 2 \sum_p A_p(y, \sigma) \log p, \qquad \omega_p \sum_{e \text{ live}} p^{-e} = \frac{1}{p-1} - \rho_p \quad (p \textit{ nonsplit}).$$

Proof. Expand the von Mangoldt function by prime powers. The factor $1 - \chi_K(p^e)$ is 2 at odd powers of an inert prime and vanishes at even powers; it is 1 at all positive powers of the ramified prime, and vanishes at split primes. For the second assertion sum the odd geometric series $p/(p^2-1)$ in the inert case and one half of the full geometric series $1/(p-1)$ in the ramified case. □

The latter identity is `hasSum_live_inv_pow` in `New/Curve/LiveSum.lean`. Introduce

$$F(p) = \sum_\sigma A_p(w_{\mathrm{nom}}(\sigma), \sigma), \qquad g_p = m \left(\sum_{\tau \in \mathcal{T}} A_p(a, \tau) + kF(p) \right). \tag{25}$$

Then

$$2 \sum_p g_p \log p = m \sum_{\tau \in \mathcal{T}} A(a, \tau) + mk \sum_\sigma A(w_{\mathrm{nom}}(\sigma), \sigma).$$

The integer m appears because there are m kernels, each with a distinguished derivatives at the tops and approximately $kw_{\mathrm{nom}}(\sigma)$ other derivatives at place σ. We shall derive the positive parts in (24).

4.2 The CM period and the differential coordinate

The Weil representation on a standard positive definite hermitian space uses a finite lattice normalized by an idele α; in the source the rational vector w has finite image $\alpha^{-1}w$. Write $\Lambda = \Omega_0 \mathcal{O}_K$ for the chosen elliptic curve lattice, $\Omega_0 \neq 0$, and ϑ for its normalized theta section, with $\vartheta(0) = 1$. Invariant derivatives are derivatives in the complex linear covering coordinate, not in an integral formal group coordinate. To read their Taylor series in the p-adic field we transfer the coefficients by ι; we use no analytic continuity of that field map.

Set $s_{\mathrm{ar}} = (2\pi/\sqrt{d})^{1/2}$. The source introduces the factors s_0 (`vacuumRoot`), $s_1 = s_0/u_0$ (`periodFactor`), and

$$\mathrm{per} = s_1 s_{\mathrm{ar}}/\Omega_0.$$

Here u_0 is the nonzero normalizing constant of the Chowla–Selberg identity and s_0 is a chosen f-th root of the reciprocal of the finite splitting character evaluated at α. The precise definitions are in `New/Curve/PeriodDefs.lean`. The positive constant in the period identity is

$$\Omega = \frac{2\pi}{\mathrm{rd}_E} \exp(\kappa_K), \qquad \kappa_K = \gamma - \Re \frac{L'}{L}(1, \chi_K).$$

Proposition 4.2 (Formal-source input: the CM normalizations)**.** *For every automorphism σ of $\mathbb{C}$ over K, $|\sigma(s_0)|^2 = \mathrm{rd}_E/\sqrt{d}$. The factor u_0 is the one prescribed by the normalized period formula `CMLattice.chowla_selberg`. For every $\tau \in \Phi$ and p-adic embedding, the norm of $\iota(s_1)$ is the reciprocal of the norm of the component of α read through the embedding $w \mapsto \iota(\tau(\overline{w}))$. Consequently, if $w \in E \cap \alpha\widehat{\mathcal{O}}_E$, then*

$$|\iota(s_1 \overline{\tau(w)})|_p \leq 1, \qquad |\iota(s_1)|_p = 1 \quad \text{if } p \nmid N_1.$$

The exact first identity is `Parameters.norm_gal_vacuumRoot` in `New/Curve/PeriodDefs.lean`; the Chowla–Selberg formula is `CMLattice.chowla_selberg` in `Lib/CM/FormalGroup/FormalHeight.lean`. The finite norm identity and its consequences are `norm_iota_periodFactor`, `norm_periodFactor_mul_conj_le_one` and `norm_iota_periodFactor_eq_one` in `New/Curve/TightFrameExists.lean`. They are used as normalized CM inputs, not as independently verified references.

4.3 Integral and logarithmic theta identities

Fix p and ι; all Taylor series of complex functions in the next paragraphs are read through ι. Write $\phi(z)$ for the integral parameter `translatedWpDeriv`, normalized to start with z, and $\xi(z)$ for `translatedWp`, which begins in degree two. Put $\mathfrak{R}_\phi = \mathfrak{o}[[\phi]] \subset \mathbb{C}_p[[z]]$. This denotes the image of the integral series ring by substitution, not the ring of integral coefficients in the differential coordinate.

Proposition 4.3 (Formal-source input: normalized CM identities)**.** *The ring $\mathfrak{R}_\phi$ contains ξ and is stable under d/dz. Moreover $\vartheta = \exp(-\mathscr{L}(z))$ with $\mathscr{L}(0) = \mathscr{L}'(0) = 0$ and*

$$\mathscr{L}''(z) = L_2 + \xi(z), \qquad |\iota(\sqrt{-d})L_2|_p \leq 1. \tag{26}$$

If (c, D) is in the closure of the pairs $(\iota\delta, \iota(\delta\bar{\delta}))$ for $\delta \in \mathcal{O}_K$, then

$$\phi(cz), \quad \exp(D\mathscr{L}(z) - \mathscr{L}(cz)), \quad (D\mathscr{L}(z) - \mathscr{L}(cz))' \ \in \mathfrak{R}_\phi. \tag{27}$$

If $0 \neq \delta \in \mathcal{O}_K$, $\delta y \in \Lambda$ and $D = \delta\bar{\delta}$, the twisted theta translate satisfies

$$(U_y\vartheta)(\iota(\bar{\delta})z) = \vartheta(z)^D V(z), \qquad V \in \mathfrak{R}_\phi. \tag{28}$$

These statements are `CMLattice.exists_log_theta` in `Lib/CM/KroneckerLimit/LogTheta.lean`, `taylorVia_translatedWp_mem_phiSeriesRing` and `deriv_mem_phiSeriesRing` in `Lib/CM/LocalHeight/IntegralParameter.lean`, and `rescale_translate_eq_theta_pow_mul` in `Lib/CM/LocalHeight/IsogenyIntegrality.lean`. The logarithmic identity follows from differentiation of the theta addition formula; the normalized integrality assertions use the integral model and integrality of torsion evaluations. Those underlying results are formal-source inputs in the present exposition.

Lemma 4.4 (Crude integrality of jets)**.** *Let $F = \prod_i (U_{y_i}\vartheta)(\sum_j c_{ij} z_j)$ for $\delta y_i \in \Lambda$, $0 \neq \delta \in \mathcal{O}_K$. If $|c_{ij}/(\iota(\bar{\delta})d)|_p \leq 1$, then for every polynomial Q with integral coefficients $(Q(\partial)F)(0)$ is integral.*

Proof. Let $D = \delta\bar{\delta}$. By (28), each factor comes from $\vartheta^D V$ with $V \in \mathfrak{R}_\phi$ by substituting $z = \sum_j (c_{ij}/\iota(\bar{\delta}))z_j$. Let $\mathfrak{R}$ be the $\mathfrak{o}$-algebra generated by the images under these substitutions of $\mathfrak{R}_\phi$, $\mathscr{L}'$, and ϑ for each i. All constants are integral. By the chain rule, a partial derivative of a generator is an integral multiplier times the substituted series obtained by applying d times the derivative in z. For $\mathscr{L}'$ the latter is $d(L_2 + \xi)$, in which dL_2 is integral; for ϑ use $\vartheta' = -\mathscr{L}'\vartheta$. The derivative stability of $\mathfrak{R}_\phi$ treats the remaining generators. Thus $\mathfrak{R}$ is stable under all partial derivatives. It contains F and hence all its iterated derivatives. Evaluating at zero and taking integral linear combinations proves the assertion. □

This is `CMLattice.jet_theta_prod_integral` of `New/Curve/FormalParameter.lean`. It applies also at split primes.

4.4 The regularized CM parameter at nonsplit primes

Suppose p is nonsplit. Write $q_p = p^\epsilon$, with $\epsilon = 2$ inert, $\epsilon = 1$ ramified, and $Q_0 = q_p - 1$. Thus $\omega_p = \epsilon/2$, $\rho_p = \omega_p/Q_0$. Choose a primitive Q_0-th root ζ in the p-adic field. The assertion that $(\zeta^j, \zeta^{j(p+1)})$ belongs to the norm-pair closure in Proposition 4.3 is the formal-source input `mem_normPairClosure_of_pow_eq` in `Lib/CM/KroneckerLimit/LogTheta.lean`. Average the parameter:

$$T(z) = Q_0^{-1} \sum_{j=0}^{Q_0-1} \zeta^{-j} \phi(\zeta^j z).$$

By (27) it lies in $\mathfrak{R}_\phi$, is tangent to the identity, and has support in degrees $1 \pmod{Q_0}$. It is an integral generator of the ring, since as a series in ϕ it has linear coefficient one and its inverse in that coordinate is integral by recursion. Let Z denote its inverse in the differential coordinate and write $Z(u) = u + z_{q_p} u^{q_p} + \cdots$. Uniqueness of compositional inverses implies $Z(\zeta u) = \zeta Z(u)$. For every $F \in \mathfrak{R}_\phi$ the series $F(Z(u))$ is integral. Differentiation of $T(Z(u)) = u$ therefore yields

$$Z'(u) \in \mathfrak{o}[[u]], \tag{29}$$

since $T' \in \mathfrak{R}_\phi$ has constant term one.

Take $c = p$ in the inert case and $c = \sqrt{-d}$ in the ramified case. Its norm-square is q_p, its p-adic valuation is ω_p, and the isogeny has degree q_p. The integral series $[c](u) = T(\iota(c)Z(u))$ is equivariant for the above roots of unity and begins with the nonunit $\iota(c)u$.

Lemma 4.5 (Formal-source input: the degree of the reduced isogeny)**.** *Under these hypotheses $[c](u)$ has a unit coefficient in degree at most q_p.*

The source is `CMLattice.exists_unit_coeff_degree_le` in `Lib/CM/FormalGroup/FormalHeight.lean`, the odd-degree CM isogeny height input for such an integral coordinate.

Since all intermediate coefficients vanish, the coefficient in degree q_p is a unit. Comparing through that degree in the composition gives $(\iota(c)-\iota(c)^{q_p})z_{q_p}$, and hence $|z_{q_p}|_p = p^{\omega_p}$. Choose ϖ with $\varpi^{Q_0} = -z_{q_p}^{-1}$ and put $\pi = \varpi^{Q_0}$; then

$$|\varpi|_p = p^{-\rho_p}, \qquad |\pi|_p = p^{-\omega_p}. \tag{30}$$

Lemma 4.6. *There is an integral series R for which*

$$\varpi^{-1} Z(\varpi u) = u - u^{q_p} + \pi R(u).$$

Its compositional inverse S is integral tangent to the identity. In particular for an integral series E_0,

$$T(\varpi z) = \varpi S(z), \qquad S - S^{q_p} = z + \pi E_0(z). \tag{31}$$

Proof. The first two nonzero coefficients are as stated by the choice of ϖ. In degree $n = 1+Q_0 j$, $j \geq 2$, (29) gives $v_p(z_n) \geq -v_p(n)$. After scaling the valuation of this coefficient is at least $j\omega_p - v_p(n)$. If $p \nmid n$ this is at least ω_p; otherwise $j \equiv 1 \pmod p$ and $j \geq p+1$. Then $n \leq p^{\epsilon} j$, and $v_p(n) \leq \epsilon + \log_p j \leq \epsilon(j-1)/2$, the last inequality holding for $p \geq 13$, $j \geq p+1$ by comparison at $j = p+1$ and monotonicity of the difference. This proves integrality of the remainder after division by π. Invert the integral series recursively and compose with this inverse, taking $E_0 = -R \circ S$. □

Thus the invariant differential is regularized at scale $p^{-\rho_p}$. The gap between $1/(p-1)$ and ρ_p is the source of the local gain.

4.5 The exceptional exponential factor

Let $c_2 = [z^2]\mathscr{L}$ and put

$$M(u) = (\mathscr{L}(z) - c_2 z^2)\big|_{z=Z(u)}, \qquad n(u) = (\mathscr{L}'(z) - 2c_2 z)\big|_{z=Z(u)}.$$

Then $M = O(u^3)$, $n(0) = 0$, and

$$M' = nZ', \qquad n' = (\xi \circ Z)Z' \in \mathfrak{o}[[u]], \qquad Z' \in \mathfrak{o}[[u]].$$

For a series G, let $G_{[j]}$ denote the part in degrees $j \pmod{Q_0}$. Since Z is equivariant for the Q_0-th roots of unity, this projection commutes with substitution of Z. Define $\kappa = \varpi^{p+1}/p$ and $\mu(u) = \kappa^{-1} M_{[p+1]}(\varpi u)$.

Lemma 4.7. *All coefficients of μ are integral. Modulo $\pi\mathfrak{o}[[u]]$, only its terms of degrees $p+1$ and $p+q_p$ may be nonzero.*

Proof. Since M vanishes through degree two, the possible degrees are $N = p+1+Q_0 I$, $I \geq 0$ (when $\epsilon = 1$ the degree-2 term of the congruence class vanishes). The coefficient of M in degree N is

$$\frac{1}{N} \sum_j n_{p+Q_0(I-j)} (1+Q_0 j) z_{1+Q_0 j},$$

where $j \geq 0$ and the subscript $i = p + Q_0(I-j)$ of n_i is positive. Thus $0 \leq j \leq I+1$ when ramified, and $0 \leq j \leq I$ when inert. By integrality of n', Z' a summand has valuation at least $-v_p(i)$, with an additional ω_p when $j = 1$ because $v_p(q_p z_{q_p}) = \epsilon - \omega_p = \omega_p$. Scaling by $\kappa^{-1}\varpi^N = p\pi^I$ makes its valuation at least

$$1 + \omega_p I - v_p(N) - v_p(i) + \omega_p \mathbf{1}_{j=1}. \tag{32}$$

For $I = 0$, $v_p(N) = 0$ and $v_p(i) \le 1$, so the expression is nonnegative. For $I = 1$ one has $N = p + q_p$ and $v_p(N) = 1$. At $j = 1$, $i = p$, and (32) is $2\omega_p - 1 \ge 0$; otherwise $p \nmid i$. If $2 \le I < p$, $N \not\equiv 0 \pmod p$ and the only divisible i is $i = p$ at $j = I$, giving valuation at least ω_p. For $I \ge p$, both positive integers N, i are at most $p^\epsilon(I+2)$. It follows that

$$v_p(N) + v_p(i) \le 2\epsilon + 2\log_p(I+2) \le 1 + \omega_p(I-1).$$

The last inequality holds at $I = p \ge 13$ since $\log_p(p+2) < 2$, and persists as I grows because the right hand side increases faster than the left. Thus (32) is at least ω_p whenever $I \ge 2$. The non-Archimedean triangle inequality finishes the proof. □

This reconstructs the valuation calculation of `norm_coeff_le_of_deriv` in `New/Curve/FormalParameter.lean`. The estimate works because the coefficients which can cause two divisions by p simultaneously fall into the first two selected degrees; in the second selected degree the exceptional coefficient of Z' is small enough.

Define $\Psi_0(z) = \kappa^{-1}\mathscr{L}_{[p+1]}(\varpi z)$. Using (31), the contribution of the non-quadratic part is $\mu(S(z))$. The quadratic term is retained only when $p = d$, in which case its coefficient is $u_0 = \kappa^{-1}\varpi^2 c_2$. From (26), $v_p(c_2) \ge -1/2$ (here 2 is a unit); hence $v_p(u_0) \ge 1 - \omega_p - 1/2 = 0$. Let $u_0 = 0$ in the inert case. We obtain

$$\Psi_0(z) = u_0 z^2 + u_1 S(z)^{p+1} + u_2 S(z)^{p+q_p} + \pi R_0(z), \tag{33}$$

with integral coefficients and $\Psi_0(0) = 0$. Moreover $|\kappa|_p = p^{\omega_p - 1/(p-1)}$, as is checked in the two cases by substituting ρ_p in $v_p(\kappa) = (p+1)\rho_p - 1$.

The formal source additionally proves that u_1 is a unit at inert primes. It uses the following formal-group input: if every $F \in \mathfrak{R}_\phi$ has $|[z^{p-1}]F'|_p < 1$, then $|[z^{p-1}]\xi|_p \ge 1$. This is `CMLattice.one_le_norm_coeff_translatedWp_nonsplit` of `Lib/CM/FormalGroup/FormalHeight.lean`. The required hypothesis follows from $T(z) = z$ through degree $p < q_p$: an integral series in T has integral coefficient in degree p, so its derivative coefficient in degree $p-1$ is divisible by p. Equation (26) now gives $|[z^{p+1}]\mathscr{L}|_p \ge p$; multiplication by $\kappa^{-1}\varpi^{p+1} = p$ makes the corresponding coefficient of Ψ_0 of norm at least one. The coefficient estimate gives the reverse inequality. This proves the unit assertion conditional on that formal-height input.

Proposition 4.8 (Local normal form for torsion translates)**.** *If $p \nmid n$ and $ny \in \Lambda$, there are a unit N and a constant κ_n in the p-adic field satisfying $|\kappa_n|_p = p^{\omega_p - 1/(p-1)}$ and an integral series V_y such that*

$$(U_y\vartheta)(\varpi z) = \big(\exp(\kappa_n \Psi_0(z))V_y(z)\big)\big|_{z \mapsto Nz}.$$

The same Ψ_0, N, κ_n work for all y with the given n.

Proof. By (27) for the norm pairs of the powers of ζ, the series $\exp(\zeta^{j(p+1)}\mathscr{L}(z) - \mathscr{L}(\zeta^j z))$ is integral in T. Replace z by $\zeta^{-j}z$ (which preserves the ring) and multiply these series over all j. Character orthogonality makes their product $\exp(Q_0(\mathscr{L}_{[p+1]} - \mathscr{L}))$. Apply (28) with $\delta = nQ_0$ to get

$$(U_y\vartheta)(nQ_0 z) = \exp(-(nQ_0)^2\mathscr{L}(z))V(z), \qquad V \in \mathfrak{R}_\phi.$$

Replace $\mathscr{L}$ by $\mathscr{L}_{[p+1]}$ in the exponential: the residual product just obtained, raised to the integer power n^2Q_0, is integral in T. Scaling by ϖ then gives integral coefficients for this product and V, since $T(\varpi z) = \varpi S(z)$. Finally take $N = (nQ_0)^{-1}$, $\kappa_n = -(nQ_0)^2\kappa$. This proves the required normal form. □

Equation (31), (33) and the proposition above give the nonsplit part of `CMLattice.exists_formal_parameter` in `New/Curve/FormalParameter.lean`.

4.6 Tight frames and theta vectors

The theta kernel on a space of matrices is not an unrelated product of one-variable theta functions. The product is built from a tight frame in the CM lattice.

Definition 4.9. A tight frame of length M is a finite collection of vectors $w_i \in E \cap \alpha\widehat{\mathcal{O}}_E$ with nonnegative integer multiplicities m_i such that

$$\sum_i m_i \sigma(w_i)\overline{\tau(w_i)} = M\delta_{\sigma\tau} \quad (\sigma, \tau \in \Phi).$$

After repetition of elements we may assume each multiplicity is one.

This is `IsTightFrame` in `Lib/Weil/Construction/ConjugateJetDefs.lean`. For instance the N_1 powers of ζ_{N_1} form a tight frame of length N_1: their integrality follows from the different normalization of α, and the frame equality is the geometric sum of $(\sigma(\zeta_{N_1})/\tau(\zeta_{N_1}))^j$. This is `Parameters.isTightFrame_zeta_pow`. To ensure good p-adic integrality of all theta vectors one needs lengths that are prime to p, and for the lower bounds also frames having some vector attaining all the finite norm bounds of Proposition 4.2.

Proposition 4.10 (Formal-source input: large frames with a unit vector)**.** *For every $R \geq 0$, every integer length $M \geq (dN_1)^6(R+1)$, and each p-adic embedding ι, there is a tight frame with multiplicities $m_i \geq R$ containing a vector w_j with $|\iota(s_1\overline{\tau(w_j)})|_p = 1$ for all τ.*

The precise arithmetic existence result is `Parameters.exists_isTightFrame`, in `New/Curve/TightFrameExists.lean`. It constructs additional frames using the cyclotomic different generator and solves for the desired length using the coprimality of two frame lengths. We have not independently reconstructed the general arithmetic frame construction. For upper bounds take $M = p(dN_1)^6 + 1$, large enough and prime to p.

For a standard positive definite hermitian space V of rank h, denote by $\mathscr{F}_V$ the finite part of its Weil representation with Heisenberg operator ϱ and theta functional ϑ_V. The Taylor model of $F \in \mathscr{F}_V$, at differential scale c, is the series $\mathcal{T}_c F$ with coefficients

$$[z^\beta]\mathcal{T}_c F = \frac{c^{|\beta|}}{\beta!}\vartheta_V(v_\beta \otimes F) \qquad (\beta \in \mathbb{N}^{\Phi\times\{1,\dots,h\}}), \tag{34}$$

where v_β is the Fock vector of the monomial and $\beta! = \prod_x \beta_x!$. This is `Space.jetSeries` in `Lib/Weil/Construction/DataProduct.lean`. Evaluating a monomial differential operator introduces $\beta!$ once again. No factorial may be lost by confusing the Taylor coefficient with the invariant derivative.

Let $\delta \in \mathcal{O}_K$ satisfy $\delta\bar{\delta} = D_\delta > 0$, and assume $v \in \Lambda$, with the curve's area $|v|^2 M\sqrt{d}/(2D_\delta)$. For a frame define linear maps on the complex coordinates of V by

$$A_{ic}(z) = \frac{v}{\bar{\delta}s_{\mathrm{ar}}}\sum_\tau \overline{\tau(w_i)}z_{\tau,c}.$$

The sum of pullbacks along A_{ic} of the normalized curve's hermitian form, counted with multiplicity m_i, equals the standard form: expand the scalar products of $A_{ic}(z)$ and $A_{ic}(z')$ and use Definition 4.9. The factor $|v/\bar{\delta}|^2 M$ is exactly absorbed by the area normalization. In particular the product of the pulled-back theta sections has the appropriate transformation law.

Proposition 4.11 (Formal-source input: theta product and finite vector)**.** *Under these frame and area choices, a product of the pullbacks of degree-m_i theta monomial sections is the entire*

theta function of some vector $F \in \mathscr{F}_V$. For rational Heisenberg translates this theta function is the product of the translated curve sections with arguments prescribed by the same linear forms. All finite point evaluations of F are integral at all p-adic embeddings. There is a finite vacuum expansion

$$\sum_{j\in I} M_\delta \ell(\varrho(\alpha x_j)\,\mathrm{vac})\varrho(-\alpha x_j)F = D_\delta^{2fh}\ell(F)\,\mathrm{vac}$$

for every complex linear functional ℓ, where M_δ is a natural number, $|I| = D_\delta^{2fh}$, and δx_j is integral at the finite places.

This is `Parameters.exists_finitePart_of_isTightFrame`, in `New/Curve/FrameSections.lean`; its vacuum expansion is the structure `Space.VacuumExpansion` of `Lib/Weil/Construction/ConjugateJetDefs.lean`. The finite vector reconstruction and its integrality are formal-source inputs here. The identity of the corresponding Taylor series follows by evaluating along every line and differentiating at zero (homogeneous polynomials are determined by their values on all lines). The corresponding lemmas are `Space.taylor_line_eq_subst_jetSeries` and `Space.jetSeries_eq_prod`. The source also proves `Space.eq_zero_of_jetSeries_eq_zero` by the identity theorem for entire functions.

4.7 Frobenius pairings in tight-frame products

Fix a nonsplit p. Consider a frame product of type `Parameters.thetaProducts` with all multiplicities repeated, frame length M prime to p, torsion denominators n prime to p, and multiplier b' of p-adic norm at most one (not the scale b). Its factors for pairs (i,c) are

$$(U_{y_{ic}}\vartheta)\left(\sum_\tau \iota(b'\overline{\tau(w_i)})z_{\tau,c}\right), \qquad ny_{ic} \in \Lambda.$$

At first there is no extra denominator from K. By Proposition 4.8, after substituting ϖz the product has the form

$$F(\varpi z) = \exp(\kappa_n \Psi(z))V(z), \quad V \in \mathfrak{o}[[z]], \quad \Psi(0) = 0, \tag{35}$$

where $\Psi = \sum_{i,c}\Psi_0(L_{ic})$, $L_{ic} = \sum_\tau r_{\tau i}z_{\tau,c}$, and $r_{\tau i} = N\iota(b'\overline{\tau(w_i)})$. At a nonsplit prime, $p \nmid N_1$, so Proposition 4.2 gives $r_{\tau i} \in \mathfrak{o}$.

Lemma 4.12 (Formal-source input: Frobenius orthogonality)**.** *In characteristic p, modulo $\pi\mathfrak{o}$, the frame matrix satisfies, for any live exponent e and for $e = 0$ when $p = d$,*

$$\sum_i r_{\sigma i}^{p^e} r_{\tau i} = (N\iota(b'))^{p^e+1} M\,\mathbf{1}_{\tau=\mathfrak{p}_e(\sigma)}. \tag{36}$$

Here the multiplier is read in the same reduction, with N regarded as the p-adic element from the normal form.

The formula is `Parameters.exists_frame_rows_frob` in `New/Curve/ThetaProducts.lean`; more precisely it holds under every map to characteristic p killing all elements of valuation at least ω_p. The proof uses the Frobenius permutation of the cyclotomic embeddings and the frame identity. At an inert prime the odd Frobenius interchanges the CM embeddings, explaining the minus sign in $[-p^e]$; at the ramified prime complex conjugation acts the same way modulo the local uniformizer. These arithmetic identifications of embeddings modulo the indicated ideals are used as input.

For a finite truncation degree $D > 0$ let $I_D = \{0 \le e \le \lfloor \log_p D \rfloor : e \text{ odd or } p = d\}$. For integral series $Z_{\tau,c}$ with zero constant term in any finite variable set, suppose the pairings

$$\sum_{\tau,c} Z_{\tau,c}^{p^e} Z_{\mathfrak{p}_e(\tau),c} = 0 \quad (e \in I_D) \tag{37}$$

hold after reduction modulo π and terms of degree $> D$. We shall prove they kill the exceptional exponent in (35) and render the exponential integral through degree D.

Lemma 4.13 (Artin–Schreier cancellation)**.** *Let R be a commutative ring of characteristic p, and let finitely many elements t_i, L_i satisfy $t_i - t_i^{p^\epsilon} = L_i$. Assume each t_i is nilpotent, and $\sum_i L_i^{p^j+1} = 0$ for all odd j if $\epsilon = 2$, or all $j \ge 0$ if $\epsilon = 1$. Then $\sum_i t_i^{p+1} = \sum_i t_i^{p+p^\epsilon} = 0$.*

Proof. Iterate the given equation:

$$t_i = \sum_{a=0}^{A-1} L_i^{p^{\epsilon a}} + t_i^{p^{\epsilon A}}.$$

Nilpotence kills the last term for large A. Expanding $\sum_i t_i^p t_i$ by Frobenius gives terms $\sum_i L_i^{p^u} L_i^{p^v}$ for $u = \epsilon a + 1$, $v = \epsilon b$. Each equals $(\sum_i L_i^{p^{|u-v|}+1})^{p^{\min(u,v)}} = 0$; indeed the exponent difference is odd if $\epsilon = 2$. Expand $\sum_i t_i^p t_i^{p^\epsilon}$ in the same way, taking $v = \epsilon(b+1)$. This proves both identities without requiring that R be reduced or perfect. □

This is `sum_pow_eq_zero_of_null` in `New/Curve/ThetaProducts.lean`.

Proposition 4.14 (Null-cone integrality)**.** *Whenever the integral zero-constant series $Z_{\tau,c}$ satisfy (37), the coefficients of $F(\varpi Z)$ through degree D are integral.*

Proof. Work in the quotient of the ring of integral series by the ideal of coefficients divisible by π together with terms of degree $> D$. For $t_{ic} = S(L_{ic}(Z))$, (31) gives $t_{ic} - t_{ic}^{p^\epsilon} = L_{ic}(Z)$, and t_{ic} is nilpotent by degree. Using (36) we have

$$\sum_{i,c} L_{ic}(Z)^{p^e+1} = (N\iota(b'))^{p^e+1} M \sum_{\tau,c} Z_{\tau,c}^{p^e} Z_{\mathfrak{p}_e(\tau),c} = 0$$

for $e \in I_D$. Larger e give zero automatically by degree. Lemma 4.13 shows that the two S-powers in (33) vanish after summing; at $p = d$ the quadratic terms vanish by the same computation with $e = 0$. It follows that the coefficients of $\Psi(Z)$ through degree D are divisible by π. The exponential in (35) is integral through this degree because

$$v_p((\kappa_n \pi)^j / j!) = j/(p-1) - v_p(j!) \ge 0.$$

There is no formal convergence problem, since the substituted series have zero constant term and we work in a finite truncation. Multiplication by $V(Z)$ preserves integrality. □

This is the first conclusion of `thetaProducts_normal_form` in `New/Curve/ThetaProducts.lean`, whose hypothesis is the predicate `IsIntegralOnNullCone`. In the source the couplings are required to vanish under every characteristic-p map annihilating π; it suffices in the proof to apply that assumption to the indicated quotient.

In the inert case, when b' has norm one, one also knows which coefficients of the exponent are units. Reducing in the residue-field series ring and differentiating (33), $u_0 = 0$ and the term with u_2 has zero derivative. Since $S - S^{p^2} = z$, $S' = 1$ in this reduction, and

$$\frac{\partial \overline{\Psi}}{\partial z_{\tau,c}} = \bar{u}_1 \sum_i r_{\tau i}\, t_{ic}^p, \qquad t_{ic} = \sum_{a \ge 0} L_{ic}^{p^{2a}}.$$

By the frame identity a coefficient is nonzero exactly at $z_{\sigma,c}^{p^e}$ with odd e and $\mathfrak{p}_e(\sigma) = \tau$. It is a unit because M, Nb', u_1 are units. No coefficient supported on several variables occurs. This proves `coeff_sum_frob_ne_zero_iff` and explains the sharpness assertion of `thetaProducts_normal_form`. The same pairings which reduce the guaranteed upper divisibility are genuinely present in the series. This fact is also used in the lower volume estimates.

4.8 Schur matrices and determinant losses

Let R_1, R_2 be the source and target ranks, $N = R_2$, and let $\kappa_\sigma(h)$ count columns of height $h < N$ in the prescribed tableau. A filling assigns h row indices to each column of height h; its polynomial $b_i(Z)$ is the product of minors in the indicated rows and first h columns. Put $D = \sum_\sigma \sum_{h<N} h\kappa_\sigma(h)$. The jet estimate must hold for every Schur matrix coefficient, not just for the highest weight vector.

Proposition 4.15 (Formal-source input: integral Schur expansion)**.** *For a weight satisfying the rank condition `FitsRanks` the Schur matrix $S_\kappa(X)$ satisfies*

$$b_{i'}(XY) = \sum_i S_{i'i}(X) b_i(Y) \prod_\sigma \det(Y_\sigma)^{d_\sigma}, \tag{38}$$

where the determinants use the first R_1 columns of the variable matrix Y_σ, and $d_\sigma = \kappa_\sigma(R_1)$ if $R_1 < N$, zero otherwise. This holds over every ring and commutes with base change; the coefficient coordinate maps in the standard tableau basis are integral on the associated Borel semi-invariants.

This is the standard integral tableau theorem as used by `PolyRep.sum_schur_smul_nablaBasis` and `schur_map` in `Lib/GLn/SchurModule/SchurMatrix.lean`. Its basis and weight are `nablaBasis` and `diagramWeight` in `Lib/General/Algebra/ZerosCostDegree.lean`; the integral basis theorem is in `Lib/GLn/SchurModule/MinorBasis.lean`. Write $J_Q(F) = Q(\partial)F(0)$ and $\mathscr{J}_\kappa(F)_{i'i} = J_{S_{i'i}}(F)$ (`schurJet` in `Lib/Weil/Construction/DataProduct.lean`).

Let ev be a finite set of places whose height-one couplings will be exempted. Define

$$a_\sigma^{ev} = \kappa_\sigma(1)\mathbf{1}_{\sigma\notin ev}, \quad c_\sigma = \sum_{\substack{h<N\\ h\neq 1}} \kappa_\sigma(h), \quad D_\sigma = \sum_{\substack{h<N\\ h\neq 1}} h\kappa_\sigma(h).$$

For the ordered coupling $\sigma, \tau = \mathfrak{p}_e(\sigma)$ put

$$\mathfrak{C}_{p,e}(\sigma,\tau) = \min(a_\sigma^{ev}/p^e, c_\tau + a_\tau^{ev}) + c_\sigma/p^e + \min(D_\sigma/p^e, D_\tau). \tag{39}$$

This is `couplingCount` in `New/Curve/LiveSum.lean`. Its three terms distinguish the three types of pairs of columns.

Lemma 4.16 (Simultaneous tilting)**.** *For any integral matrix arrays $Y_{\sigma,c,j}$ with source rows c and column indices $j \geq 0$, there are integral $h \times h$ matrices $M_{\sigma,h}$, $1 \leq h < N$, as follows. Write $Y_\sigma^{(h)}$ for the transposed array of the first h columns. For each $e \in I_D$, $\tau = \mathfrak{p}_e(\sigma)$, and rows u, u' from $M_{\sigma,h}Y_\sigma^{(h)}$, $M_{\tau,h'}Y_\tau^{(h')}$,*

$$\sum_c u_c^{p^e} u'_c \equiv 0 \pmod{\pi}$$

unless $h = 1$ and either $\sigma \in ev$ or $(h' = 1, \tau \in ev)$. Moreover

$$\prod_{\sigma,h} |\det M_{\sigma,h}|_p^{\kappa_\sigma(h)} \geq p^{-\omega_p \sum_{e\in I_D} \sum_\sigma \mathfrak{C}_{p,e}(\sigma,\mathfrak{p}_e(\sigma))}. \tag{40}$$

Proof. An elementary observation about valuation rings: if coefficient vectors in $\mathfrak{o}^h$ must satisfy finitely many conditions $L_i(v) \in u_i\mathfrak{o}$ with integral linear functionals L_i and nonzero $u_i \in \mathfrak{o}$, there is a matrix of h admissible rows of determinant valuation at most $\sum_i v_p(u_i)$. Inductively a unimodular change on the current coefficient basis makes the next linear functional depend on just one coordinate: choose among its coefficients one of largest norm and use divisibility to eliminate the others. Multiplying that coordinate by u_i suffices; earlier conditions are maintained by linearity. Requiring all h coordinates to be divisible by u_i instead costs $hv_p(u_i)$.

Fix e, σ, τ as in the statement. For source height one, the untransformed row y is fixed. If $\sigma \notin ev$, either scale it by an integral p^e-th root of π, of valuation ω_p/p^e, or impose the single congruence $\sum_c y_c^{p^e} u'_c \equiv 0 \pmod{\pi}$ on each target height except the exempt height-one row. The weighted costs are $\omega_p a_\sigma^{ev}/p^e$ and $\omega_p(c_\tau + a_\tau^{ev})$; choose the cheaper. If $\sigma \in ev$ there are no conditions of this type.

For source height greater than one and target height one, choose integral p^e-th roots of the entries of the untransformed target row y' and require

$$\sum_c (y'_c)^{1/p^e} u_c \in \pi^{1/p^e}\mathfrak{o}.$$

Taking p^e-th powers in characteristic p gives the required congruence modulo π. Further transformation of y' is scalar multiplication and preserves this relation. The weighted cost is $\omega_p c_\sigma/p^e$. Finally, for both heights greater than one, annihilate all pairings by scaling all coefficient rows on either the source or the target side, at the smaller cost $\omega_p \min(D_\sigma/p^e, D_\tau)$. All the conditions are on the original coefficient coordinates or the fixed linear forms specified above. Impose them simultaneously by the first observation; integral row operations preserve previously zero pairings, since the quotient has characteristic p. The costs add to the exponent of (40). □

This is the argument of `exists_null_coupling_frames` in `New/Curve/JetDivisibilit y.lean`. The source implements the valuation-ring observation with `IndexDvd.of_det_dvd`. In particular the matrices are not claimed unimodular; their determinant costs must be subtracted.

For each box with height h and row label a prescribe a nonnegative integer charge $g(\sigma, h, a)$. Require

$$g(\sigma, 1, a) > 0 \quad \text{or} \quad g(\mathfrak{p}_e(\sigma), h, a) > 0$$

whenever $\sigma \in ev$ or $(h = 1, \mathfrak{p}_e(\sigma) \in ev)$. For a tableau i' write $G_{i'}$ for the total charges of its boxes, with multiplicity. The charges will kill the exempt pairings.

Theorem 4.17 (Schur-jet bound)**.** *Suppose F satisfies the null-cone property of Proposition 4.14 for degree $D > 0$ and the scales (30). Then*

$$\begin{aligned} v_p(\mathscr{J}_\kappa(F)_{i'i}) \geq (1/(p-1) - \rho_p)D - \omega_p \sum_{e \in I_D} \sum_\sigma \mathfrak{C}_{p,e}(\sigma, \mathfrak{p}_e(\sigma)) \\ - fN^3(\lfloor \log_p D \rfloor + 1) - G_{i'}. \end{aligned} \tag{41}$$

Proof. Apply the jet functional in X to (38). The integral basis property converts any coefficient bound on the polynomial $J_{b_{i'}(XY)}(F)$ in Y to a bound on the entries of the Schur jet. The largest coefficient norm of a polynomial over an algebraically closed valued field is attained at some integral specialization: divide by a largest coefficient, reduce modulo the maximal ideal, and choose a point of the infinite residue field where the resulting nonzero polynomial does not vanish. Thus it suffices to bound the expression for all specialized integral Y.

Apply Lemma 4.16. Substitute in the theta series fresh variables $u_{\sigma,h,j,a}$, $0 \le j < h < N$, $1 \le a \le N$, via

$$X_{\sigma,a,c} = \varpi \sum_{h,j} p^{g(\sigma,h,a)} (M_{\sigma,h} Y_\sigma^{(h)})_{j,c} u_{\sigma,h,j,a}.$$

Remove the scaling factor ϖ from the substitution. Expanding the Frobenius pairings of matrix coordinates (a, c) in characteristic p gives exactly the pairings of tilted coefficient rows. Unexempted terms vanish, and exempt terms vanish by their positive charges since $p \in \pi\mathfrak{o}$. The null-cone condition gives integral Taylor coefficients of the substituted theta series through degree D.

Now apply the product of specified minors in the fresh variables as a differential operator. It is homogeneous of degree D, with integral coefficients. There are at most $fN \sum_{h<N} h \le fN^3$ fresh variables. A monomial u^β of degree D contributes a factor $\beta!$; the base-p digit formula gives

$$v_p(\beta!) = (D - \sum_i s_p(\beta_i))/(p-1) \ge D/(p-1) - fN^3(\lfloor \log_p D \rfloor + 1).$$

The same bound holds for the sum of the monomial pairings. The chain rule identifies that sum with

$$\varpi^D p^{G_{i'}} \left(\prod_{\sigma,h} \det M_{\sigma,h}^{\kappa_\sigma(h)} \right) J_{b_{i'}(XY)}(F).$$

Within each minor the substituted differential rows are transformed by the same $M_{\sigma,h}$, yielding its determinant, and each box contributes ϖ and the power of p prescribed by its charge. Removing these factors subtracts $\rho_p D$, $G_{i'}$, and at most the determinant loss (40). This proves (41). □

This is `norm_schurJet_le` in `New/Curve/JetDivisibility.lean`. The factorial gain is diminished by the coordinate and determinant costs, and the explicit logarithmic factorial correction. The minima in (39) are essential: replacing them by the more expensive alternatives would eliminate the positive-part terms in (24).

4.9 Application to the prescribed weights

For the arithmetic data $\mathcal{D}$, let $R_j = k + (j+1)(t+1)$, $0 \le j < m$, be the target ranks of the definite chain. Put $w_\sigma = w_{\mathrm{nom}}(\sigma)$,

$$a_\sigma = a\mathbf{1}_{\sigma\in\mathcal{T}}, \quad c_{j,\sigma} = \sum_{\substack{h<R_j \\ h\neq 1}} \kappa_\sigma(h), \quad D_{j,\sigma} = \sum_{\substack{h<R_j \\ h\neq 1}} h\kappa_\sigma(h), \quad D_j = \sum_\sigma (a_\sigma + D_{j,\sigma}).$$

The weight multiplicities $\kappa_\sigma(h)$ are `Data.`κ in `New/Curve/LiveSum.lean`: height one occurs at tops with multiplicity a; height k has multiplicity $w_1(\sigma) - \lfloor (k+t+1)/2 \rfloor$; heights above k congruent to k modulo $t+1$ have multiplicity $2h_2 - (t+1)$; all other multiplicities vanish. These expressions are nonnegative by admissibility.

Lemma 4.18 (Weight counts)**.** *For $0 \le j < m$ one has*

$$kw_\sigma \le D_{j,\sigma} \le kw_\sigma + x^{60}, \qquad 0 \le c_{j,\sigma} \le w_\sigma + x^{60}.$$

Proof. The multiplicity of height k is at least w_σ because the rounded weight w_1 includes the gap $3h_2$. The difference from w_σ is at most the rounding step plus $3h_2$. There are at most $j+1$ higher heights, each bounded by R_j and of multiplicity at most $2h_2$. Thus the excess box count is at most $k(N_1 + 3h_2) + 2(j+1)R_j h_2$ (here N_1 in the rounding term

denotes the cyclotomic order, not a target rank). With $k \ll x^5$, $m \ll x^{1/2}$, $R_j \ll x^{11/2}$, $h_2 = r^8 \ll x^{40}$ and the cyclotomic bound, this is smaller than x^{60} on the admissible asymptotic range. The column count is bounded by ignoring the heights. The formal source establishes these comparisons for every admissible datum as `Data.lowColumns_lowBoxes_bounds` in `New/Columns/JetExponentBound.lean`. □

For a nonsplit prime $p < x_1$ define the exponent

$$E_{p,j}(ev) = \left(\frac{1}{p-1} - \rho_p\right) D_j - \omega_p \sum_{e \in I_{D_j}} \sum_{\sigma} \mathfrak{C}_{p,e}(\sigma, \mathfrak{p}_e(\sigma)) - f R_j^3(\lfloor \log_p D_j \rfloor + 1). \tag{42}$$

This is `Data.jetExponent` in `New/Columns/JetExponentBound.lean`. Define

$$\mathrm{coll}_p(ev) = \sum_{\sigma} \sum_{e \text{ live}} (a/p^e) \mathbf{1}_{\sigma, \mathfrak{p}_e(\sigma) \in \mathcal{T} \setminus ev}.$$

The sum converges by geometric decay but need not have finite support. It records the loss when a long row has a long partner.

Proposition 4.19 (Prime-power gain of one jet)**.** *For nonsplit $p < x_1$ and $0 \le j < m$,*

$$E_{p,j}(ev) \ge \frac{g_p}{m} - \mathrm{coll}_p(ev) + \omega_p |\mathcal{T} \cap ev|\, a/x_1 - \frac{f a (\log b)^2}{pm}. \tag{43}$$

Proof. We give the comparison on the asymptotic admissible range, sufficient for the main theorem; the formal source proves it uniformly for all admissible data. Set $u = a/(pm)$. The separated scales imply $x^{60}, R_j^3 \le u$, $c_{j,\sigma}/p \le u$, and $5(\lfloor \log_p D_j \rfloor + 1) \le (\log b)^2$. For example $u \ge xb/(x_1 m)$ and $x_1 \le b^{10^{-6}}$ control all polynomial terms. The column bound and $m/x \to 0$ give $c_{j,\sigma}/p \le (a/x + x^{60})/p \le u$. The last inequality follows from $D_j \ll b^2$ for sufficiently large x using $p \ge 2$.

Fix a live exponent $e \in I_{D_j}$, write $P = p^e$, and let $\tau = \mathfrak{p}_e(\sigma)$. The elementary factorial contribution less the cost for that coupling is

$$\begin{aligned} &\frac{a_\sigma + D_{j,\sigma}}{P} - \mathfrak{C}_{p,e}(\sigma, \tau) \\ &= \frac{a_\sigma}{P} - \min(a_\sigma^{ev}/P, c_{j,\tau} + a_\tau^{ev}) \\ &\qquad + \left(\frac{D_{j,\sigma}}{P} - D_{j,\tau}\right)_+ - \frac{c_{j,\sigma}}{P}. \end{aligned}$$

The ordinary-box term is at least $k(w_\sigma/P - w_\tau)_+ - x^{60}$, by Lemma 4.18. The last term costs at most u. For the distinguished height-one term, if σ is not a top then the contribution and requested gain are zero. If both σ, τ are nonexempt tops, the loss relative to $(a/P - w_\tau)_+$ is at most a/P, the collision summand. If σ is a nonexempt top but τ is not a nonexempt top, then $a_\tau^{ev} = 0$ and $c_{j,\tau} \le w_\tau + x^{60}$, giving at least $(a/P - w_\tau)_+ - x^{60}$. If σ is an exempt top, the minimum is zero and a/P exceeds the requested positive part by $\min(a/P, w_\tau)$; at $e = 1$ this is at least a/x_1 since $p < x_1$ and $w_\tau \ge a/x_1$. These four cases give the pointwise desired bound with at most $3u$ of error per place and exponent, plus the exempt contribution for $e = 1$.

For live e above the truncation cutoff there is no coupling cost; the factorial contribution then bounds both requested positive parts. The extra contribution at $e = 1$ does occur within the truncation since $D_j \ge a > x_1$. In summing the factorial contribution over all live exponents use Lemma 4.1. The $e = 0$ term allowed at $p = d$ occurs only in the coupling

cost, and costs at most $4u$ per place. Indeed $p = d$ is fixed and $c_{j,\sigma} \le u$ from $w_\sigma \le a/x$ and $m/x \to 0$. The place $[-1]\sigma$ cannot have the same class of tops as σ, nor can both be thick. Hence the first minimum is bounded by u, and the last by $ka/x_1 + x^{60} \le u$, since at least one term corresponds to the small weight. Summing, the comparison error is at most $4fu(\lfloor \log_p D_j \rfloor + 1)$, and the factorial error at most $fu(\lfloor \log_p D_j \rfloor + 1)$. Multiply the comparison by $\omega_p \le 1$ and use the last scale inequality above. The sum of the desired positive parts with factor ω_p is exactly g_p/m by (25). □

The exact uniform source declaration is `Data.le_jetExponent`, in `New/Columns/JetExponentBound.lean`. Formula (43) is the numerical interface to the chain. When $ev = \varnothing$ there are no charges, but long-row collisions can be expensive. The Hecke repair uses exempt sets which remove all such collisions; the extra positive term in (43) compensates part of the charge incurred in doing so.

4.10 From theta products to the standard integral kernel data

A bound for arbitrary formal products becomes a bound for the automorphic theta kernel only after identifying the normalized finite integral structures. We give this interface explicitly. For a standard space V let vac_V be the standard finite vacuum. Fix a nonzero factor s_{vac} chosen for rank one and define the scaled vacuum of rank h to be $\mathrm{vac}_V^{\circ} = s_{\mathrm{vac}}^{h}\,\mathrm{vac}_V$. Its point evaluations are the products over the h coordinates of the rank-one evaluations. The integral module $\mathrm{Std}_\delta(V, \iota)$ is the span over $\iota^{-1}(\mathfrak{o})$ of $\varrho(\alpha x)\,\mathrm{vac}_V^{\circ}$ with δx integral at all finite places above p (x is a finite adelic vector). This definition is `Parameters.standardData` in `New/Curve/IntegralVacuum.lean`; it depends on $\delta \in \mathcal{O}_K$ and on the p-adic embedding.

Proposition 4.20 (Formal-source input: integral vacuum and conjugates)**.** *The scalar s_{vac} is nonzero, and for all p and all p-adic embeddings every finite point evaluation of vac_V° is integral. For $\sigma \in \mathrm{Aut}(\mathbb{C}/K)$ its conjugates are bounded by $\exp(CN_1^2 h)$ for a constant C depending only on the fixed K, N_0. Moreover there is a finite vector generating the finite part by Heisenberg translates and with theta functional equal to 1; after applying σ, its jets at scale* per *coincide, up to permutation of the places, with jets of translates of another such finite vector at a scale of squared absolute value $1/\Omega$.*

The precise uniform generator statement is `Parameters.exists_isJetConjugate_generator` in `New/Curve/ConjugateGenerator.lean`; integrality and the conjugate bound for the scaled vacuum are `norm_intVacuum_le_one` and `norm_intVacuum_le` in `New/Curve/IntegralVacuum.lean`. These inputs use the normalized CM reciprocity law. For higher rank the vacuum value bounds follow multiplicatively from rank one, as indicated by `intVacuum_finVec_eq_prod`.

The integral module is stable under the Heisenberg translations by the indicated lattice because the central factors in their multiplication law are roots of unity and hence units at p. It is also stable under the finite Weil action of an element integral above p. This follows from the finite vacuum-translation sum for the Weil action on a vacuum: it is a sum of translates by integral vectors with unit multipliers, and conjugation by the integral element preserves the translation lattice. This bounded-action assertion is used in its exact form `Parameters.omega_mem_standardData` and `finHeisOp_mem_standardData`, in `New/Curve/ConjugatesOfJets.lean`. We assume the finite Weil model's vacuum sum identity in this application.

Lemma 4.21 (Transfer of a jet bound)**.** *Let $\mathcal{L}$ be a complex linear functional on the Taylor series space (in particular any specified Schur jet entry), $c \in \mathbb{C}$ a scale multiplier, and Q a*

nonnegative real bound. Suppose the norm of $\iota(\mathcal{L}(F))$ is at most Q for every F whose image under ι is a frame theta product with the torsion condition $n\delta y_i \in \Lambda$ and with scalar norm $|\iota(cs_1)|_p$. Then for every $f \in \mathrm{Std}_\delta(V,\iota)$,

$$|\iota(\mathcal{L}(\mathcal{T}_{c\ \mathrm{per}} f))|_p \le Q.$$

Proof. Take a frame with $M = p(dN_1)^6 + 1$, of length prime to p, using Proposition 4.10. Repeat its vectors so that all multiplicities are one. Use Proposition 4.11 with $\delta = M$ and the corresponding repeated frame of length M^2, choosing the monomial sections equal to the normalized theta section. Then $v = \Omega_0$ and $D_\delta = M^2$ satisfy the required area identity. The resulting finite vector F_0 has theta value one, since every normalized theta section has constant value one. In the vacuum expansion take ℓ to be the theta functional and solve for the vacuum. After scaling by s_{vac}^h, the coefficients are integral: they are the normalized vacuum evaluations times the natural multiplier divided by M^{4fh}, a p-adic unit. The phase factors from translating these expansions have absolute value one.

Consequently a lattice translate of the normalized vacuum is an integral linear combination of translates of F_0. Approximate each finite translate by a rational one, using density of rational points of the finite additive adeles and local constancy of the Heisenberg orbit on the finite model. The rational approximation can be made so that δ times the translation vector is integral above p. Denominators at other primes are cleared by a rational integer prime to p. Applying the frame product identity gives the class of theta products in the hypothesis: the scalar factor is the product of cs_1 and factors whose norm is one from M; the torsion coordinates have denominators dividing $n\delta$ with $p \nmid n$. This exact rational approximation and denominator assertion is the one used in `Parameters.norm_jetSeries_le` in `New/Curve/ConjugatesOfJets.lean`. Linearity of $\mathcal{L}$ and the non-Archimedean inequality then bound every generator of the integral module, hence its entire span. □

This argument uses the stated finite-vector and period inputs rather than replacing the actual kernel data by a single formal theta product. For $\delta = 1$ at a nonsplit prime, Proposition 4.14, Theorem 4.17, and Lemma 4.21 bound the Schur jets on the entire module whenever the scalar multiplier has integral norm. For general $\delta \neq 0$, Lemma 4.4 applied using Proposition 4.2 gives the elementary integrality at scale $\overline{\delta}d$ per; this is precisely `Parameters.norm_jet_jetSeries_le_one` in `New/Curve/ConjugatesOfJets.lean`.

Finally, a finite theta kernel acts not just on one fixed Weyl coefficient vector but on vectors at all adelic arguments. The Weyl gauge at g of a vector ξ is the maximum absolute value of its coordinates after reading $g^{-1}\xi$ through the specified p-adic embeddings (where the action contains the determinant character as well as the Schur representation). This is `weylGauge` of `Lib/GLn/SchurModule/WeightRepBasic.lean`. An integral change of adelic frames above p has integral matrices in both directions under this representation, so the gauge is unchanged. For the standard data f the kernel is equivariant for these integral changes.

Proposition 4.22 (Formal-source input: gauge bound for a translated theta kernel)**.** *If the pair of adelic changes of the matrix variable factors into integral changes and a diagonal element with parameters $u_\sigma(a)$, then a uniform bound on every jet coefficient after weighting by*

$$\prod_\sigma \left(\prod_a |u_\sigma(a)|_p \right)^{e_\sigma} \prod_{\text{boxes } (\sigma,h,a) \text{ of } i'} |u_\sigma(a)|_p$$

is a bound on the factor by which the theta kernel raises the Weyl gauge. Here e_σ is the determinant exponent of the output weight. In particular the kernel with standard datum and

scale $c\,\mathrm{per}$*, with no diagonal change, raises gauges by at most*

$$\left(\frac{|\iota(c)|_p}{|\iota(\bar{\delta}d)|_p}\right)^D,$$

where D *is the number of boxes of the output tableau.*

The exact interface is `Parameters.mem_raises_of_jet_bound` and `mem_raises_of_standardData` in `New/Curve/ConjugatesOfJets.lean`. Its finite-place factorization and equivariance are used as inputs. The matrix estimate itself is just the ultrametric operator bound, and the weights come from evaluating the diagonal matrix in the Schur representation and its determinant character. The final displayed bound follows by homogeneity and the crude integrality at scale $\bar{\delta}d$ per proved above. More importantly, at a non-split prime away from the conductor and level, the sharper Schur bound replaces that crude estimate by (41), with any extra weights charged box by box. This completes the local differential calculation needed for the Hecke and seesaw argument.

5 The definite chain, collisions and Hecke repair

We continue the construction of the preceding section. Throughout this section the prime d of K is fixed. Write $g = g_0 g_1$ for the product of the two level primes, and let $\pi_g \in \mathcal{O}_K$ have norm g. The letter π without a subscript denotes the element of norm N used to define the extra column level. The spaces W and C are, respectively, the good part of the row space and the entire initial column space, of dimensions n and c. We use $n_j = k + j(t+1)$ for the ranks of the definite Hermitian spaces X_j, $0 \le j \le m$; thus n_j is not a dimension of a space of automorphic forms. Set $w(\tau) = w_{\mathrm{nom}}(\tau)$. The local ask A_p, the prime sum A, the gain g_p, and the direction cost $\mathcal{U}$ are as in the preceding section; in particular

$$g_p = m\left(\sum_{\tau\in\mathcal{T}} A_p(a,\tau) + k\sum_{\tau\in\Phi} A_p(w(\tau),\tau)\right). \tag{44}$$

Here we prove the quantitative implications between the jet theorem and Theorem A. We also isolate the remaining nontrivial automorphic and local representation-theoretic assertions on which the argument depends. They are identified explicitly as formal-source inputs when we do not give an independent proof of their full strength.

5.1 Coefficient lattices, class spaces and index

It is important to formulate the lattice estimates without assuming that the coefficient field of the automorphic forms is a fixed number field. For a rational prime p, fix a field isomorphism $\iota : \mathbb{C} \simeq \overline{\mathbb{Q}}_p$ as in the coefficient construction, and transport the valuation ring and absolute value to $\mathbb{C}$. We denote this ring by O_ι and normalize $|p|_p = p^{-1}$. All assertions below are quantified over ι. For a Hermitian space X and adelic point g', let $\|v\|_{\iota,g'}$ denote the Weyl gauge of its coefficient vector in the integral structure at g'. Thus, if ρ_κ is the algebraic coefficient representation, the gauge is obtained by applying the inverse of the adelic coefficient matrix to v and then taking the coefficient supremum norm transported by ι. It is non-Archimedean; multiplication by a scalar z multiplies the gauge by $|\iota(z)|_p$. At the ordinary complex absolute value the Weyl norm is instead the chosen Hermitian norm; the two kinds of estimate should not be confused.

For level K_j and weight κ of the given definite space X_j, define $M_j(\iota)$ to be the O_ι-module of forms of this type, every value of which has gauge at most 1. Write $M = M_0(\iota)$, a submodule

of C. All finite-place estimates in the proof are made on this adelic integral module, not just on abstract polynomials in a single weight representation. The exact definitions are `Data.integralColumns` and `integralFormsAt` in `New/Exchange/CrudeModularData.lean` and the background files respectively. The notation M in this section always depends on the chosen embedding.

We recall the index convention used in the proof.

Definition 5.1. Let P, Q be submodules over a valuation domain O, and let $z \in O$. Write $I(P, Q; z)$ if for every finite family $(v_i)_{i=1}^{h}$ of elements of P there exists a square h by h matrix D over O such that $\det D \mid z$ in O and

$$\sum_j D_{ij} v_j \in Q \quad (1 \le i \le h).$$

For an O-linear endomorphism T preserving P, write $T \in \mathrm{Type}(p, P, s; e)$ if there is an integer $l \ge 0$ with $l \le e$ and

$$I(P, TP + p^s P; p^l).$$

The parameter s is a nonnegative integer and e is allowed to be a nonnegative real number.

These are `IndexDvd` and `HasType` in `New/Core/ModuleIndex.lean`. The divisibility of the determinant is in the stated direction: $\det D \mid p^l$ implies $|\det D|_p \ge p^{-l}$. It does not mean that $\det D$ is a multiple of p^l. Nor is it necessary to assign a finite cardinality to a lattice quotient. This definition is appropriate even when O is not discretely valued, and is sufficient for the later height argument.

Lemma 5.2 (Elementary index calculus)**.** *Let $I(P, Q; z)$ and $I(Q, R; a)$ hold. Then $I(P, R; za)$ holds. If u is an O-linear map with $u(P') \subset P$ and $u^{-1}(Q) \cap P' \subset Q'$, then $I(P', Q'; z)$ holds. If $u(Q) \subset Q'$ and $P' \subset u(P) + Q'$, then $I(P', Q'; z)$ also holds. Consequently for $T_1, \dots, T_h$ preserving P,*

$$T_i \in \mathrm{Type}(p, P, s; e_i)\ (1 \le i \le h) \Longrightarrow T_1 \cdots T_h \in \mathrm{Type}(p, P, s; \textstyle\sum_i e_i).$$

Finally, if $u_i : P \to V'$ are linear and $I(P, u_i^{-1}(Q'); z)$ holds for each i, then

$$I\left(Q' + \sum_{i=1}^{h} u_i(P), Q'; z^h\right).$$

Proof. The first assertion is obtained by applying the index witness matrix for $I(Q, R; a)$ to the family already obtained using the matrix for $I(P, Q; z)$, and taking their product. The determinant is multiplicative. For pullback, apply the index condition to $(u(v_i))_i$ and use linearity. For pushforward, write $v_i' = u(v_i) + q_i'$ with $q_i' \in Q'$ and use the same linear combinations of these expressions.

For the product assertion it suffices to consider two factors. Suppose we have index witnesses for $T_i P + p^s P$, with determinant bounds p^{l_i}. Apply the second operator to the index condition for the first, if necessary reversing the order in this statement. Since $T_i(p^s P) \subset p^s P$, the pushforward rule gives the required condition from $T_1 P + p^s P$ to $T_1 T_2 P + p^s P$, and the transitivity rule gives the sum $l_1 + l_2$. Now use induction. For the final assertion enlarge Q' successively by the images. In each step, push the index condition along u_i and use transitivity. □

For comparison, if $P = O^D$ is free over a discretely valued ring and T has elementary divisors $p^{\nu_1}, \dots, p^{\nu_D}$ (with $\nu_i = +\infty$ allowed), the determinant valuation of the canonical lattice $TP + p^s P$ is

$$\sum_i \min(\nu_i, s).$$

Thus a bound on the type measures a truncated sum of elementary divisors. The Hermite and Smith factorization arguments in `ModuleIndex.lean`, in particular `IndexDvd.of_det_dvd`, implement this relation for the valuation domains of the proof. We will use type bounds, not an assertion about each individual eigenvalue of a Hecke operator.

5.2 The groups and the chain of theta kernels

At stage j the source is the standard definite space X_j of rank n_j. The column level K_j is the intersection of the chosen principal level at the places dividing $\overline{\pi_g}$ with the subgroup whose last row is congruent to $(0, \dots, 0, 1)$ modulo π. The element π is chosen so that $\pi\bar{\pi} = N$. There are f places dividing π, one member of each conjugate pair over the completely split prime N. In particular K_j contains the principal level of $\overline{\pi_g}\pi$ and is compact open. The spaces of forms are finite sums of coefficient spaces indexed by rational adelic classes. The chosen principal level makes the automorphism groups of the classes trivial. These claims are `Data.columnLevel`, `Data.principalLevel_le_columnLevel`, `Data.card_piPlaces`, and `Data.classAut_columnLevel_eq_bot` in `New/Columns/ColumnDefs.lean`. The analogous triviality at the intermediate stages follows from `Data.classAut_eq_bot_of_le_chosenLevel`. We use this specific property as a formal-source input for the adelic lattice model; no estimate for a weighted stabilizer sum can replace it in the following argument.

Here is how the coefficient weight is structured. With $h = 2h_2$ as in the admissible data, the number of columns of height j' in the Young diagram at τ is

$$\kappa_\tau(j') = \begin{cases} a, & j' = 1,\ \tau \in \mathcal{T}, \\ 0, & j' = 1,\ \tau \notin \mathcal{T}, \\ w_1(\tau) - \lfloor (k+t+1)/2 \rfloor, & j' = k, \\ h - (t+1), & j' > k, \quad t+1 \mid j' - k, \\ 0, & \text{otherwise.} \end{cases}$$

Only columns of height below n_j contribute to the polynomial part at stage j. The determinant exponent of the representation is 0 at stage zero, and $e_j(\tau) = \lfloor n_{j-1}/2 \rfloor - w_j(\tau)$ for $j \geq 1$. The coefficient representation on X_0 is in particular

$$\bigotimes_{\tau \in \mathcal{T}} \mathrm{Sym}^a \mathbb{C}^k.$$

All factors at the other places have degree zero. The restriction on the places of height-one columns is central in the collision argument: once a p-adic Frobenius orbit meets both top places, the two factors of degree a can interact at precisely the prime-power levels that contribute to the required gain.

Let θ be an admissible twist of finite conductor of rational norm q_θ, and let μ_{j+1} be the prescribed splitting character of the $j+1$st lift, with twist θ^{j+1}. The theta kernel K_j is the one for $\mathrm{Hom}(X_j, X_{j+1})$ with weight κ, character determined by $\mu_{j+1}\theta^{j+1}$, and complex Taylor scale $\overline{\pi_g}$ per. The identities between weights that make the lift available use the fact that there is no nonzero column height strictly between n_j and n_{j+1}. The declarations are `Data.isLiftWeight_kappa` and `Data.liftKernel` in `New/Exchange/CrudeModularData.lean`; the formulas for κ, e_j are in `New/Curve/LiveSum.lean`.

For a tuple of finite adelic matrices $\xi = (\xi_j)_{0\le j<m}$ in the corresponding Hom spaces, let $\phi_{\xi,j}$ be the finite Heisenberg translate by ξ_j of the normalized integral vacuum of the preceding section. Let

$$\mathsf{C}^{j}_{\theta,\xi} = \mathsf{K}_{j-1}(\phi_{\xi,j-1}) \circ \cdots \circ \mathsf{K}_0(\phi_{\xi,0}), \qquad \mathsf{C}^{0}_{\theta,\xi} = 1$$

be the composite definite lift to stage j. Individual terms need not be invariant under the intermediate level. A useful datum is instead a finite linear combination of tuples $\Psi = \sum_\xi c_\xi[\xi]$, which is required to be jointly invariant under the levels in the sense that the corresponding tensor of kernel data is fixed by their simultaneous actions.

At the final stage X_m has rank one greater than Y_m. The map to the last row space is not simply evaluation on $U(Y_m)$: it uses branching of the algebraic coefficient representation. Let D_{br} be the integral branching matrix and put

$$e_{\mathrm{br}}(\tau) = \sum_{j'<n_m} (\kappa_\tau(j') - \kappa'_{\mathrm{kept}}(j')) + e_m(\tau).$$

Here the summands count the columns at that height whose height is reduced in branching; the formal definition of κ'_{kept} is $h_2 - \lfloor (t+1)/2 \rfloor$ if $j' \ge k$ and $t+1 \mid j'-k$, and zero otherwise. Thus the linear restriction is

$$(\mathsf{R}v)(u) = \left(\prod_{\tau\in\Phi} \tau(\bar\pi)^{-e_{\mathrm{br}}(\tau)} \right) D_{\mathrm{br}} v(\mathrm{diag}(u,1)). \tag{45}$$

The relevant branching identity is `Data.restrictToY`, and the estimate `Data.le_branchExp` is

$$e_{\mathrm{br}}(\tau) \ge \kappa_\tau(1) - 2b.$$

For a prime $p \ne N$, the scalar in (45) is a p-adic unit; using the lattice-compatible branching identity the restriction cannot increase the gauge. At a complex conjugate σ the estimate is stronger:

$$\|\sigma\mathsf{R}v(u)\|^2 \le \exp\left(x^{70} - \textstyle\sum_\tau e_{\mathrm{br}}(\tau)\log N\right) \|\sigma v(\mathrm{diag}(u,1))\|^2. \tag{46}$$

Indeed under each conjugate $|\tau(\bar\pi)|^2 = N$, so the scalar gives the displayed negative exponent; the branching matrix contributes at most $e^{x^{70}}$ by the coefficient estimate. The precise lattice and norm compatibility statements `Data.GaugeLe.restrictToY` and `Data.weylNormSq_restrictToY_le`, in `New/Seesaw/LiftCharacters.lean`, are formal-source inputs. Their significance is that the archimedean gain of level N is large, while the prime N itself will be treated separately.

Write

$$\mathsf{B}^{\Psi}_{\theta} = \mathsf{R} \sum_{\xi} c_\xi \mathsf{C}^{m}_{\theta,\xi}.$$

This is `Data.Bμ` in the source. For the local calculations the datum must have the following exact properties (the record `Data.IsGoodDatum`, in `New/Columns/ChainValuation.lean`):

1. If ξ lies in the support, its components multiplied by π_g are integral away from the places of $q_\theta N$.
2. The coefficients c_ξ are integral at every ι with $p \nmid q_\theta$. For $q_\theta = 1$ the sum of the absolute values of their conjugates is at most $e^{x^{60}}$.
3. The total tensor is jointly invariant under the intermediate column levels and the final action specified in the source.

4. At N the chain has an additional paired description. Project each kernel datum to the places dividing π; sum the pairwise translations of adjacent data by last-column $-y/N$ and last-row $-\bar{y}/N$ for y modulo N. The resulting combination of paired lifts, with the same coefficients c_ξ, equals the original chain of m lifts. In the source $m = 2(\varrho + 1)$.

The fourth property, `chain_at_N`, will make it possible to move large apparent denominators at N onto a single column operator, instead of losing them at every Taylor coefficient.

Lemma 5.3 (Gauge bounds for a jointly invariant chain)**.** *Suppose $p \nmid q_\theta$, Ψ is good and each translated kernel after a chosen modification obeys*

$$\|\mathsf{K}_j(\phi_{\xi,j})v\|_{\iota,u} \le p^{r_j}\|v\|_{\iota,g'}$$

at the prescribed coefficient frames, uniformly for each pair of points and every ξ in the support. Assume also that the modifying operators commute with the required source and target level actions on the total invariant tensor. Starting with an integral column, the total modified chain has gauge at most $p^{\sum_{j<m} r_j}$ in the final frame.

Proof. For jointly invariant tensors all the integrals over definite groups reduce to finite class sums; an individual term is a composition, indexed by a tuple of representatives of the successive class spaces. The triviality of the stabilizers means that such a sum is a plain sum of the corresponding values, not a sum with rational denominators from the stabilizer orders. Each composite has gauge at most the product of the individual bounds. The coefficients c_ξ are integral since p is prime to q_θ. Thus the non-Archimedean triangle inequality leaves the largest gauge bound unchanged under summation. This proves the result. The source formalizes the reduction to the class sum and this argument as `Data.IsGoodDatum.gaugeLe_chain` via `Data.exists_chain_eq_sum` in `New/Columns/ChainValuation.lean`. □

The same argument at an archimedean place bounds the square root of the squared norm by the product of individual operator norm estimates, the sum of absolute values of the coefficients, and the product of the numbers of integral classes at each stage. This explains why the class-counting estimate and coefficient-size estimate must be checked even when no such factor is present in the non-Archimedean estimate.

5.3 The geometric seesaw and its denominator control

Let P be the (unprojected) Poincaré construction from columns to initial row forms. It comes from a Bessel period for the unitary space Y_0; the successive indefinite theta lifts run through the spaces $Y_0, \dots, Y_m$ of ranks $r + jt$. Our notation Θ_θ is reserved for the composite on W (or its extension to the unprojected principal-level row space before selecting W). Initially we do not identify it with any definite chain.

Theorem 5.4 (Seesaw identity, formal-source input)**.** *There is a nonzero $z_0 \in \mathbb{C}$, independent of the twist, and for every admissible twist θ a good datum Ψ_θ such that*

$$\Theta_\theta P(z_0 u) = \mathsf{B}_\theta^{\Psi_\theta} u \qquad (u \in C^{\mathrm{ram}}).$$

The declaration is `Data.seesaw` in `New/Seesaw/SeesawTheorem.lean`. The exact content includes the four properties of good data above, not just an equality of two complex-valued functions. In the following we explain the denominator mechanism and give the finite cut argument in a form that is independent of the deeper theta calculations.

In the adelic construction a stage j has a rational Bessel subgroup, with finite-adelic component denoted here by S_j, and a compact open subgroup $J_{j,\theta}$ of the row group. The

function arising after the first j unfoldings is P_j, with P_0 the generic function from which the original Poincaré series is formed. The rule $P_{j+1} = \mathsf{U}_j(P_j)$ is a Bessel unfolding by a single step of theta lifting, at parameters (ξ_j, y_j) of the archimedean and finite row spaces.

Three deeper assertions of the source are needed in this construction; we do not reprove the full geometric or local theta identities.

1. The local constant-term identity preserves the generic character, and consecutive pairs of unfolding steps give matrix coefficients of a definite paired chain. At the even step $2i + 2$, the equality is

$$P_{2i+2}(\iota_X v \cdot (x, y_{2i+2})) = \mathrm{algCoeff}_{t_i}\,(\mathsf{C}^{\mathrm{pair}}_{i+1}(u))(\mathrm{emb}_X v)$$

   for all ramified columns u and all configurations y with initial point y_0 in the chosen level. The linear functional t_i depends on the archimedean parameters but not on the twist or on u. These are `Data.isGeneric_unfoldStage` and `Data.unfoldStage_even` in `New/Seesaw/UnfoldingStages.lean`; the local arrow-pair tensor and integral identity used in their proof are in `New/Seesaw/EdgePair.lean` and `New/Seesaw/DoubleLiftValue.lean`. In the formal proof one unfolds the two consecutive theta kernels, takes their constant terms with respect to the unipotent subgroup, and applies the identity for integrating a product of kernels. This gives $i + 1$ paired lifts from the assertion for i.

2. The support of each stage function is controlled by the lattice lines of its Hermitian space. If P_j is nonzero, the lines at places of the twist move at most j steps, and the integral lines outside those places do not move (`Data.hasNearLines_of_unfoldStage_ne_zero`, same file). In the induction, even steps are covered by the line estimate for the theta kernel. At odd steps the possible exceptional local lines at the level prime are eliminated by writing a ramified column as $u' - z_w u'$ for a central generator z_w at a place over N. The two terms of the matrix coefficient are then equal there, so cancel. This use of ramification is why the seesaw statement is on C^{ram}.

3. For each stage one can choose a compact open A_j independent of the twist and a finite cut $T_{j,\theta} \subset G_{j,f}/J_{j,\theta}$ stable under $S_j \cap A_j$ and contained in the level outside the twist conductor. Here $G_{j,f} = U(Y_j)(\mathbb{A}_f)$. The cut covers the support after translation by S_j and has cover multiplicities which divide powers of q_θ. At $j = 0$ it is contained in the chosen level. This assertion is `Data.exists_stage_cut` in `New/Seesaw/SeesawTheorem.lean`.

Lemma 5.5 (Finite-cut multiplicities)**.** *Let G be a locally compact totally disconnected group, S a closed subgroup, and $J \subset A$ compact open subgroups of G. Let $D \subset G$ be compact, and suppose that the set to be covered is in SD. Suppose for each $x \in D$ there is an e such that*

$$(xux^{-1})^{q^e} \in A \qquad (u \in J).$$

There is a finite $(S \cap A)$-invariant subset $T \subset G/J$ whose S-saturation covers the image of SD. For v in this saturation, write

$$m(v) = \#\{s(S \cap A) : s \in S,\ s^{-1}v \in T\}.$$

Then $m(v)$ divides a power of q.

Proof. The image of D in the discrete space G/J is finite. Choose one element $[x]$ in that image from each S-orbit, and take T to be the union of their $(S \cap A)$-orbits. These are finite

since $S \cap A$ is compact. If H_x is the stabilizer of $[x]$ in S, then the s with $s^{-1}[x] \in T$ constitute $H_x(S \cap A)$. Thus

$$m([x]) = [H_x : H_x \cap A].$$

This index is finite, and is invariant under translating $[x]$ by S. Pass from H_x to its quotient by the normal core of $H_x \cap A$; this is a finite group. As $H_x \subset xJx^{-1}$, every element in the finite quotient has order dividing a power of q. Indeed for each $h \in H_x$ and each conjugate by an element of H_x, that conjugate's q^e-power belongs to A, hence h^{q^e} belongs to the normal core. By Cauchy's theorem every prime divisor of the order of this finite group divides q. The index divides its order, proving the assertion. This is the finite-index argument in `exists_cut_coverMultiplicity_dvd`, `New/Seesaw/IntegralAsCosetSum.lean`. □

The cut hypothesis is justified in the source by the support of the unfolded stages. Outside the twist the local groups are integral. At the finitely many places of the twist, taking a sufficiently large power of the residue characteristic puts an element of the conjugated compact open level back into the fixed level. The admissible translations therefore lie in finitely many local compact sets modulo S_j; the stage-cut theorem applies the lemma to those sets. Thus the importance of the local geometric support bound is not just convergence: it is the prescribed prime support of denominators.

At the analytic level the unfolding reduces an integral over the rational Bessel quotient to a sum over a finite cut. In the notation of the lemma the formula has the form

$$\int_{B(F)\backslash G(\mathbb{A}_F)} P_j(g)z(g)\,dg = \lambda_j \sum_{v \in T_{j,\theta}} m(v)^{-1} b_j(v).$$

The scalar λ_j is determined by the fixed archimedean integral and Haar normalizations; it is independent of the twist. To see the finite sum mechanism, the weight $\mathbf{1}_T(v)/m(v)$ is a partition of unity after summing over $S_j/(S_j \cap A_j)$ on the support. Indeed exactly $m(v)$ terms contribute, with the same orbit multiplicity. It may therefore be inserted in the integral without changing it. Integration over a coset of the compact subgroup $J_{j,\theta}$ contributes its Haar mass; the standard level-index normalizations in the period over B cancel this factor. The remaining archimedean integral is λ_j. The analytic hypotheses are precisely the generic character identity, automorphy of z, integrability of $P_j z$ and the product form of its Bessel-group integral. The formal-source input for this quotient-integral statement, including its measure normalizations, is `Grading.exists_integral_eq_sum_over_cut` in `New/Seesaw/IntegralAsCosetSum.lean`.

The product of the resulting finite combinations along the chain is again a finite combination: its coefficients are sums of products of $1/m(v)$. For $p \nmid q_\theta$ these are integral. If the twist is trivial, $q_\theta = 1$, the multiplicities are all 1 wherever the cut is nonempty. For quantitative complex sizes one counts the cut at each stage by the index of the chosen principal level; this gives a power of `Data.cutBound` independent of the actual columns. The formal assertion `Data.exists_isAdmissibleComb_unfolding` in `New/Seesaw/SeesawTheorem.lean` carries out this induction, testing against every suitable form z and linear functional on the tensor space. The terminal stage identifies the coefficient with paired lifts, using the preceding even-step identity.

To turn the resulting tensors into the vacua of the definite chain one needs a further finite Weil representation calculation. For the evaluated kernel tensor at each configuration y which is integral outside q_θ, the source proves that it is an admissible linear combination of reduced vacua, up to a single constant C and with denominator $\mathrm{cutBound}^{2m}$, bounded size when the twist is trivial. The reduced vacua are the paired combinations of Heisenberg translates described above. The exact declarations are `Data.exists_isAdmissibleComb_reducedVacu`

`um`, `Data.isAdmissibleComb_reducedVacuum`, and `Data.exists_isGoodDatum_of_invariant` in `New/Seesaw/GoodDatumExists.lean`. We take the three assertions as formal-source inputs; they contain the explicit evaluation of finite Weil data and the local fiber geometry, rather than just the group-theoretic finite-cut calculation. The scalar C and the archimedean normalization from the integrals are absorbed into the one scalar z_0. The theorem then follows by pairing against the test vectors at stage m; equality of all those matrix coefficients gives equality of the lifts. These are the mechanisms underlying Theorem 5.4.

5.4 Ramification and short generators on the definite group

The set of ramified columns has an elementary description which is compatible with its adelic lattice. For each $v \mid \pi$ choose a central element of the local norm-one torus, integral everywhere, supported on $v, \bar{v}$, whose component at v reduces to a generator of $\mathbb{F}_N^\times$. Let z_v be its translation operator on columns. Define

$$R_N = \prod_{v|\pi}(1 - z_v), \qquad C^{\mathrm{ram}} = R_N C.$$

These operators commute. In particular every element of the image can be written $(1 - z_v)u'$ at each $v \mid \pi$. The definition and the lift of the generators are in `Data.ramifyingOp`, `Data.ramifiedColumns` and `Data.exists_centralGen` in `New/Columns/ChainValuation.lean`.

The following dimension and norm observations will be useful both here and for the spanning theorem.

Lemma 5.6. *For each $v \mid \pi$,*

$$\dim \ker(1 - z_v) \le \frac{c}{N-1}.$$

Consequently for any subspace $U \subset C$ and any subset T of the places dividing π, with $R_T = \prod_{v\in T}(1 - z_v)$,

$$\dim R_T^{-1}(U) \le \dim U + |T|\frac{c}{N-1}.$$

If u is integral at every embedding and every value of its conjugates has squared norm at most C_0, $R_T u$ is likewise integral, with squared norm at most $4^{|T|}C_0$.

Proof. The forms fixed by z_v have level enlarged by the cyclic group generated by the central translation. The index of the old level in this enlargement is at least $N - 1$: a positive power lying in the old level reduces to the identity at v. Since both levels are contained in the chosen level and their class automorphism groups are trivial, their spaces have dimension proportional to the numbers of adelic classes. The ratio of these class counts is the level index; thus the dimension of the fixed space is at most $c/(N - 1)$. This is the class-count formula specialized in `Data.finrank_ker_one_sub_centralGen_le`. For a linear operator L, $\dim L^{-1}(U) = \dim \ker L + \dim(U \cap \operatorname{im} L)$. Applying this successively proves the second assertion. The central element preserves the integral gauge and permutes the arguments of an algebraic form. Therefore $u - z_v u$ is still integral by the ultrametric inequality, and its squared complex norm is at most $4C_0$ by the ordinary triangle inequality. Induction gives the bound. This is `Data.ramifyingOp_finrank_comap_and_short`. □

To produce the initial short vectors the source uses a geometry-of-numbers estimate for definite Hermitian lattices. We record its role explicitly. Let I be a level whose class stabilizers are trivial, let the coefficient representation of rank u have a Young diagram with D_κ boxes (counted over the infinite places), let $Q \ge 1$ be a rational integer in the different ideal of

E, and suppose $M_0 \geq uf + D_\kappa$. Every class admits an algebraic basis of its associated lattice with the uniform bounds summarized in `exists_short_basis_stdSpace`. Put $R_0 = 6Q(fu)^2$. Then the coefficient forms which are everywhere integral with squared norms of all conjugates bounded by $R_0^{6M_0}$ span the whole class space (`span_isShort_eq_top`, in `New/Columns/ChainValuation.lean`). This is a formal-source input for the quantitative short basis bound; the coefficient argument is as follows. The class space is freely the direct sum of weight modules at the class representatives, so it suffices to span one such weight module. Apply the Schur representation to the bounded lattice basis. In the standard coefficient model the Weyl vectors obtained by evaluating integral flag polynomials at roots of unity span the weight module: if a linear functional annihilates all such vectors, successive application of the root bound for a one-variable polynomial shows that the corresponding polynomial vanishes identically. Multiplication by the determinant factor for the determinant weight also preserves integrality in the chosen lattice basis. Every conjugate of a root of unity has absolute value 1, so the norm bound follows from the power of R_0 for the bounded basis and the degree of the coefficient representation. This is why one obtains a spanning set rather than simply a single short vector. Applying this construction at level K_0 with the admissible scale estimates, and then applying R_N, gives the needed conclusion:

$$\operatorname{span}_{\mathbb{C}}\{u \in C^{\mathrm{ram}} : u \text{ is everywhere integral, } \|\sigma u(g')\|^2 \leq e^B \ \forall \sigma, g'\} = C^{\mathrm{ram}}. \tag{47}$$

Specifically $|\{v \mid \pi\}| = f$, and the bound before application of R_N is chosen at most $e^B/4^f$. The exact scale inequality and conclusion are `Data.span_short_ramifiedColumns`. We will use (47) with the ordinary complex norm squared, not its logarithm, bounded by e^B.

5.5 The collision loss and the marked set

For a rational prime p, put

$$\omega_p = \frac{1 - \chi_K(p)}{2}, \qquad s_p = \lceil \log_p m \rceil, \qquad \delta_p = s_p f(n_m + s_p).$$

A prime is called short if $p \equiv 1$ or $-1 \pmod m$. In this terminology shortness refers to a congruence and not to the size of p. Define

$$\ell_p = \frac{fa(\log b)^2}{p} \begin{cases} m, & p \text{ short}, \\ 1, & \text{otherwise}. \end{cases}$$

We will call p small nonsplit if $p < x_1$ and $\chi_K(p) \neq 1$. At such a prime, g_p in (44) is the desired total valuation. In the formulas below, all relevant place actions are the actions of the CM cyclotomic type, not of the single quadratic extension $K/\mathbb{Q}$ alone.

For a subset $E_0 \subset \Phi$ of places declared even, set

$$C_p(E_0) = \sum_{\tau \in \Phi} \sum_{\substack{e \geq 1,\ e \text{ odd or } p=d \\ \tau, -p^e\tau \in \mathcal{T} \setminus E_0}} \frac{a}{p^e}.$$

This collision count removes all pairs for which one of the two places belongs to E_0. Let o_0 be the order of the class of $-p$ modulo N_0 (using the place convention of the data). We call p served if it is inert in K, not short, and o_0 is odd with $o_0 \leq s_p$. These are the definitions `Data.IsShortPrime`, `Data.loss`, `Data.ceilLogM`, `Data.deltaP`, `Data.IsServed` and `Data.collisionSum` in `New/Curve/LiveSum.lean`.

Lemma 5.7 (Collisions without exceptional treatment)**.** *For p small nonsplit and not served,*

$$mC_p(\varnothing) \le \ell_p.$$

Moreover, for every small nonsplit p one has $C_p(\varnothing) \le 2fa/p$.

Proof. If all the exponents appearing in a collision sum are at least o, then

$$C_p(E_0) \le fa \sum_{e \ge o} p^{-e} \le 2fa/p^o,$$

since $p \ge 2$. The second assertion follows by taking $o = 1$. The first assertion for short p follows as soon as $(\log b)^2 \ge 2$. If p is not short, recall that a collision has $p \ne d$ and odd e divisible by o_0 (`Data.IsCollision.ne_and_odd_and_dvd`, `New/Columns/JetExponentBound.lean`). If there are any collisions for such a prime, o_0 is odd; the assumption that p is not served therefore gives $o_0 > s_p$. Consequently

$$mC_p(\varnothing) \le 2mfa/p^{s_p+1} \le 2fa/p \le \ell_p.$$

The source's assertion is `Data.mul_collisionSum_le_loss`. The arithmetic fact excluding collisions at d uses the congruence hypotheses on d; without those hypotheses the argument at a ramified prime would require modification. □

The issue at a served prime is different: exponents $e \le s_p$ can contribute, and multiplying a loss of size a/p by m cannot be absorbed in ℓ_p. The prime-by-prime repair uses a set of split places of E/E^+ even though p is inert in $K/\mathbb{Q}$. Fix ι and denote by $l(\tau)$ the prime of E defined by the embedding attached to τ and ι.

At a served prime, write $u = -p$ in the group of infinite places. Its order is divisible by m:

$$m \mid \mathrm{ord}(u). \tag{48}$$

For clarity, the basic 2-power estimate behind this assertion is the following. If an odd integer v is congruent to neither 1 nor -1 modulo 2^h, then the order of its image modulo 2^s has 2-part at least 2^{s-h+1} provided s is sufficiently large. For an even exponent L the identity

$$v_2(v^L - 1) = v_2(v-1) + v_2(v+1) + v_2(L) - 1$$

follows by repeated factorization of $v^{2j}-1$. Since one of $v-1, v+1$ has valuation 1 and the other less than h, the required bound follows from divisibility by 2^s. Here $m = 2^h$, the cyclotomic exponent is $s = 22h$, and the shortness exclusion allows the estimate with ample slack. The local embeddings of the cyclotomic tower, including primes dividing the fixed conductor, are handled in `Data.IsServed.m_dvd_orderOf` in `New/Columns/JetExponentBound.lean`. We use that exact assertion for the admissible data.

The other relevant arithmetic property is

$$l(u^e\tau) = \overline{l(\tau)} \quad (e \text{ odd}), \qquad l(\tau') \in \{l(\tau), \overline{l(\tau)}\} \iff \tau' = u^e\tau \text{ for some } e. \tag{49}$$

These are `Data.inducedPlace_neg_pow_mul_of_odd` and `Data.inducedPlace_eq_or_star_iff`, same file. In particular, a served prime of the data is unramified in the cyclotomic tower. One way to see the parity assertion is to express the decomposition group through Frobenius in $\mathrm{Gal}(E/\mathbb{Q})$: its restriction to K is conjugation and its restriction to the cyclotomic part is the pth power. The element $-p$ of the CM place group is obtained by composing this Frobenius with complex conjugation. Thus it interchanges the two induced primes above the real place. The order condition (48) ensures that these two primes are distinct. The full arithmetic assertions `smul_eq_iff_inert`, `smul_eq_star_iff_inert`, `q_eq_inert` and the unramified/splitting hypotheses used for these formulas are formal-source inputs for the specific extension E.

Lemma 5.8 (Marking)**.** *Fix a small served prime and a p-adic embedding. There is a finite set Λ of primes above p such that:*

1. *no $l \in \Lambda$ has $\bar{l} \in \Lambda$;*
2. *each marked prime l is induced by a top;*
3. *for every top τ, either $l(\tau)$ or $\overline{l(\tau)}$ belongs to Λ.*

Write $E_0(\Lambda) = \{\tau : l(\tau) \in \Lambda\}$ for the set of even places. Then

$$|\Lambda| \le f, \qquad C_p(E_0(\Lambda)) = 0.$$

Every τ with $l(\tau)$ or $\overline{l(\tau)}$ marked has $\tau \notin \mathcal{S}$.

Proof. For each top, multiplication by u^{o_0} stays within the tops and, because o_0 is odd, interchanges the two induced primes by (49). Hence the primes induced by tops occur in conjugate pairs. Choose one from each pair. This gives all three properties and the cardinality bound. Every collision has e odd, so the primes induced by its two top places are conjugate. At least one is marked, and consequently no collision remains outside $E_0(\Lambda)$.

For the last assertion, the orbit condition (49) gives a power of u taking the place τ to a top. Its class modulo N_0 is therefore in the cyclic subgroup generated by the class of u. That subgroup has odd order. A thick place has class -1, of order two in the residue group of the fixed modulus N_0; it cannot lie in an odd-order subgroup. The exact statements are `Data.exists_isMarked`, `Data.IsFullMark.collisionSum_evenPlaces_eq_zero` and `Data.IsMarked.notMem_thick`. □

We will also need the balancing property of marking. If $G(c, l)$ is any function of an integer weight and a prime, zero unless $l \in \Lambda$, then

$$\sum_{\tau} G(w_j(\tau), l(\tau)) = \sum_{\tau} G(w_j(\tau), \overline{l(\tau)}). \tag{50}$$

Indeed multiplication by u exchanges the two primes. All places in these paired orbits are non-thick by the previous lemma, so the functions w_j are constant there, with the common weight determined by b_2. Reindexing the first sum by $\tau \mapsto u\tau$ proves the equality. The source's exact assertion is `Data.IsMarked.sum_inducedPlace_sub_star_eq_zero`. The equality will ensure that the local factors from the splitting characters in the Hecke transfer are integral.

5.6 Moving the marked Hecke operators through the chain

For each marked l, consider the spherical Hecke operator T_{l,s_p} corresponding to

$$\mathrm{GL}_k(\mathcal{O}_{E_l}) \,\mathrm{diag}(p\mathbf{1}_{s_p}, \mathbf{1}_{k-s_p}) \,\mathrm{GL}_k(\mathcal{O}_{E_l})$$

on the column space. We use the unnormalized coset-sum convention of the source's `heckeSplit`. Put $A_l = a\#\{\tau \in \mathcal{T} : l(\tau) = l\}$ and set $Y_l = p^{A_l} T_{l,s_p}$ and $Y = \prod_{l \in \Lambda} Y_l$ in a fixed order. By the scale constraints p is outside the column levels g and N. The local group is split and p is a uniformizer at l. The use of s_p rather than just the rank-one Hecke operator is part of the local repair strategy.

Let $J_{p,j}(E_0)$ denote the jet exponent for the jth lift established in the preceding section, and let $L_j(\tau)$ denote the number of columns in the weight at τ, below n_{j+1} and of heights other than one. Define

$$D_{p,j}(\Lambda) = s_p \sum_{\tau \in E_0(\Lambda)} L_j(\tau) - J_{p,j}(E_0(\Lambda)).$$

The quantity $D_{p,j}(\Lambda)$ may be negative; it is a logarithmic p-adic gauge exponent, not a loss required to be nonnegative. The definition is `Data.liftLoss` in `New/Columns/LocalStatement.lean`. For the following discussion a jet datum multiplying gauges by at most p^u will mean exactly membership in the source's submonoid `Kernel.raises` with prescribed frames and bound p^u.

Proposition 5.9 (Two local compatibility assertions, formal-source inputs)**.** *Let p be small nonsplit, let Ψ be good for a twist with $p \nmid q_\theta$ and let $l \in \Lambda$. For every $j < m$:*

1. *There are integral coefficients $c_0, \dots, c_{s_p}$ such that, on the partial definite chain $\mathsf{C}^{j+1}_{\theta,\xi}$ with any $\xi \in \operatorname{supp} \Psi$,*
$$\sum_{i=0}^{s_p} c_i T^{(j+1)}_{l,i} \mathsf{C}^{j+1}_{\theta,\xi} v = p^{\delta_p} \mathsf{C}^{j+1}_{\theta,\xi} T^{(0)}_{l,s_p} v$$
for spherical columns v, with spherical target datum at l.

2. *If each u_l belongs to one of the basic double cosets at l of index $i_l \le s_p$, the modified jth datum, translated on the target side by the product of the u_l, multiplies gauges by at most*
$$p^{\sum_{l \in \Lambda} A_l + D_{p,j}(\Lambda)}.$$

The declarations are `Data.IsGoodDatum.exists_hecke_transfer` and `Data.IsGoodDatum.mem_raises` in `New/Columns/LocalStatement.lean`. We explain precisely how the stated estimates enter the chain, and the source of each term.

At a split unramified place, the Hecke polynomial of a theta lift differs from that of the preceding stage by a polynomial with explicit powers of the size of the residue field and values of splitting characters. In the proof the source polynomial is evaluated with a rank-dependent normalization, and the constant term of the extra polynomial is 1. One can therefore recover its coefficient of degree s_p by multiplying the target polynomial by the truncated formal inverse of the extra polynomial. This is simple polynomial algebra: for $U(X) = 1 + u_1 X + \cdots$ over a ring, $U(X)^{-1} \bmod X^{s_p+1}$ has coefficients in the same ring, recursively $v_j = -\sum_{i=1}^{j} u_i v_{j-i}$. The unramified transfer identity for the single kernel is applied successively to all earlier lifts before extracting the coefficient. The weight balance (50) controls the norms of the values of the splitting characters at l; the coefficient of extraction uses at most p^{δ_p} to clear the rank normalizations. The exact normalized coefficient identity is the deeper part of assertion (1); it is not a consequence of unnormalized Satake transfer without the denominator computation.

For assertion (2), write each basic double coset representative as an integral invertible matrix times a diagonal matrix times an integral invertible matrix. At the marked places its diagonal has p in a set of at most s_p coordinates, and at conjugate places its entries are inverted by the unitary action. At infinite places this changes the coefficient frame for the local Schur-jet estimate. The top height-one columns contribute a units of gauge cost for each top inducing l, yielding A_l. At the other columns one incurs at most s_p times the number of columns at even places. The tilted Schur-jet bound supplies divisibility $J_{p,j}(E_0(\Lambda))$. The factors from determinant characters cancel between conjugate marked places, again by (50). This gives the exponent in assertion (2). It is crucial that the theta-jet estimate applies to

the actual tilted data: the Hecke operator has modified the kernel vectors, rather than merely changed their scalar coefficient by p^{A_l}. The proof of `mem_raises` composes the diagonal-frame jet estimate of the preceding section with integral covariance.

Lemma 5.10 (Iteration of the repair)**.** *For integral columns v and $p \nmid q_\theta$, with Ψ good,*

$$\|\mathsf{B}_\theta^\Psi(Y^m v)\|_\iota \le p^{\sum_{j<m}(|\Lambda|\delta_p + D_{p,j}(\Lambda))}.$$

Here the bound is on every gauge of the restricted chain.

Proof. Move one marked operator through the partial chain ending at $j+1$, using Proposition 5.9(1). On the jth kernel datum this replaces it by the operator

$$p^{A_l-\delta_p}\sum_{i\le s_p} c_i T_{l,i}^{(\mathrm{datum})}$$

on the finite Weil representation on the target side of that kernel. On its gauge the scalar factor contributes $p^{\delta_p - A_l}$. Target-translation covariance and the spherical invariance of the datum guarantee this operator commutes with the required simultaneous level actions. When the same operation is applied at a second marked place, the first remains in the same finite kernel; the local groups at distinct places commute. Repeating for a list I of distinct marked primes gives a new datum at stage j and, for the as yet unprocessed local translates, gauge exponent

$$\sum_{l\in I}(\delta_p - A_l) + \sum_{l\in\Lambda} A_l + D_{p,j}(\Lambda).$$

Indeed the remaining Hecke actions are sums of basic coset translates as in assertion (2), with integral coefficients; the ultrametric inequality does not change the exponent when adding these terms. Taking $I=\Lambda$ leaves $|\Lambda|\delta_p + D_{p,j}(\Lambda)$. Modify every stage by this construction. Inductively the resulting first j kernels act on v as the original partial chain on $Y^j v$: the new last kernel has exactly absorbed one more Y through the partial chain. Thus the modified full chain equals the original chain on $Y^m v$. Lemma 5.3 adds the exponents. Finally the restriction to Y_m does not increase gauges, since $p \ne N$. This is the deduction `Data.IsGoodDatum.existsts_move_markOperator` and `Data.IsGoodDatum.gaugeLe_markOperator_pow`. □

Here is the numerical consequence. The jet exponent of the preceding section obeys the exact inequality

$$J_{p,j}(E_0) \ge \frac{g_p}{m} - C_p(E_0) + \omega_p \#(\mathcal{T}\cap E_0)\frac{a}{x_1} - \frac{fa(\log b)^2}{pm}. \tag{51}$$

For marked primes the other bound required is

$$|\Lambda|\delta_p + s_p \sum_{\tau\in E_0(\Lambda)} L_j(\tau) \le \frac{\ell_p}{m} + \#(\mathcal{T}\cap E_0(\Lambda))\frac{a}{x_1}. \tag{52}$$

This is `Data.IsMarked.sum_lowColumns_le` in `New/Columns/JetExponentBound.lean`. To explain the special top contribution here, the diagram count gives $L_j(\tau) \le w(\tau) + x^{60}$ and the number of boxes in these columns is between $kw(\tau)$ and $kw(\tau)+x^{60}$. Every marked place is non-thick, hence $w(\tau) = a/x_1$. For marked nonempty Λ the prime is served and not short. If $p \le m$, the factor $s_p \le 1 + \log_p m$ times a/x_1 is negligible in ℓ_p/m because of the separation $x \ll x_1$. If $p > m$, then $s_p = 1$ and the served condition forces $o_0 = 1$. Any even place in a marked orbit of tops is therefore itself a top. The contribution a/x_1 per such place is accounted for exactly by the term on the right of (52). In both cases $|\Lambda| \le f$, and the

terms $|\Lambda|\delta_p$ and $s_p f x^{60}$ are absorbed by the remaining margin $fa(\log b)^2/(pm)$. The formal calculation makes the constants explicit uniformly up to $p < x_1$; its scale inequalities are used here as formal-source bounds. The large a/x_1 term in (52) could not simply be discarded: for p near x_1 it would be too large. It is exactly matched by the positive term in (51).

For a full marking, $C_p(E_0(\Lambda)) = 0$ and $\omega_p = 1$. Combining (51), (52) and Lemma 5.10 gives

$$\|\mathsf{B}_\theta^\Psi(Y^m v)\|_\iota \le p^{2\ell_p - g_p}. \tag{53}$$

There is also a useful estimate with no marking at all. One may take $\Lambda = \varnothing$ in Lemma 5.10, with no nontrivial operators. Then (51) gives

$$\|\mathsf{B}_\theta^\Psi(v)\|_\iota \le p^{mC_p(\varnothing) + \ell_p - g_p}. \tag{54}$$

Writing

$$j_p = \left\lceil \frac{2mfa}{p} \right\rceil,$$

the estimate $C_p(\varnothing) \le 2fa/p$ and scalar linearity imply

$$\|\mathsf{B}_\theta^\Psi(p^{j_p} v)\|_\iota \le p^{\ell_p - g_p}. \tag{55}$$

These are the declarations `Data.IsGoodDatum.gaugeLe_Bmu_of_isMarked`, `Data.IsGoodDatum.gaugeLe_collisionSum` and `Data.IsGoodDatum.gaugeLe_pow_jetCeil_smul`, all in `New/Columns/LocalStatement.lean`. Thus the submodule

$$Y^m M + p^{j_p} M$$

has good divisibility for every admissible twist prime to p. In particular, its definition and the gain bound do not involve the twist. What remains is to prove that it has small index; this is a question about algebraic modular forms and the special Frobenius-twisted Hecke operator.

5.7 The Frobenius-twisted elementary-divisor theorem

We give the local theorem in its representation-theoretic form, because its application to the automorphic lattice is not valid unless both the operator and its coordinate normalization agree.

Fix a small served prime p, and a marked place l of E. Put $u = -p$ as above, $P_0 = p^{o_0}$, and let ν be the order of u^{o_0} in the CM place group. The tops inducing l or $\bar{l}$ are exactly the orbit of a chosen one under this element, so they may be labelled by $s \in \mathbb{Z}/\nu$. Since o_0 is odd, the parity of $l, \bar{l}$ alternates along the orbit. The residue field of l has cardinality $Q_l = P_0^\nu$ and ν is even. This uses the equality $q_l = p^{\mathrm{ord}(u)}$ in the served case, as in `Data.IsMarked.hasType_markOperator`. Here ν is a local orbit length and is not the auxiliary level prime N.

Write R_l for the valuation ring of E at l. There are embeddings $\iota_s : R_l \to O_\iota$ whose reductions satisfy

$$\iota_{s+1}(x) \equiv \iota_s(x)^{P_0} \pmod{p}.$$

In the odd letters the embedding in the coefficient representation also includes conjugation, so that the action of $g \in \mathrm{GL}_k(R_l)$ in the sth factor is by

$$g_s = \begin{cases} \iota_s(g), & s \text{ even}, \\ \iota_s((g^{-1})^t), & s \text{ odd}. \end{cases}$$

The weight module is $\mathcal{P} = \bigotimes_{s\in\mathbb{Z}/\nu} \mathrm{Sym}^a O_\iota^k$, represented by polynomials in variables x_{bs}, $0 \le b < k$, of degree a in each letter s. Let $\rho(g)$ denote the above representation. The other coefficient factors, not in this orbit, are left as a separate factor. Write $r_p = s_p$ for the size of the first block of coordinates. A variable is deep if either s is even and $b \ge r_p$, or s is odd and $b < r_p$. Denote the indicator of this condition by $\Delta(b, s)$. For a monomial x^α its depth is $\sum_{b,s} \Delta(b, s)\alpha_{bs}$. These are the exact definitions `TwistingData.letterRep`, `TwistingData.deepIndicator` and `TwistingData.depth` in `New/TwistedHecke/TwistingData.lean`. The dimension of $\mathcal{P}$ over the coefficient field is $D_\mathcal{P} = \binom{a+k-1}{k-1}^\nu$.

The following Euler- and weight-like notation is only local. Put $\lambda = \nu/2$ and $c_* = 2(k + k^2)o_0\nu^2 + 2$. The numerical conditions of the local theorem, with parameters ϵ, y, z, are

$$\begin{aligned}
&0 < 2r_p \le k, \quad \epsilon \le 1 \le y,\\
&\nu ak(\lambda + 2) + a + k < P_0^\lambda, \quad 6c_* + 2 \le y,\\
&2(\nu + 2)k/y \le \epsilon,\\
&(y + \nu(k + 1))kP_0^{3r_p} + k + 2 \le a/y,\\
&\nu a\left(P_0^{-2r_p} + \frac{r_p^2}{kP_0}\right) \le z.
\end{aligned} \tag{56}$$

This is `HeckeSizes` in `New/TwistedHecke/Counting.lean`. They allow the Frobenius degree to be large enough for interpolation while retaining a small elementary-divisor budget.

Why the normalized spherical operator has this weight. Write $d_p = \mathrm{diag}(p\mathbf{1}_{r_p}, \mathbf{1}_{k-r_p})$. In a single even letter, the contragredient coefficient action of d_p in the class-sum formula scales monomials by negative powers of p supported on the first block. Multiplying the action by p^a changes the power to the degree in the second block. At an odd letter the inverse transpose interchanges the two prescriptions. Thus multiplying the total Hecke operator by $p^{A_l} = p^{a\nu/2}$ gives precisely the diagonal matrix

$$D_p = \mathrm{diag}(p^{\mathrm{depth}(\alpha)}).$$

There is one factor p^a for every even letter. This explains the normalization of Y_l at the automorphic lattice, not just its complex normalization.

The root inequality. For distinct coordinates b, b' let $y_s = x_{b's}$, $z_s = x_{bs}$ if s is even, and interchange the choices if s is odd. For a fixed letter i define the excess functional

$$\mathrm{exc}_i(\alpha; b, b') = \frac{1}{Q_l}\sum_s \left(P_0^{1+[s-1-i]_\nu}\alpha(y_s) - P_0^{[s-i]_\nu}\alpha(z_s)\right), \tag{57}$$

where $[u]_\nu$ denotes the residue in $[0, \nu - 1]$. This is exactly the weight in `TwistingData.excessWeight`. Let $w \in O_\iota$ be a square root of p. For a set of real variable weights c_v, use the Gauss filtration: f is of order at least A if every coefficient of x^α is divisible by $w^{\max(0,\lceil A - \sum_v c_v\alpha_v\rceil)}$. This filtration is multiplicative: for every product of monomials the desired divisibility follows by adding the inequalities, then summing proves it for polynomials.

Lemma 5.11 (The elementary multiplicity count)**.** *If $\Phi \in O[T]$ is a nonzero polynomial, O a ring, and the images u_i of q distinct residue classes of a local ring have pairwise invertible differences in O, and $(T - u_i)^E \mid \Phi$ for all i, then $qE \le \deg\Phi$.*

Proof. The monic linear factors are pairwise comaximal and their powers are pairwise comaximal. Thus their product divides Φ. Multiplication by a monic polynomial adds its degree to the degree of a nonzero polynomial over any nonzero ring. Compare degrees. □

The precise root lemma of the local proof is stronger than a multiplicity count, and the following statement is the formal-source input `TwistingData.root_lemma` in `New/TwistedHecke/RootLemma.lean`: suppose b, b' are in the same block of a Levi subgroup and, for every letter s,

$$\frac{2}{5} \le c_{y_s} - c_{z_s} \le \frac{8}{5}.$$

Suppose every Levi translate of F is in the Gauss filtration of order $l + \sum_v c_v \alpha_v$, and $w^{l+1} \nmid F_\alpha$. Then $\mathrm{exc}_i(\alpha; b, b') \le 0$ for each i. One cannot obtain this directly by counting zeros of the individual coefficient polynomials of degree a, because the Levi action uses different Frobenius powers in different letters.

We describe the reduction in the formal proof. Choose the transvection from b to b' composed with interchange of the two coordinates, with parameter $t \in R_l$. At a letter s the parameter in the variable substitution reduces to a signed $P_0^{[s-i]_\nu}$-power of the parameter at i. The coefficient of a selected monomial is read after identifying the two variables in letter i, and substituting appropriate powers of an auxiliary variable T in the other letters. More precisely, the exponents assigned to z_s in the selected coefficient are $Q_l P_0^{[s-i]_\nu}$ and those recording the forced zeros from y_s are $P_0^{1+[s-1-i]_\nu}$. Modulo w^{l+1} the hypothesis on F_α gives a nonzero polynomial in T whose degree is at most

$$Q_l \sum_s P_0^{[s-i]_\nu} \alpha(z_s).$$

Its Taylor coefficients in degrees below $\sum_s P_0^{1+[s-1-i]_\nu} \alpha(y_s)$ vanish at every residue class t in R_l. The nonzero coefficient assertion in this calculation uses the minimum for the Gauss filtration: otherwise different monomials could contribute to the selected coefficient and invalidate the argument. The discarded substitutions differ by terms of higher Gauss order; this follows from the Frobenius congruence modulo $p = w^2$ and the two-sided bounds on the differences of variable weights. Thus the multiplicity count gives $Q_l \sum_s P_0^{1+[s-1-i]_\nu} \alpha(y_s) \le Q_l \sum_s P_0^{[s-i]_\nu} \alpha(z_s)$, which is the asserted inequality. This describes the coefficient calculation of the root lemma; the general-ring Gauss-filtration substitution estimates which justify its nonzero coefficient and all its divisibilities are part of that input. The essential gain is that each residue class contributes the same weighted multiplicity, and there are Q_l of them.

Two dimension estimates. Reduce coefficients modulo the maximal ideal and call the residue coefficient space $\overline{\mathcal{P}}$. Let $\mathcal{P}_{\mathrm{lev}}$ be the two-block parahoric of $\mathrm{GL}_k(R_l)$, namely the inverse image of the two-block parabolic over its residue field. Fix the adjacent coordinates $(r_p - 1, r_p)$. Define the subspace U_h of $\overline{\mathcal{P}}$ by requiring in every block-Levi translate that monomials of excess greater than h at this adjacent pair vanish, in every letter. Let V_j be the subspace of $\overline{\mathcal{P}}$ of forms such that every $\mathrm{GL}_k(R_l)$ translate has depth at least j. These are `TwistingData.excessCond` and `TwistingData.deepForms`.

The two relevant estimates are

$$[\mathrm{GL}_k(R_l) : \mathcal{P}_{\mathrm{lev}}] \dim U_y \le (1+\epsilon) D_{\mathcal{P}}, \tag{58}$$

$$\sum_{0 \le j < J} \dim V_j \le (1+z) D_{\mathcal{P}} \quad (J \ge 0). \tag{59}$$

These are `TwistingData.index_mul_finrank_excessCond_le` and `TwistingData.sum_finrank_deepForms_le`, in `New/TwistedHecke/Counting.lean`. We explain the combinatorics yielding these bounds. The uniform Frobenius minimum-weight calculations in their proof are additional formal-source inputs, specified below.

The first count uses the following cyclic interlacing cone. For a real array v_{bs} put $\beta(b,s) = \pm(P_0 v_{b,s+1} - v_{b,s})$, with sign positive in the even letters and negative in the odd letters. This is an invertible linear transformation: if $P_0 v_{b,s+1} = v_{b,s}$ for every letter, then cycling gives $(P_0^{\nu} - 1)v_{bs} = 0$. Substitution into (57), telescoping around the letters, gives

$$\mathrm{exc}_i(\beta; b, b') = (1 - Q_l^{-1})(\beta(y_i) - (v_{b,i} - v_{b',i})).$$

Thus the zero-excess inequalities are exactly the interlacing conditions

$$\begin{cases} P_0 v_{b+1,s+1} \le v_{b,s}, & s \text{ even}, \\ v_{b+1,s} \le P_0 v_{b,s+1}, & s \text{ odd}, \end{cases} \qquad 0 \le b < k-1.$$

Let $\Sigma(D)$ be the product of slabs $\beta_{bs} \ge 0$, $D_s - 1 \le \sum_b \beta_{bs} \le D_s$. The cone of zero excess, intersected with $\Sigma(D)$, has volume exactly

$$\frac{\mathrm{vol}(\Sigma(D))}{\prod_{j=1}^{k} \dfrac{Q_l^j - 1}{Q_l - 1}}. \tag{60}$$

Here is the signed-volume identity underpinning this formula. For decreasing real vectors x, y of length k, the matrix $\mathbf{1}_{y_c < x_b}$ has determinant 1 precisely when $x_1 > y_1 \ge x_2 > y_2 \ge \cdots \ge x_k > y_k$; otherwise it has determinant zero. This follows because the row of index b ends with a string of ones, preceded by zeros, and the nested rows are independent exactly when all their threshold positions differ. Take one such matrix of interlacing indicators for each pair of adjacent letters, using transpose in the alternating letters. Expanding the product of determinants yields a signed sum over tuples of permutations, whose product permutation will be denoted π'. The corresponding linear changes of variables $v \mapsto \beta$ have determinant, up to a common factor independent of π',

$$\det(1 - Q_l^{-1} A_{\pi'}),$$

where $A_{\pi'}$ is the permutation matrix. This follows by solving the equations around the cyclic chain after successively reordering the coordinates: the sole boundary condition is given by the product permutation. The walls where equalities occur have measure zero. Consequently integrating the signed indicators and accounting for sorting yields

$$\frac{\mathrm{vol}(\mathrm{cone} \cap \Sigma(D))}{\mathrm{vol}(\Sigma(D))} = \frac{(1 - Q_l^{-1})^k}{k!} \sum_{\pi' \in S_k} \frac{\mathrm{sgn}(\pi')}{|\det(1 - Q_l^{-1} A_{\pi'})|}.$$

The formula, including the cancellation of all intermediate permutation multiplicities, is `volume_inter_slab_eq_signed_sum` in `New/TwistedHecke/InterlacingCone.lean`. To evaluate the signed sum, it suffices to compare with the single-letter case with parameter Q_l instead of P_0. In that case the inequalities are $\beta_b \ge Q_l \beta_{b+1}$. The triangular change

$$\beta_b \longmapsto d_b(\beta_b - Q_l \beta_{b+1}), \qquad d_b = 1 + Q_l + \cdots + Q_l^b \quad (b = 0, \ldots, k-1)$$

with $\beta_k = 0$ carries the cone onto the nonnegative orthant. It preserves the sum of all coordinates because $d_b - Q_l d_{b-1} = 1$ and $d_0 = 1$; its determinant is $\prod_b d_b$. This gives (60). In the source this step is `prod_mul_volume_inter_slab`. The factor in the denominator is exactly the number of complete flags in a k-dimensional vector space over the residue field of size Q_l. This exact match explains why a leading constant $1 + \epsilon$ in (58) is possible.

The root lemma relates this volume to the dimension of U_y. To see the requisite amplification, one passes from a polynomial F to blocks of polynomials of the form $x^{\gamma(u)} F_u(x^D)$ with

$D = Q_l^J$, and $u \in \{0, \dots, D-1\}^{k\nu}$. The exponents $\gamma(u)$ are congruent to u modulo D and have degree in each letter between kD and $(k+1)D$. Hence these summands are independent, and substitution $x \mapsto x^D$ commutes with the residual group action because D is a full Frobenius period. One chooses a least-weight monomial among its translates over the flags in the Levi subgroup. The root lemma at adjacent roots within either block forces the corresponding excess inequalities, with the chosen bound y retained at the single boundary root between blocks. A count of its available coefficients and the amplification limit $D \to \infty$ bounds the dimension by the number of Levi flags times the volume of the thickened cone of boundary excess $y+\nu(k+1)$. This is the formal-source input `TwistingData.finrank_excessCond_le_volume`; it identifies the injection obtained by taking the relevant coefficients of the translates of the amplified polynomials.

To remove the thickening, the proof constructs a nonnegative vector σ with all coordinates at most $P_0^{3r_p}$, satisfying each zero-excess inequality and having boundary excess at most -1. One can see such a vector by taking $v_{b,s}$ proportional to $2P_0^{2(r_p-1-b)}$ for $b < r_p$ and zero for $b \geq r_p$, with the odd letters multiplied by P_0. Substituting into the linear map defining β and into the interlacing inequalities gives this bound. The assertion is `TwistingData.exists_mem_excessCone_mem_Icc`. Translation by $(y + \nu(k+2))\sigma$ embeds the thickened cone into the zero-excess cone; it changes the upper endpoint of each slab by at most $(y + \nu(k+2))kP_0^{3r_p}$, with allowance for the initial degree padding. The precise choice in the Lean proof uses $y + \nu(k+2)$ with each letter's shift and the size conditions (56). The volume of a slab of upper endpoint D_s is at most $D_s^{k-1}/(k-1)!$, as is seen by subtracting the volume of two simplices. With the size conditions the changed volume of all slabs is at most $(1+\epsilon)D_{\mathcal{P}}$ after multiplying by the flag count and canceling it against the denominator in (60). The exact inequalities including the roundings are those of `index_mul_finrank_excessCond_le`. They supply (58).

For the depth bound assign to each deep variable of coordinate $b \geq 2r_p$ cost $P_0^{-2r_p}$, and to each deep variable with $b < 2r_p$ cost $1/P_0$. All other variables have cost zero. The formal-source input `TwistingData.exists_le_weight_of_mem_deepForms` in `New/TwistedHecke/DeepFormCost.lean` says that any nonzero polynomial in V_j has a nonzero coefficient at a monomial of total cost at least j, provided $2r_p \leq k$ and $2r_p < \nu$. The source uses the upper unipotent matrix $1 + A$ with A supported in the block from the first r_p rows to the remaining rows and $A^2 = 0$. In the rational function field of the variables, the coefficients of A can be recovered by linear equations in Frobenius powers of the deep variables; evaluating each letter's variables in the specified order gives a triangular matrix with nonzero determinant. The equations express each matrix entry at residue points by polynomials of cost less than one. The condition of depth j in every translate then means vanishing to that weighted order on the whole family of residue points. Polynomial interpolation forces the existence of the claimed costly monomial. The recovery matrix and the weighted vanishing estimate are the substantive content of this input, rather than a generic vanishing-order assertion.

The rest of the depth argument is a direct count. Restriction to coefficients of cost at least j is injective on V_j by the stated input. Summing the resulting dimension bounds counts each monomial at most 1 plus its cost. There are $\nu/2$ deep letters for each fixed coordinate. Symmetry of the monomial basis gives average exponent a/k at each coordinate of each letter. Thus the mean cost over all monomials is at most

$$\frac{a}{k}\frac{\nu}{2}\left(2r_pP_0^{-1} + (k-2r_p)P_0^{-2r_p}\right) \leq \nu a\left(P_0^{-2r_p} + \frac{r_p^2}{kP_0}\right) \leq z.$$

This proves (59), using only the lower bound on cost in V_j and the last size condition in (56).

The approximate projector and the integral pairing. The dimension estimates must be converted into a determinant bound; they do not by themselves control Y_l. In particular, bounding the dimension of a conditioned representation modulo p does not give a bound on all Smith divisors of an arbitrary matrix with that representation in its target.

In the source, the approximation is constructed using an invariant diagonal pairing on degree-a polynomials. For a multi-index α of degree less than p^F, its multinomial coefficient has p-valuation at most (#variables)F. Indeed Legendre's formula for the factorial expresses the valuation as a sum of carry contributions at at most F levels, bounded by the number of input summands at each level. Consequently one may choose integral diagonal entries proportional to the inverse multinomial coefficients, preserving transpose covariance of the action. The exact formula is `ClassMean.exists_invariant_diagonal_form` in `New/TwistedHecke/ClassMeans.lean`.

A second input in that file, `ClassMean.exists_mean`, constructs an invariant linear form on polynomials of a generic Levi matrix. If x is a weight function on the matrix coordinates and P is a condition on the row weight, the linear form kills monomials whose row and column weights differ; on monomials satisfying P it agrees with evaluation at the identity multiplied by p^{k^2F}, up to the prescribed degree cutoff. It is in the span of Levi translates of linear forms killing monomials not satisfying P. The hypotheses are degree cutoff $D < p^F$ and $D + k < \#(R_l/\mathfrak{m}_l)$. For orientation, the finite-residue interpolation behind this fact is elementary: evaluation on the matrices with unit determinant spans the dual of the polynomial coefficient space in degrees D if $D + k$ is less than the residue cardinality. A polynomial vanishing on all such matrices, multiplied by their determinant, vanishes on the entire finite matrix space. The degree bound and successive one-variable root bounds force it to be zero. Lifting linear independence modulo the maximal ideal gives the integral statement on the finite free coefficient module. The invariant trace construction which incorporates the row-weight restriction is part of `exists_mean`.

Taking tensor products of these class means in the cyclic letters gives the approximate projector. In the Frobenius product of generic matrices the factor at the dth position is the entrywise P_0^d power of the generic matrix from letter $s - d$. At each letter one takes the first λ factors, plus one additional factor in each deep row; adjugate transpose takes the place of ordinary multiplication in odd letters. The construction tracks the Gauss weight and hence the excess of each output monomial. If the generic monomial has equal row and column weights outside a single letter, and the discrepancy between its weight at that letter and the output excess is greater than $2c_* + 1$, the coefficient of this monomial is divisible by w^{c_*}. This is the formal-source input `TwistingData.pow_dvd_coeff_conj_frobOperator` in `New/TwistedHecke/FrobeniusFactors.lean`; it follows from the multiplicative Gauss weight bound `TwistingData.frobOperator_mem` and the Frobenius congruence for conjugated generic matrices in the same file. That bound is consistent with the local scaling $p = w^2$.

The resulting precise assertion is `TwistingData.exists_projector`, in `New/TwistedHecke/ElementaryDivisors.lean`. There are a diagonal pairing matrix b_* and an operator Z_* such that Z_* maps every coefficient vector into the integral excess condition of bound y modulo w^{c_*}, and

$$y_0^t b_* Z_* F \equiv p^{(k+k^2)o_0\nu^2} y_0^t F \pmod{w^{c_*}}$$

for y_0 in the smaller excess condition with threshold $y - (4c_* + 2)$ modulo w^{c_*}. Here Z_* is not asserted to be an idempotent. The transpose compatibility $b_*\rho(\gamma^t) = \rho(\gamma)^t b_*$ is part of the statement. The precision $c_* = 2(k + k^2)o_0\nu^2 + 2$ leaves two extra powers of w beyond the scalar in the congruence. This pairing estimate is the nontrivial bridge between the counting result and the determinant calculation.

We describe the final formal-source determinant theorem, so that its applicability to the

coefficient lattice can be verified. In the following statement let X be a finite class set, I a coefficient index set identified with the monomial indices of $\mathcal{P}$ times a finite set H_0 of spectator indices (the remaining coefficient factors). Let $J = \mathrm{GL}_k(R_l)/\mathcal{P}_{\mathrm{lev}}$. Suppose for $e = (c, b) \in X \times J$ we have class $s(e) \in X$, matrix $g_e \in \mathrm{GL}_k(R_l)$ and integral invertible matrix u_e on the spectator indices, with

$$e \longmapsto (s(e), [(g_e^t)^{-1}]) \in X \times J \quad \text{injective.}$$

If γ_b are the chosen parahoric-coset representatives and $F = (F_c)_{c \in X}$ is a family of coefficients, consider

$$(TF)_c = \sum_{b \in J} \Big(u_e \otimes (\rho(g_e) D_p \rho(\gamma_b))^t \Big) F_{s(e)}.$$

With the orientation of representatives chosen in this formula, the theorem `TwistingData.hasType_twisted_hecke` states that on any linear image of the coordinate module to which T descends, the descended map preserves that module and has type bound

$$(z + \epsilon s + y)|X||I| \quad \text{modulo } p^s \tag{61}$$

when (56) holds. It is a formal-source input.

For its proof the map $F \mapsto (w^{\mathrm{depth}} \rho(\gamma_b) F_{s(e)})_e$ is split by a Hermite–Smith factorization into a primitive factor and a square factor. The determinant of the square factor divides a power of w with exponent at most the sum of dimensions of the depth spaces over the required truncation range. One way to read this bound is to reduce a column of the Smith factor modulo the maximal ideal: a column whose divisor exceeds w^j is forced into the jth depth condition in all frames. The independence of the reduced primitive columns bounds the count of such divisors by the dimension of that condition. This is `TwistingData.exists_depth_factorization`, which also ensures the primitive factor's coefficients lie in the integral excess condition with prescribed precision. Apply this factorization to each of the two half-depth factors of D_p (using $w^2 = p$). The half-factors are paired through Z_* so that the scalar $p^{(k+k^2)o_0\nu^2}$ controls the primitive part of the product, except for excess-condition vectors of total dimension bounded by (58). There is a second Smith factorization on the discrepancy matrix between the two primitive images. It gives the abstract pairing assertion `TwistingData.of_pairing`: if $RQ = 1$, $\beta M \equiv a_0 I \pmod{a_0 w}$, and the normalized difference $Qu - Mv = a_0 w F$ imposes the excess condition on F, then the image of βQ enlarged by p^s has index dividing $a_0^{|I'|} p^{sl}$, with $|I'| + l$ bounded by the dimension of the excess condition (including the class and coset multiplicities). Here I' is the domain index set of the primitive matrices. This is proved in the source by applying Smith form to $(QRM - M)R$, assessing each divisor relative to $a_0 w$, and using independence modulo the maximal ideal. The counts (58) and (59) now account for the three terms ϵs, z, and y in (61): the first is the exceptional share of the cokernel charged at depth s, the second is the aggregate half-depth cost, and the third pays for the controlled scalar in the projector. These contributions are additive by Lemma 5.2. This description specifies the remaining determinant theorem needed in the proof, including its hypotheses; proving the full matrix approximation independently would require the exact Smith-factor and projector identities, beyond the counting arguments already given here.

5.8 Application to the actual column lattice

The abstract coefficient operator above occurs on M itself, rather than on a putative subspace with convenient coefficients. This fact uses the trivial class stabilizers at K_0 and the availability of class representatives integral above p. The latter assertion is `exists_reps_integralAbove`, used in `New/Columns/TwistedHeckeType.lean`. For each class representative γ_c, write

the Hecke sum at the point γ_c using the local double-coset representatives indexed by $\mathrm{GL}_k(R_l)/\mathcal{P}_{\mathrm{lev}}$. Each product is rationally equivalent to one chosen representative $\gamma_{s(e)}$, times a compact-level element. Triviality of class automorphism groups gives injectivity of $e \mapsto (s(e), [(g_e^t)^{-1}])$: if the class and the parahoric coset agree, the two chosen rational identifications differ by an automorphism of the representative, which must be the identity, and the local cosets agree. The other primes above p give integral invertible matrices on the spectator factors. Thus in the coordinate identification of algebraic forms the normalized Hecke operator has exactly the shape of (61), including the depth diagonal. The exact comparison, with the class-representative and gauge conditions, is `Columns.hasType_markOperator` in `New/Columns/TwistedHeckeType.lean`. We use this coordinate comparison as a formal-source input. It is essential that the integral image addressed by that theorem is all of M, as opposed to the coefficient module with unrelated adelic transition functions.

The local congruence and numerical hypotheses are verified for our marked place by `Data.IsMarked.hasType_markOperator`, in `New/Columns/MarkedPrimeType.lean`. The precise result is the following.

Proposition 5.12 (Marked operator type)**.** *For each $l \in \Lambda$ and each integer $s \geq 0$, the operator Y_l preserves M and*

$$Y_l \in \mathrm{Type}\left(p, M, s; \left(\frac{128 A_l \log b}{mp} + \frac{s}{x^{14}} + x^{40}\right) c\right).$$

Indeed on the local orbit discussed above one has $A_l = a\nu/2$, the parity of the local embeddings is alternating, and the orbit size is even. The choices of the generic theorem's parameters are

$$\epsilon = x^{-14}, \qquad y = x^{40}, \qquad z = \frac{128 A_l \log b}{mp}.$$

The relation $P_0^{\lambda} \geq p^{m/2}$, from the divisibility of the order of u by m, supplies the interpolation inequality in (56). The served condition $o_0 \leq s_p$ gives $P_0 \leq p^{s_p} \leq pm$. The bounds $f \ll x^{11}(\log b)^{1/5}$ up to fixed factors, $k \asymp x^5$, $\nu \leq f$ and the hierarchy $x \ll x_1 \ll b$ leave room in the remaining inequalities; for example $P_0^{-2r_p} \leq m^{-2}$, and $r_p^2/(kP_0)$ is negligible against $\log b/(mp)$ on the admissible range. More precisely $P_0^{2r_p} \geq p^{2s_p} \geq mp$ since $P_0 \geq p$ and $m, p \leq p^{s_p}$. The large-prime range is covered by the separate estimate $s_p = 1$ when $p \geq m$. The stated Lean declaration checks these estimates with the fixed constants in (56), including the degree condition when ν and x_1 vary. Thus the proposition is a quantitative application of the coordinate comparison and the local theorem, not an assumption about Hecke eigenvalue slopes.

Corollary 5.13 (Type bound for the repair)**.** *With Y and j_p as above there exists an integer $e_p \geq 0$ such that*

$$I(M, Y^m M + p^{j_p} M; p^{e_p}), \qquad e_p \leq 3\ell_p c.$$

Proof. By Lemma 5.2 and the proposition, the exponent bound for the product of the marked operators, repeated m times, is at most

$$cm \sum_{l \in \Lambda} \left(\frac{128 A_l \log b}{mp} + \frac{j_p}{x^{14}} + x^{40}\right).$$

Now $\sum_l A_l \leq fa$, $|\Lambda| \leq f$, and $j_p \leq 2mfa/p + 1$. Thus the first contribution divided by $c\ell_p$ is at most $128/\log b$, the second at most

$$\frac{2m^2 f}{x^{14}(\log b)^2} + \frac{mp}{ax^{14}(\log b)^2},$$

and the third at most $mx^{40}p/(a(\log b)^2)$. All are at most 1 on the admissible scale, the second using the bound on f and $m^2 \le 4x$, and the third using $p < x_1$ and $x_1^{10^6} \le b$. The precise three-term comparison with the fixed constants is `Data.IsMarked.hasType_prod_markOperator_pow`. The type condition gives integrality and an integer witness e_p for the desired index. □

The crucial point is that the type estimate grows with the *truncated mean* of elementary divisors rather than with their largest possible value. Charging the whole space for one large elementary divisor would multiply ℓ_p by m and defeat the saving.

5.9 The prime N: moving denominators to a Fourier operator

The additional prime N splits completely in E and does not appear among the small nonsplit primes. Its effect in the complex bound is the large negative term in (46). At the finite places of N one must recover a usable lattice estimate despite the large negative powers of π in the restriction and the translations by y/N in the good datum.

Fix ι above N and let S be the set of conjugates of the primes induced by the given CM-type embeddings at ι (`Data.nonInducedPlaces`). At these split places use the coordinate embedding of the general linear group into the unitary group. Let d_j be the finite point whose component at the places in $S \cap \{v \mid \pi\}$ is $\mathrm{diag}(1, \dots, 1, N)$, with the corresponding conjugate component and identity elsewhere. These are `Data.diagPointN`. In the representation conjugated by d_j, the column level has mirabolic reduction, i.e. its last column is the final basis vector modulo N. The relevant level identity is `Data.atPlaces_conj_mem_mirabolic` in `New/Seesaw/LiftCharacters.lean`.

Let F_N be the Fourier operator on columns, given in this conjugated representation by multiplication by the inverse last-coordinate diagonal and a sum over integral matrices with prescribed final row and column, weighted by the additive character of their final residue coordinate. The formula in `IsSplitLevel.fourierOp` (`New/Seesaw/NamesAtSplitLevel.lean`) is

$$F_N = \rho(\mathrm{diag}(1, \dots, 1, N^{-1})) \sum_{y,\ y_{\mathrm{last}}\ \mathrm{unit}} \psi_N(y_{\mathrm{last}})\rho(\widetilde{g}_y).$$

All coordinates here are at the places of S, with the conjugated representation understood. Each $\widetilde{g}_y$ is an integral invertible lift of a matrix having final row and column determined by y. The additive character takes root-of-unity values. This operator preserves the column level (`Data.fourierOp_mem_column`) and is compatible with the ramification condition at N.

The key finite calculation is local and independent of the large weights. If L is a single theta-lift law, $\mathbf{b}$ is the indicator of the final-column condition and $\mathbf{c}$ the associated additive-character datum, then moving the diagonal off the character datum gives the Fourier operator applied to the indicator lift; it also equals a scalar times the indicator lift applied to the Fourier operator on the source. In the notation of `LiftLaw.u$_0$_charDatum_eq_fourierOp`, this is

$$\rho'(u_0)L[\mathbf{c}]v = F_N' L[\mathbf{b}]v = \gamma L[\mathbf{b}]F_N v.$$

The equality uses ramification: vectors in the image of $1 - z_v$ kill the terms in the finite character expansion which are invariant under that local scalar translation. One can consequently repeat the identity $m/2$ times to move the weighted Fourier operators from the paired chain onto the beginning of the column chain. This iteration is `compUpTo_pairs_eq_smul` in `New/Seesaw/NamesAtSplitLevel.lean`; its proof tracks the number of residue columns nonzero at every place and the scalar splitting-character factors at each stage. The full local identity, with these normalizations, is a formal-source input. Applying it to the paired

expression of the good datum yields `Data.chain$_2$_pairTranslates_eq_smul_chain`, `New/Columns/TwistedHeckeType.lean`: for every admissible twist, the paired chain of m kernels on a ramified column is a scalar of N-adic absolute value at most N^{2mfb} times a chain with data cut by $\mathbf{b}$, read in the frames d_j, acting on $F_N^{m/2}v$.

Each new cut kernel multiplies gauges by at most N^{fb}. This is `Data.name_lastColDatum_mem_raises`, same file; it is the crude CM integrality estimate applied to the cut datum rather than to arbitrary translates with denominator N. The branching restriction has the same gauge bound at N when its source gauge is computed at d_m: the negative powers of $\bar{\pi}$ are canceled by the corresponding final-coordinate powers in the branching formula. This is `Data.GaugeLe.restrictToY_of_diag` in `New/Seesaw/LiftCharacters.lean`. Both are used as precise formal-source inputs. Consequently, by Lemma 5.3,

$$\|\mathsf{B}_\theta^\Psi(v)\|_\iota \le N^{3mfb} \tag{62}$$

if $v \in C^{\mathrm{ram}}$ and $F_N^{m/2}v$ belongs to the column lattice measured at the frame d_0. This is `Data.gaugeLe_Bmu_at_N`.

The index cost of imposing this condition is much smaller than charging every coefficient its worst weight at the last coordinate. To see this, for each coefficient index I let $\alpha_I(\tau, b')$ be its tableau weight in coordinate b' at τ. For the initial column space these are just monomial exponents, and the sum over the coordinate b' in any factor is at most a: the factor is of degree a on the tops and degree zero elsewhere. With $\mathcal{X}$ the class set and $\mathcal{I}$ the Weyl indices put

$$T_{b'} = \sum_{c' \in \mathcal{X}} \sum_{I \in \mathcal{I}} \sum_{\tau \in \Phi} \alpha_I(\tau, b').$$

Coordinate symmetry of the initial weight implies $T_{b'}$ is independent of b', and the degree bound gives $\sum_{b'} T_{b'} \le fac$. In particular $T_{\mathrm{last}} \le fac/k$. The class-space action in the chosen coefficient frames is diagonal for the last-coordinate scaling: it multiplies a coefficient by $N^{\sum_\tau \alpha_I(\tau,\mathrm{last})}$, so its determinant is $N^{T_{\mathrm{last}}}$.

Precisely, if the column lattice at a frame $d_0 d(S')$ is changed to the lattice at d_0, where $d(S')$ is the inverse last-coordinate diagonal supported on a set of split places, the corresponding index divides $N^{T_{\mathrm{last}}}$. The adelic integrality here uses rational class representatives adjusted at N; the coefficient norm computation is a direct diagonal determinant comparison. The proof in the source pulls back the coordinate-diagonal index condition using the integral representative correspondence. Applying the Fourier operator consists of such a diagonal frame change and a sum of translations by integral invertible matrices; its character coefficients are N-adic units. Applying the pullback and transitivity rules of Lemma 5.2 once for each application of F_N, in addition to the initial frame change, gives an index bound dividing $N^{(m/2+1)T_{\mathrm{last}}}$. This is the assertion `Data.exists_indexDvd_fourierOp_pow` (`New/Seesaw/LiftCharacters.lean`), including the conditions on the changes of representatives. Since $a \le 2xb$ and $k \gg x^5$, the exponent is at most $4mfbc$ on the admissible data. Thus there is a submodule M'_N of M independent of the twist with

$$I(M, M'_N; N^e), \quad e \le 4mfbc,$$

on which (62) holds for all ramified columns. The presence of the average tableau weight rather than the largest coefficient degree is another place where symmetry of the integral weight representation is important.

5.10 The simultaneous local divisibility statement

Define a function on rational primes by

$$\epsilon_p = \begin{cases} 3\ell_p, & p < x_1,\ \chi_K(p) \neq 1, \\ 4mfb, & p = N, \\ 0, & \text{otherwise.} \end{cases}$$

Extend g_p by zero off the small nonsplit primes. For p outside those primes and $p \neq N$ the elementary integrality bound from the theta data suffices: every translated kernel is standard with integral gauge away from the twist conductor, and its scale $\bar{\pi}_g$per compensates for the denominators. There is no division by d, since in that case $p \neq d$. The resulting all-column estimate $\|\mathsf{B}_\theta^\Psi v\|_\iota \leq 1$ is `Data.IsGoodDatum.gaugeLe_one` in `New/Columns/LocalStatement.lean`. It is the application of the crude jet bound and Lemma 5.3 to the exact unmodified data.

Proposition 5.14 (Local divisibility and index)**.** *For every p, every ι above p, and every integer $\beta < g_p - \epsilon_p + 1$, there is a submodule M'_p of M and an integer $e \geq 0$ such that*

$$I(M, M'_p; p^e), \qquad e \leq \epsilon_p c.$$

For every admissible twist θ prime to p, every good datum Ψ and every $v \in M'_p \cap C^{\mathrm{ram}}$, the restricted lift $p^{-\beta}\mathsf{B}_\theta^\Psi v$ is integral at ι.

Proof. For a served small nonsplit prime take $M'_p = Y^m M + p^{j_p} M$ and the index supplied by Corollary 5.13. The two terms have gauges at most $p^{2\ell_p - g_p}$ and $p^{\ell_p - g_p}$ respectively by (53) and (55); the same bound $p^{2\ell_p - g_p}$ holds for their sum. Since $\ell_p \geq 1$, the condition on β gives $\beta + 2\ell_p - g_p \leq 0$, so multiplication by $p^{-\beta}$ makes the gauge at most one. For an unserved small nonsplit prime use $M'_p = M$, the collision bound Lemma 5.7, and (54). At N use the preceding frame-change index and (62); the condition $\beta < 1 - 4mfb$ ensures $\beta + 3mfb \leq 0$ since $mfb \geq 1$. At the other primes integrality is unconditional and the hypothesis $\beta < 1$ gives $\beta \leq 0$. This proves the proposition, including the common choice of submodule before the twist. It is exactly `Data.exists_indexDvd_Bmu_divisible`. □

The loss function is summable with the following quantitative bound:

$$\sum_p \epsilon_p \log p \leq B. \tag{63}$$

Here is a direct calculation, with generous constants for the scale comparison. If $p \equiv 1 \pmod m$ is prime, $p = um + 1$ with $u \geq 1$, so $m/p \leq 1/u$. If $p \equiv -1 \pmod m$, then $p = um - 1$ with $u \geq 1$, so $m/p \leq 2/u$ since $m \geq 2$. Dropping the primality conditions and bounding $\log p$ by $\log x_1$, one obtains

$$\sum_{p < x_1} \frac{\log p}{p} \begin{cases} m, & p \text{ short}, \\ 1, & \text{otherwise} \end{cases} \leq 4 \log x_1 (1 + \log x_1).$$

Thus the small-prime contribution is bounded by $12fa(\log b)^2 \log x_1(1 + \log x_1)$. Since $(\log x_1)^6 \leq \log b$ and $\log b$ is large, this is less than $B/2$. At N, one has $4mfb \log N \leq 120fm^2 b \log b \leq 480fa \log b$, using $m^2 \leq 4x$ and $a \geq xb$. This too is less than $B/2$ for the data in question. The statement with the precise admissibility thresholds is `Data.sum_allowedLoss_mul_log_le` in `New/Columns/SizeEstimates.lean`. It is the sparseness of the congruences for short primes which allows the factor m in ℓ_p and makes the global error budget independent of an overall factor m.

5.11 Archimedean growth and the level normalization

We now compare the finite-prime gains with the growth at every complex conjugate. Define

$$h_0 = m\left(\sum_{\tau\in\mathcal{T}}(J(a)-A(a,\tau))+k\mathcal{U}\right)-\#\mathcal{T}\,a\log N.$$

Since $A(y,\tau)=2\sum_p A_p(y,\tau)\log p$, it follows identically from the definition of $\mathcal{U}$ and g_p that

$$h_0+2\sum_p g_p\log p = m\left(\sum_{\tau\in\mathcal{T}}J(a)+k\sum_{\tau\in\Phi}J(w_0(\tau))\right)-\#\mathcal{T}\,a\log N. \tag{64}$$

This quantity is the leading exponent of the purely archimedean chain. In particular its expression in terms of prime asks is not a second analytic estimate; it is an exact algebraic identity, used to record the finite gain to be divided out.

The precise norm assertion of the definite chain is `Data.weylNormSq_Bmu_le`, `New/Columns/LocalStatement.lean`. For the trivial twist, good datum Ψ, and any conjugate σ, if all values of a column v have squared Weyl norm at most e^{η}, all values of its restricted lift obey

$$\|\sigma\mathsf{B}_1^{\Psi}v(g')\|^2\le\exp\left(\eta+h_0+2\sum_p g_p\log p+B\right). \tag{65}$$

Here the norm assertion is conditional on the conjugate-jet bounds of the preceding section. We explain the size comparison to isolate the exact quantities going into its error budget.

For one lift the principal contribution from the Young diagram is the sum of $J(l)$ over its row lengths. The uniform diagram estimate `Data.diagramWeight_le_by_index` gives each row length at place τ the bound

$$\begin{cases}\kappa_\tau(1)+b_1+b_2+h_2^2, & \text{row zero},\\ w_0(\tau)+h_2^2, & \text{rows } 1,\dots,k-1,\\ h_2^2, & \text{remaining rows.}\end{cases}$$

Positive error terms in the first row can be regarded as additional degree beyond the a-row (at a top) or the degree-zero row. For any $l\le r'+e$, one has the elementary bound

$$J(l)\le J(r')+e\big(\log(r'+e+1)-\log(\Omega/g)\big).$$

Indeed $l!\le(r'+e)!$ and $(r'+e)!\le r'!(r'+e+1)^e$; the exponential scale $\Omega/g\le 1$ on the data. This gives `Data.jetSize_le_add` in `New/Columns/SizeEstimates.lean`. Using it for the rows of the diagram gives an upper bound with leading contribution $\sum_{\tau\in\mathcal{T}}J(a)+k\sum_{\tau\in\Phi}J(w_0(\tau))$. All lower-order row contributions, conjugate-jet errors and the difference between actual and nominal placement are bounded by a polynomial in the ranks and f times $b\log b$ or by similar smaller expressions. The source's size calculation following `diagramWeight_le_by_index` checks the corresponding inequalities with the fixed scale constants. The other terms in the complex triangle inequality of Lemma 5.3 are tightly bounded: the number of classes at each stage is at most $e^{x^{30}}$ (a uniform bound `Data.card_stdClasses_le_exp`, same file), the sum of absolute values of coefficients of the good datum is at most $e^{x^{60}}$, and the final branching produces the large negative exponent in (46), with error at most $x^{70}+2fb\log N$. Since $\log N\ll m\log b$, the latter error is bounded within the budget. Combining the per-lift factorial estimates with this branching calculation gives the right side of (64), plus at most B of error. The exact inequality (65), including the conjugation uniformity and class-count normalization, is a formal-source input at the quantitative scale being used.

To see how the level N is chosen, retain the notation ℓ_{CS} for the Chowla–Selberg constant and put

$$\overline{w} = \frac{1}{f}\sum_{\tau} \log w(\tau), \quad G_0 = \log(g\,\mathrm{rd}_E/(2\pi)), \quad M_0 = a(\log a - 1 - \ell_{\mathrm{CS}} - \overline{w}).$$

Thus $\Omega = e^{\ell_{\mathrm{CS}}} 2\pi/\mathrm{rd}_E$, $\mathcal{R} = \sum_{\tau\in\mathcal{T}}(A(a,\tau) - M_0)$ and $J(a) = \log(a!) - a\log(\Omega/g)$. The elementary Stirling upper bound for $a \geq 1$ is

$$\log(a!) \leq a\log a - a + 1 + \tfrac{1}{2}\log a.$$

It follows from the monotonicity of the Stirling remainder with the constant fixed at $a = 1$, or by the standard Stirling bound with positive explicit error. Substituting in $J(a) - M_0$ gives

$$J(a) - M_0 \leq a(\overline{w} + G_0) + 1 + \tfrac{1}{2}\log a.$$

By construction of the level window

$$\log N \geq m(\overline{w} + G_0) - G_0 - \tfrac{15}{2}\log b.$$

Consequently

$$\begin{aligned} h_0 &= H - m\mathcal{R} + \#\mathcal{T}\{m(J(a) - M_0) - a\log N\} \\ &\leq H - m\mathcal{R} + \#\mathcal{T}\left\{m(1 + \tfrac{1}{2}\log a) + a(G_0 + \tfrac{15}{2}\log b)\right\} \\ &\leq H - m\mathcal{R} + B. \end{aligned} \tag{66}$$

The last estimate uses $\#\mathcal{T} \leq f$, $G_0 = O(\log b)$ and $m \leq 2\sqrt{x} \ll a$, with the large $\log b$ admissibility margin. This is precisely the calculation in `Data.heightPerRow_le` (`New/Endgame/ShortAndEnough.lean`). The level is chosen to remove the large contribution $m\#\mathcal{T}a(\overline{w} + G_0)$, leaving the prime remainder itself.

5.12 Projection, product formula scaling and Theorem A

The good projection e_0 on the finite level quotient G is normalized by $|G|$. Write $P_s = |G|e_0 sP(z_0\cdot)$ with z_0 from Theorem 5.4. The equality at the level of lifts is translation by finite compact elements followed by the same group-ring projection: the value of the lift of $|G|e_0 sz$ at g' is

$$\sum_{u\in G} (|G|e_0(u))\,(\Theta_\theta z)(g'\widetilde{us}).$$

This is `Data.iota_goodProjTranslate_apply` (`New/Rows/MonomialCoefficientVector.lean`). The coefficient $|G|e_0(u)$ is an algebraic integer: the group-ring central idempotents are projections onto specified irreducible types and their coefficients are characters multiplied by dimensions, divided by $|G|$. Finite group characters take algebraic-integer values. Thus the normalized projection preserves integrality at every finite place, including twists prime to that place. This yields `Data.goodProjTranslate_mem_integralForms`. For the trivial twist its complex norm contributes at most $e^{x^{99}}$ times the previous squared-value norm bound when converted to the squared class-sum norm. Indeed the absolute values of the coefficients of $|G|e_0$ are bounded in terms of the finite group order, and the number of classes is bounded as above. The statement with exact constant is `Data.normSqAt_goodProjTranslate_le`, same file.

Theorem 5.15 (A: short forms)**.** *There is a positive rational q such that, with $Q_s = qP_s$, the everywhere-integral columns $u \in C^{\mathrm{ram}}$ for which*

$$\|\Theta_1 Q_s u\|_\sigma^2 \le \exp(H - m\mathcal{R} + L) \qquad (s \in G,\ \sigma \in \mathrm{Aut}(\mathbb{C}))$$

span C^{ram}. For every finite $J \subset G$ and every finite family y of elements of the images $Q_s(C^{\mathrm{ram}}_{\mathrm{int}})$, $s \in J$, the logarithmic index of y in the twisted integral structures is at most

$$|J|cB.$$

Proof. Take the spanning set of (47). Put $h_p = \lfloor \epsilon_p - g_p \rfloor$ and $q = \prod_p p^{h_p}$, with the product finite by support of the loss and gain functions. In particular

$$\log q = \sum_p h_p \log p, qquad |\iota(q)|_p = p^{-h_p}.$$

For each p the integer $-h_p = \lceil g_p - \epsilon_p \rceil$ satisfies the strict hypothesis of Proposition 5.14. Therefore on M'_p the factor q times the restricted lift is integral for every admissible twist prime to p. The upper bound (65) with $\eta = B$ now gives

$$\begin{aligned} \log \|\Theta_1 Q_s u\|_\sigma^2 &\le 2\log q + h_0 + 2\sum_p g_p \log p + 2B + x^{99} \\ &\le h_0 + 2\sum_p \epsilon_p \log p + 2B + x^{99} \\ &\le H - m\mathcal{R} + 5B + x^{99} \le H - m\mathcal{R} + L. \end{aligned}$$

The last inequality uses the numerical hypothesis of the data, $L = B \log b$ and x^{99} negligible compared to B. A zero norm already satisfies the desired bound; the logarithmic calculation is applied to nonzero norms.

For the index, fix p and ι. On $M \cap C^{\mathrm{ram}}$ pull back the condition $I(M, M'_p; p^e)$ of Proposition 5.14 by inclusion; thus its determinant bound is unchanged. Under every particular Q_s the restricted submodule has integral lifted rows by the seesaw and the normalized projection. If M'' is the inverse image of the integral lifted rows under Θ_θ, this gives $I(M \cap C^{\mathrm{ram}}, Q_s^{-1}(M''); p^e)$. The last part of Lemma 5.2 therefore shows that any family from the specified union of $|J|$ images has a square coefficient matrix over O_ι turning its lifts into integral forms, with determinant dividing $p^{|J|e}$. In particular, its p-adic absolute value is at least $\exp(-|J|c\epsilon_p \log p)$. The numerical costs are independent of ι and of the twist, though the matrices need not be. This exactly meets the definition of logarithmic index in the row space: there is one function of p bounding the costs for every embedding and every twist prime to p. Finally (63) gives an upper bound $|J|cB$ for the sum of the costs over every finite set of primes. This proves the second part and completes the proof from the stated inputs. The source declaration is `Data.short_forms`, in `New/Endgame/ShortAndEnough.lean`. □

We emphasize which part of the mathematical reconstruction is still an input. The theorem's last deduction uses only the explicitly stated local divisibility, conjugate-size and index bounds, and their proofs as far as given above. The deepest local subtheorems are the CM jet normal form and Schur bound of the preceding section, the genuine seesaw with uniformly integral finite data, and the Frobenius-twisted Hecke divisor theorem `TwistingData.hasType_twisted_hecke`, with its root inequality and approximate projector. The presentation above specifies the actual versions needed, including the level and gauge compatibilities. They are formal-source inputs, not consequences merely of the usual theta correspondence or Satake parameter formulas.

6 Quantitative spanning by projected Bessel–Poincaré series

The preceding bound for the columns is useful only if the resulting rows span a sufficiently large space. A merely qualitative nonvanishing or spanning theorem does not suffice: the integral-index bound is proportional to the number of translates used. We give the quantitative statement, develop the relative trace calculation on which it rests, and prove the finite-dimensional deduction. This section separates out the analytic and arithmetic estimates that are specialized to the large ranks, levels and weights of this proof. The term *Formal-source input* below denotes a result extracted from the indicated formal declaration; it is not asserted to have been independently reproved here. We retain the actual constants when they are used in another estimate. The Hilbert space arguments and the finite-field sieve deductions are given directly.

6.1 Statement and finite-dimensional framework

Let C denote the column space of algebraic modular forms, C^{ram} its subspace ramified at the chosen auxiliary prime, W the row space, and $\mathfrak{G}$ the finite group of level translations. Put $c = \dim_{\mathbb{C}} C$ and $n = \dim_{\mathbb{C}} W$. The maps $P_s : C^{\mathrm{ram}} \to W$ are the projected translates defined above, before the common rational rescaling of Theorem A. This rescaling has no effect on the image of any map.

Theorem 6.1 (Quantitative spanning, Theorem B)**.** *For admissible data $\mathcal{D}$ we have $c, n > 0$. There exists a finite subset $T \subset \mathfrak{G}$ for which*

$$|T|c \le 5n\log(1+b^2), \qquad \dim\Big(\sum_{s\in T} P_s(C^{\mathrm{ram}})\Big) \ge (1-4b^{-2})n.$$

This is `Data.enough_forms` in `New/Endgame/ShortAndEnough.lean`. We will deduce it from a local-to-global moment calculation and a finite-dimensional variance inequality, rather than take `Data.enough_forms` as an input. Notice the first inequality uses c, the dimension of the entire column space: the actual index estimate of A costs at most cB per translate. It is not enough to bound $|T| \dim C^{\mathrm{ram}}$.

For an arbitrary subspace $J \subset W$, with fraction $\alpha = \dim J/n$, the operative intermediate result is

$$\dim(P_s D_N)^{-1}(J) \le ((1+\alpha)/2 + b^{-2})c \quad \text{for some } s \in \mathfrak{G}. \tag{67}$$

Here $D_N : C \to C^{\mathrm{ram}}$ is the ramifying operator of the previous section, viewed as a map onto its image. When J has dimension appreciably less than n, the dimension of the new image modulo J is correspondingly large. There are two analytical tasks: construct the original Poincaré map on C, and control its translates after discarding bad finite types. The ramification operator imposes only a small additional cost on the source rank.

6.2 Finite level, central projection, and the transfer bound

Write E^+ for the maximal totally real subfield of E. The row forms before projection belong to a finite-dimensional Hilbert space V of cusp forms on $U(Y_0)$ of the prescribed archimedean type and of principal level at the chosen split places of g_0, g_1. Its inner product is the Petersson pairing. The group of finite level translations is

$$\mathfrak{G} = \mathfrak{G}_0 \times \mathfrak{G}_1, \qquad \mathfrak{G}_i = \prod_{v \mid g_i \text{ chosen}} \mathrm{GL}_r(\mathbb{F}_v),$$

acting via a unitary representation ρ. In each product there are precisely f factors with residue cardinality g_i. Conjugate places are omitted from the indexing because the local unitary condition recovers the data at the conjugates. For an irreducible complex character ξ of a finite group G the element

$$e_\xi = \frac{\xi(1)}{|G|} \sum_{g \in G} \overline{\xi(g)}[g]$$

is a self-adjoint central idempotent. If $e = \sum_{\xi \in F_0} e_\xi$ for a set of irreducible characters, its coefficient at the identity is the Plancherel mass $\mu(F_0) = \sum_{\xi \in F_0} \xi(1)^2/|G|$. Consequently $\mu(F_0)$ is the fraction of the regular representation occupied by these isotypic spaces; it need not be the analogous fraction of V. Distinguishing these assertions is essential in the filling proof, which bounds the trace of the discarded part on the particular image of the Poincaré series, rather than deducing it from the discarded proportion in an arbitrary representation.

At a chosen place with residue field $\mathbb{F}_q$ the regular labels used in the proof are monic squarefree polynomials of degree r with nonzero constant term. Let M_ω denote the excluded torsion order, as in the definition of the level characters. A label is allowed if it has no root $x \in \mathbb{F}_q$ of $x^{M_\omega} = 1$. A primitive label is an irreducible polynomial with one (and hence all) roots of multiplicative order $q^r - 1$. The good tuples are tuples of allowed labels for which at least one label is primitive. The definitions are `IsPrimitiveLabel` and `goodTuples` in `New/Trace/BadShare.lean`; the corresponding idempotents are `labelIdem`, `tupleIdem` and `Data.e0` in the finite character theory and `New/Seesaw/IndexTypes.lean`. Let e_i be the good central projection in $\mathbb{C}[\mathfrak{G}_i]$ and $R = \rho(e_0 \otimes e_1)$. (Here e_0 denotes the first factor's projection; the code uses `Data.e0` for their product.) Our row space is $W = \operatorname{im} R$. The two factors are both used: the goodness requirement contains a primitive label above *each* g_i.

Lemma 6.2 (Plancherel deficit)**.** *Let $1 < r < q$ and $M_\omega \geq 1$. Suppose $0 \leq \varepsilon \leq 1$ and $L_0 \geq 0$ satisfy*

$$(1 + M_\omega + 2r)/q \leq \varepsilon, \qquad 2r(2r \log q + 1)L_0 \leq f.$$

Then the coefficient at the identity of $1 - e_i$ is at most $f\varepsilon + e^{-L_0}$. In particular, for our data, the complementary coefficient is at most $b^{-8}/2$ for each i.

Proof. The representation-theoretic input is the degree estimate of `labelChar_degree_bounds` for the Green characters, used in `bad_share_le`: each regular label contributes Plancherel mass at least

$$\mu_0 = (1 - r/q)q^{-r}.$$

We take this degree formula as the *Formal-source input* from the finite general linear group character theory (see `New/Trace/BadShare.lean` for the use of this bound). The rest is counting. There are q^r monic polynomials of degree r. For each of at most $1 + M_\omega$ prohibited linear roots (including zero), there are q^{r-1} polynomials divisible by the corresponding linear factor. The nonsquarefree polynomials belong to the union of the families Q^2P, where Q, P are monic and $1 \leq \deg Q \leq r/2$, $\deg P = r - 2\deg Q$. For fixed degree j there are at most q^{r-j} choices, so a crude bound for the total is rq^{r-1}. Hence allowed labels have total uniform lower mass at least

$$(1 - r/q)(1 - (1 + M_\omega + r)/q) \geq 1 - \varepsilon.$$

Primitive roots are precisely the $\varphi(q^r - 1)$ generators of the multiplicative group of the degree-r field. Such a generator is not contained in a proper subfield, so each of their minimal polynomials is irreducible of degree r and corresponds to exactly r generators. The number of primitive labels is therefore $\varphi(q^r - 1)/r$. The elementary totient lower bound used in the source is $\varphi(u) \geq u/(2\log u + 1)$ for $u \geq 2$ (`le_totient_mul_log`). Thus the uniform mass q^{-r} assigned to these labels is at least $1/[2r(2r \log q + 1)]$. The mass of nonprimitive labels

with respect to the weight μ_0 is at most their mass with respect to q^{-r}, whence it is at most $1-L_0/f$. On the product of f factors, the mass lower estimate for all allowed labels is $(1-\varepsilon)^f$, and the uniform lower mass of the allowed tuples that are not primitive anywhere is bounded above by $(1-L_0/f)^f$. Thus the mass of the good tuples is at least $(1-\varepsilon)^f-(1-L_0/f)^f$. Subtracting from one, and using Bernoulli's inequality and $(1-u)^f \le e^{-fu}$, proves the claimed bound.

For the last specialization one takes $L_0 = 9\log b$. The size inequalities used in `Data.ramifiedColumns_fill_rows`, together with $g_i \ge b^{10}$ and f sufficiently smaller than the level, give $f\varepsilon + b^{-9} \le b^{-8}/2$. These exact numerical inequalities on the data are the *Formal-source input* `bad_share_le` as instantiated in `New/Filling/FillingTheorem.lean`. The counting proof explains the two separate contributions: excluded labels are individually rare, while failure to contain even one primitive label is exponentially rare across the many factors. □

The second input is a bound on the actual number of target rows. Let $d_\infty = \prod_{v\in\Phi} d_v$ be the product of the local formal degrees (defined through the Bessel systems below) and let $\mathrm{mass}_r(E)$ be the mass of the rank-r positive definite unimodular genus, with finite stabilizer weights. Define

$$n_0 = |\mathfrak{G}| d_\infty \,\mathrm{mass}_r(E). \tag{68}$$

Proposition 6.3 (Formal-source input: dimension transfer)**.** *One has $n \le n_0$. More precisely the trace of every element of the common split unramified Hecke algebra on good forms of the same rank, level and compatible archimedean Harish–Chandra data is independent of the unitary form, in the rigid-label situation of this proof.*

The statements are `trace_heckeOnGood_eq`, `det_heckeOnGood_eq` and `finrank_rows_eq_finrank_goodForms` in `New/Trace/GoodFormTransfer.lean`, and `Data.finrank_rows_le_mass` in `New/Filling/RowDimension.lean`. For the dimension computation, the primitive-label conditions and the local rigidity lemma in `New/Trace/LabelRigidity.lean` are needed in the trace formula, as well as the existence of projectors onto the finite Hecke eigenspaces. The proof transfers the row forms to the definite member of rank r, where good forms are a subspace of the full space of algebraic modular forms. Here the class stabilizers are trivial at the chosen level. The class sum weighted by the reciprocal group orders is $|\mathfrak{G}|\,\mathrm{mass}_r(E)$ by the level-index identity; the dimension of the finite-dimensional algebraic representation at infinity is d_∞ in the transfer normalization. Multiplying gives (68). The nontrivial passage between signatures is not a formal consequence of dimension equality and cannot be omitted; the indicated trace-transfer theorem is the imported assertion. The same trace equality is used again to compare auxiliary lifts with the main chain in the lower-bound argument.

6.3 Hermitian geometry of the Bessel subgroup

We describe the objects for the unfolding calculation. In this section use $h = m/2-1$ and $s_0 = h+1$ (not the dimension n of row forms). The admissible rank $r = k+2h+1$ is the rank of the split extension of the definite rank-k space X_0. It has ordered basis of h hyperbolic pairs e_i, e_i^* $(0 \le i < h)$, an orthonormal basis of X_0, and one middle vector z with norm $-N$. Thus the Gram matrix in the rational basis is the sum of the hyperbolic planes, I_k, and $(-N)$. Denote this form by Y. Let U be the unipotent radical of the stabilizer of the isotropic flag spanned by successive initial e_i, with its standard Bessel character, and let H_0 be the subgroup isomorphic to $U(X_0)$ acting as the identity on the frame and its dual. The Bessel subgroup is $S_B = UH_0$. Its frame is $F = (e_0, \dots, e_{h-1}, z)$; U^0 is the pointwise frame stabilizer in U. The coordinate form of these groups, including

the flag, is the `Shape` construction of `New/Filling/EichlerFlag.lean` and the groups in `Lib/Adelic/Poincare/BesselSubgroup.lean`.

At finite places put $N = \pi\bar{\pi}$, interpreted as an idelic factorization if necessary, and modify the basis of $X_0 \oplus Ez$ by replacing its selected vector e_* with $v_* = e_*/\bar{\pi} + z/N$. Its self-pairing is $1/N - 1/N = 0$ and its pairing with z is -1; the 2×2 middle Gram block becomes

$$\begin{pmatrix} 0 & -1 \\ -1 & -N \end{pmatrix}$$

with determinant -1. All other middle basis vectors retain their original pairings; the hyperbolic planes are unchanged. Thus the lattice Gram matrix is unimodular. These are the formulas of `latMiddleGram` and `gram_besselLatticeBasis` in `New/Filling/EichlerFlag.lean`. Let K_f be the stabilizer of this lattice away from the infinite places, and $K(g)$ the kernel of reduction to the above finite group. The column level is the inverse image of $K(g)$ under inclusion of H_0; this is one of the hypotheses in the filling record. Let $\mathcal{C} = H_0(E^+)\backslash H_0(\mathbb{A}_f)/K_c$ and h_u be the chosen images of representatives of classes $u \in \mathcal{C}$.

Proposition 6.4 (Formal-source input: global filling datum)**.** *The admissible data give a filling system in the sense of* ***`Filling`*** *and* ***`Filling.IsAdmissible`*** *of* ***`New/Filling/Setup.lean`****, with the prescribed r, k, h, N, g_i, b, whose column stabilizers are trivial, reduction onto $\mathfrak{G}$ is surjective, and whose Bessel coordinate spaces at each infinite place identify with the column weight representation. The resulting column coordinates give an isomorphism*

$$C \simeq \bigoplus_{u \in \mathcal{C}} \mathbb{C}^I, \qquad I = \prod_{v \in \Phi} I_v.$$

In particular $c = |\mathcal{C}||I| > 0$.

The coordinate identification is `Data.columnCoords` in `New/Seesaw/BesselGeneric.lean` (the coordinate model is built in `New/Filling/ColumnCoords.lean`); the filling construction is `New/Filling/FillingData.lean` and the admissibility check is `New/Filling/FillingAdmissible.lean`. Positivity follows since the class of the identity and every weight coordinate set are nonempty. The dimension formula also follows directly from the double-coset description when stabilizers are trivial. For the local-to-global estimates below we record several of the other numerical hypotheses, rather than replace them by unspecified implied constants. If ϖ is the purely imaginary generator of the absolute different and $\kappa_v = \Re(gv(\varpi)/(2\pi i))$, then

$$1 \le |\kappa_v| \le b^{23}, \quad |v(\varpi)| \le b, \quad N \le b^{60 s_0}, \quad b^{10} \le g_i \le b^{11}.$$

Also $h \ge 10^5$, $1000(2h+2)^9 \le k \le f$, $s_0^7 \le k$ and $(200 f k \operatorname{rd}_E)^{1000} \le b$. These are much stronger than the hypotheses of the elementary rank argument. They enter below in estimates which would be false uniformly for arbitrary families of lattices of varying dimension.

The normalized archimedean basis at v is scaled by $\kappa_v^{h-i}/\sqrt{N}$ on e_i, its inverse on e_i^*, and $1/\sqrt{N}$ on z. On X_0 the basis is unchanged. The middle norm is thereby normalized to -1. Put

$$l_{vi} = \frac{\kappa_v^{2(h-i)}}{N} \quad (0 \le i \le h), \qquad \mathcal{L} = \prod_{v \in \Phi} \prod_{i=0}^{h} l_{vi}^{k+2(h-i)}.$$

For a rational element γ of $U(Y)$ define the frame-product matrix $Z(\gamma) = (\langle F_i, \gamma F_j \rangle)_{0 \le i,j \le h}$, and write $D_i(\gamma)$ for its leading principal minor of order i. The discriminant separating the open and smaller cells is

$$\mathfrak{d}(\gamma) = D_{h+1}\overline{D_{h+1}} - N^2 D_h \overline{D_h}. \tag{69}$$

The normalized archimedean matrix Z_v has entries $\sqrt{l_{vi}l_{vj}}\, v(Z(\gamma)_{ij})$ up to signs of the real scales (the scales themselves are chosen real, not assumed positive); its normalized discriminant is $|\det Z_v|^2 - |D_h(Z_v)|^2$. Direct expansion shows that it equals $(\prod_i l_{vi})^2 v(\mathfrak{d}(\gamma))$, since the ratio between the square of the products of the full scalings and of the first h scalings is $l_{vh}^2 = N^{-2}$. The definitions are `Filling.frameProducts`, `frameScaleSq`, `termDisc` in `Setup.lean`, and `frameDisc` in `EichlerFlag.lean`. The nonnegativity of this discriminant for actual frame products and the strict inequality in the full open cell are geometric consequences of the positive definiteness of the complement of the hyperbolic planes, recorded in `NearTerms.lean`.

This particular frame is relevant both to the archimedean kernels and to counting at the finite level: the leading minors act as pivots. It is important that one does not have to count all pairs of rational points of the full rank-r group indiscriminately. Coherent pairs are sorted by common frame data and by a linear relation between the two rank-k middles; then the deepest strata admit a local orbital calculation. The smaller strata are controlled by means of the same discriminant and congruences.

6.4 The archimedean Bessel system

For each $v \in \Phi$ the group at the normalized real place is the unitary group $U(k+h, h+1)$ in the split basis just described. Denote it by G_1. We use Haar measures in the normalization of `BesselSystem.lean`; in particular the measure on $U\backslash G_1$ is the quotient by the specified unipotent measure and is not a probability measure. The compact type of $H_0(\mathbb{R}) \simeq U(k)$ is identified with the prescribed column representation (of dimension $|I_v|$). The archimedean parameter of the target representation has large gaps: the local construction uses a gap at least r^7 and each doubled parameter has absolute value at most $3a$. The finite-dimensional column type has weight $(\kappa_v^{\mathrm{col}}, 0, \dots, 0)$ in the convention of the proof: its only potentially nonzero entry is the one corresponding to the marked first column of order a at the tops. This weight should not be confused with the real scaling constant κ_v just introduced.

The representation is described in Lean via a reproducing kernel rather than via a choice of realization of a discrete-series Hilbert space. An invariant kernel of formal degree d_v, acting by convolution, reproduces the functions of its type. Its integral against the Bessel character constructs a relative kernel; the associated map on the compact type is normalized to have prescribed L^2 norm. The source constructs a full function and a cutoff with unusually strong uniform bounds despite the changing representation and rank. Here are the exact properties that will be used.

Definition 6.5 (Quantitative Bessel system)**.** For a real place v, let I_v be a nonempty finite set, let $D_v = |I_v|$, and let $d_v > 0$. Let W_v (full) and $W_v^\flat$ (cut) be continuous functions on G_1 with values in $\mathbb{C}^{I_v}$, equivariant on the left under U through the fixed character and under H_0 through the specified irreducible unitary representation. Define the relative size

$$\mathrm{bs}(y) = \inf_{s \in S_B} (1 + \|sy\|_{\mathrm{op}}).$$

The following conditions are part of the quantitative Bessel system:

1. The full function is integrable on $U\backslash G_1$, its squared L^2 norm is D_v, and the cut has squared norm at most D_v and vanishes where $\mathrm{bs}(y) > b^{6s_0+2}$. Moreover

$$\|W_v^\flat - W_v\|_2^2 \le b^{-10r^2} D_v.$$

2. There are a measurable left U-invariant majorant A and a finite measure μ on G_1 such that each of the two norms is bounded by A, and in the region where the functions

differ their norms are bounded by $b^{-3r^3 f}A$. In addition

$$A(y) \le \mu\{g : \mathrm{bs}(yg) \le b^{6s_0+2}\}, \qquad \mu(G_1)\int_{U\backslash G_1} A(y)\,dy \le b^{2r^3} D_v.$$

3. For each word α in the matrix-coordinate directions, with $|\alpha| \le 2s_0^2$, there is a continuous derived Bessel function $W_{v,\alpha}$. Its coordinate right differentiation relation is the ordinary chain rule along a curve $\gamma'(t) = \gamma(t)X$, namely the derivative of $W_{v,\alpha}(\gamma(t))$ is the sum of the coordinates X_{ij} times $W_{v,(i,j)\alpha}(\gamma(t))$. For every $P \ge b^5$ and every integrable scalar ϕ on G_1, the derived Schur bound is

$$\int_{U\backslash G_1} \left\| \int_{G_1} \phi(y) W_{v,\alpha}(\gamma y)\,dy \right\|^2 d\gamma \le P^{2|\alpha|}\frac{D_v}{d_v}\|\phi\|_2^2. \tag{70}$$

4. The Bessel correlation of the cut, after the changes of coordinates below, belongs to a tower of kernels with explicit Fourier bounds, called a sieve system. Its bottom constant is $(1 + b^{-10})/d_v$.

These are precisely `IsCutPair`, `HasSchurBound`, `BesselSystem` and `HasBandLimits` of `New/Filling/BesselSystem.lean`. The word "derived" refers here to differentiation, not to derived categories. Condition (2) is stronger than L^2 approximation alone: after summing translates over a growing lattice, a small L^2 error does not automatically imply a small contribution to the second moment.

Proposition 6.6 (Formal-source input: existence of local systems)**.** *There exist quantitative Bessel systems with the prescribed formal degrees, type, and column-coordinate representations at all $v \in \Phi$. The cutoff system satisfies the Fourier bounds of (72) below, and the open-cell estimate of (73) below.*

The existence assertion is `Data.exists_besselSystem` in `New/Filling/FillingData.lean`, obtained from `HasPeakAt.exists_cut_isCutPair` and `IsBesselFunction.hasSchurBound_of_convolve_eq` in `New/RealPlace/BesselCut.lean` and `hasBandLimits_of_isCutBessel` in `New/RealPlace/BandLimits.lean`. The local construction obtains a peak of a Bessel convolution by lifting between unitary pairs with separated parameters; localization and lower bounds on the peak are what permit normalization by the peak value. The gap and the bound on the highest parameters are used both in the tail of the kernel and to control iterated right derivatives. For example the derivative estimate for a reproducing kernel of rank r, highest doubled parameter bounded by A_0, permits a Schur constant P at least $(r+1)^4(A_0+r+1+2s_0^2)$; the admissible data absorb this in b^5. In reproducing-kernel notation the reason for the Schur estimate is that convolution is an orthogonal projection up to the factor d_v. After normalization and differentiation, its operator norm is bounded by $d_v^{-1/2}$ times the norm of the differentiated vector. The nontrivial uniform part is bounding those vectors and obtaining the stated peak and cutoff, parts of the above formal-source input; neither follows merely from the existence of a discrete-series representation.

Lemma 6.7 (Orthonormality of the full coordinates)**.** *For each v and all $i, j \in I_v$,*

$$\int_{U\backslash G_1} W_{v,i}(y)\overline{W_{v,j}(y)}\,dy = \delta_{ij}.$$

Proof. The integral matrix $T = \int W_v(y)W_v(y)^* dy$ is a Hermitian positive semidefinite operator on the coordinate space. Left invariance of the quotient measure under the compact group H_0 and Bessel equivariance show that T commutes with its irreducible unitary representation.

Schur's lemma gives $T = aI$ with $a \geq 0$. Taking traces and using condition (1) gives $aD_v = D_v$, so $a = 1$. This is the argument of `BesselSystem.orthonormal` in `New/RealPlace/BesselOrthonormal.lean`. □

We record the sharper local correlation bounds because they fix the constant in the global moment. For any vector-valued W on G_1 let

$$J_W(\gamma) = \int_{G_1} W(y) \otimes W(\gamma^{-1}y)^* \, dy$$

when this integral is defined. For $W = W_v^\flat$, the full open-cell Weyl representative with head $p = (p_0, \ldots, p_h)$ is denoted w_p. Here $p_i \neq 0$ are the successive pivot ratios in elimination of the frame-product matrix; for $i = h$ we have $p_h = \langle z, \sigma z \rangle$ for the residual middle element $\sigma \in U(X_0 \oplus \mathbb{C}z)$. Let $V(U(j))$ be the volume of the compact unitary group in the source normalization and put

$$c_G = \frac{V(U(k))}{V(U(r))} \frac{2\pi^k}{(k-1)!}, \qquad c_{\mathrm{band}} = \frac{V(U(k))}{V(U(r))} \pi^{(2h+1)k}.$$

For matching top entries $|p_h| > 1$ define the normalized correlation tensor

$$T_v(\sigma, p) = \frac{\sqrt{c_G}\,(|p_h|^2 - 1)^{(k-1)/2}}{D_v} \prod_{i<h} |p_i|^{k+1+2(h-i-1)} \, J_{W_v^\flat}(w_p \sigma). \tag{71}$$

The normalization of the tensor has a particular Jacobian meaning: the powers in the product are the half-densities of the successive rank-one steps in the open cell. In the last step the residual middle group contributes the volume of a sphere of dimension $2k-1$ and the radial factor $(|p_h|^2-1)^{k-1}$. This is the real factor that will compensate the global finite-place quadric count.

The top kernel of the local sieve system is the inner product of the tensors (71), when the two top entries both come from σ. Hence the top kernel is positive semidefinite for common top entry. The frequency bound for the system has the following explicit majorant. Write

$$\lambda_a(u) = u^a \min(1, u)^6, \qquad \Lambda_i(p) = \prod_{j<i} \lambda_{s_0^6}(|p_0 \cdots p_j|^{-1}).$$

In the integral of one common entry at level i the nonzero Fourier coefficient, at frequency ξ, is bounded in absolute value by

$$\beta_v(\xi)\Lambda_i(p)\Lambda_i(q), \qquad \beta_v(\xi) = c_{\mathrm{band}} b^{2000 s_0^9} \lambda_{k - s_0^7}(|\xi|^{-1}). \tag{72}$$

The same bound in the unequal-entry relation (with the next head weights) is used below. These are the actual constants in `HasBandLimits`. The large exponent $k - s_0^7$ is critical; a bound merely asserting rapid decay for fixed parameters would not provide the required powers after multiplying over f places. The exponent 6 in $\min(1, u)^6$ is a separate damping term used for summability of frequencies after applying the product formula. The construction of these kernels and the bounds (including the identity of the top kernel with the correlation) is the formal-source input `hasBandLimits_of_isCutBessel` and the preceding real-place calculations in `New/RealPlace/BandLimits.lean` and `LocalTheorem.lean`.

A second local estimate is used for the terms that cannot be treated by the sharp main-term calculation. Let $Z_v(\gamma)$ denote the real-place frame-product matrix and let $\mathfrak{d}_v = |\det Z_v|^2 -$

$|D_h(Z_v)|^2 > 0$. If $\rho_0 \geq b$ and the box in which all frame coordinates have norm at most ρ_0 meets its translate by γ, define

$$M_v(\rho_0, \gamma) = \left(\frac{\rho_0^{9s_0}}{\mathfrak{d}_v}\right)^{10^5 s_0^4} \frac{D_v^2}{d_v c_G(\mathfrak{d}_v/2)^{k-1}}. \tag{73}$$

The open-cell estimate (*Formal-source input*: `BesselSystem.open_cell` in `New/RealPlace/OpenCellBound.lean`) asserts the existence of a compact integration set, an averaging parameter in $U^0 H_0$, and a measure of total mass at most M_v, such that the averaged squared norm of the integrals against $W(\gamma ty)W(ty)^*$ is at most the integral of the squared modulus of the test function against this measure. The compact integration set contains the required translates of the box, almost surely in the averaging parameter. This is a local operator bound with an explicit discriminant pole, not a pointwise estimate for W. The normalization in (71) is intentionally different: in the sharp term one preserves cancellation between Fourier phases, while in (73) one uses upper bounds allowing losses. The large exponent on $\mathfrak{d}_v$ is compensated globally by discriminant bounds and congruence counts.

6.5 Construction and exact unfolding of the Poincaré map

Using the coordinate isomorphism of Proposition 6.4, endow C with its orthonormal coordinate basis $e_{u,i}$ for $u \in \mathcal{C}$, $i \in I$. Write $W(y)_i = \prod_v W_{v,i_v}(y_v)$ and similarly $W^\flat$ for the cut. The superscript $\flat$ now denotes the approximate function, not the auxiliary cutoff on unipotent cosets. The product Bessel functions live on the product of the normalized infinite groups, denoted G_∞. For any rational element γ, use γ_∞ and γ_f for its images in these normalized coordinates and in the finite lattice coordinates, respectively.

For $u \in \mathcal{C}$, let $U_u(g)$ be the rational unipotents with $h_u^{-1}\gamma_f h_u \in K(g)$. The image at infinity is a discrete group. Take a measurable nonnegative class cutoff β_u on G_∞ with

$$\sum_{v \in U_u(g)} \beta_u(vy) = 1. \tag{74}$$

Such a function may be obtained by taking a measurable fundamental domain, or by normalizing a nonnegative locally finite function that is positive somewhere on each orbit. The choice in the source is expressed as a cutoff for a discrete subgroup. The arithmetic assertion of discreteness, and the existence of a cutoff with the specified compatibility, is `exists_cutoff_classUnip` in `New/Filling/Setup.lean`. The measures are chosen so that the covolume of $U_u(g)$ in the infinite unipotents is a constant $\nu > 0$ independent of u. Explicitly, put $d_U = h(h+k)$ and $D_E = |\operatorname{disc}(E)|$. With $v_f = \operatorname{vol}(\widehat{\mathcal{O}}_E)$ in the self-dual finite measure of the source, the covolume is

$$\nu = \frac{(g^{2d_U+h}/N^k)^f}{v_f^{d_U}} \frac{D_E^{h/4}}{\mathcal{L}}. \tag{75}$$

This is `Filling.unipCovolume`. To see the origin of the exponent, the unipotent filtration has a middle step of size $k+1$ and two strings of sizes i, i for each $0 \leq i < h$; the sum of their sizes is $\sum_i (2i+k+1) = h(h+k)$. Each such adelic coordinate has congruence density and a self-dual covolume, while the normalized archimedean change of basis supplies the Jacobian $\mathcal{L}$. The norm term N^{-kf} occurs in the middle lattice modification. The exact quotient formula (75), including independence of u, is the *Formal-source input* `Filling.integral_classCutoff_eq_unipCovolume` of `New/Filling/UnipotentCovolume.lean`; the numerical factors

are those of `Setup.lean`. For an integrable function F on the unipotent quotient, the cutoff identity implies

$$\int_{G_\infty} \beta_u(y)F(y)\,dy = \nu \int_{U_\infty\backslash G_\infty} F(y)\,dy$$

provided F is invariant on unipotent orbits: integrate first over the left unipotent and then use (74) to replace that integral by its quotient covolume. This is the quotient calculation used repeatedly in the unfolding.

For (u,i) let $\phi_{u,i}$ be the seed given by $\beta_u(y)W_i(y)$ at infinity times the indicator that $h_u^{-1}g_f \in K(g)$, in the common coordinates. A fixed common adelic coordinate change is understood (the map called `Filling.Fits.finiteEquiv` and the inclusion of its fixed η in `New/Filling/PoincareSeries.lean`); it has no effect on the formulas for relative positions since it cancels from them. The Poincaré map $Q : C \to V$ is the periodization

$$Qz(g) = \sum_{u,i} z_{u,i} \sum_{\gamma\in G(E^+)} \phi_{u,i}(\gamma g). \tag{76}$$

When summing first over unipotents, the periodization replaces the cutoff β_u by 1 on the permitted finite coset, by (74) and the Bessel character being trivial on $U_u(g)$. This identifies (76) with the Poincaré series on the Bessel quotient; note that the source function called `cutBessel` in this construction contains the full W and the cutoff β_u, not the cut $W^\flat$.

To state the convergence and unfolding precisely, put

$$\mathrm{pos}_{u,u'}(\gamma) = h_u^{-1}\gamma_f h_{u'}, \qquad \mathcal{E}(u,s,u') = \{\gamma : \mathrm{pos}_{u,u'}(\gamma) \in K_f,\ \mathrm{red}(\mathrm{pos}_{u,u'}(\gamma)) = s\}.$$

Here K_f is the lattice stabilizer used in the filling datum, with reduction kernel $K(g)$. For a set T of rational terms define the amplitude matrices (coordinates $I \times I$)

$$A(\gamma;u,u') = \int_{G_\infty} \beta_u(y)\beta_{u'}(\gamma_\infty^{-1}y)W(y)\otimes W(\gamma_\infty^{-1}y)^*\,dy, \tag{77}$$

$$A_T(u,s,u') = \sum_{\gamma\in T\cap\mathcal{E}(u,s,u')} A(\gamma;u,u'). \tag{78}$$

Define $A_T^\flat$ with $W^\flat$ in place of W.

Proposition 6.8 (Formal-source input: convergence and reproduction)**.** *For every u,s,u' the sums in (78) are absolutely convergent in the finite-dimensional tensor space. For fixed u,s the sum over u' and the corresponding rational terms of the integrated norm difference of the full and cut kernels, with the product of the two cutoffs as in (77), is at most $\nu b^{-k^3 f}/2$. The sum of the norms in the inner periodization in (76) is bounded uniformly in g. The Poincaré map has image in V.*

The first assertions are `amplitudeAbs_lt_top_and_cut_le`, `summable_amplitude` in `New/Filling/AmplitudeSums.lean`; the last two are `Fits.tsum_cutBessel_le`, `Fits.poincareMap` in `New/Filling/PoincareSeries.lean`. They are not consequences of pointwise boundedness of the seed alone. The majorant in Definition 6.5(2), together with the bounded box partition at infinity and lattice estimates, is used to prove convergence before interchanging the integrals. The reproducing-kernel condition and integration on the Bessel quotient then put the Poincaré series in the space of the required archimedean cusp type and principal finite level.

Lemma 6.9 (Exact coefficient identity)**.** *In the Petersson normalization there is a constant $v_0 = \mathrm{vol}(K(g)) > 0$ such that*

$$\langle Qe_{u,i}, \rho(s)Qe_{u',j}\rangle = v_0 A_{\mathrm{all}}(u,s,u')_{i,j}. \tag{79}$$

Here and below the pairing uses the convention in which the order of i, j agrees with (77); if one takes the opposite convention, both sides are conjugated/transposed accordingly.

Proof. Unfold the first periodization in the quotient Petersson integral and expand the second into its defining series. This gives a single sum of integrals over the whole adelic group, of the product of a seed of class u with a translated conjugate seed of class u'. Absolute convergence and Fubini are justified by Proposition 6.8 and the integrability statement `Fits.integral_poincare` in `New/Filling/PoincareSeries.lean`. After left translating the finite integration variable by the class representative h_u^{-1}, its range is $K(g)$. Normality of $K(g)$ in K_f implies that the second indicator is identically 1 on this range precisely when the relative element has reduction s, and otherwise identically zero. Hence the finite integral contributes v_0 and restricts the rational terms to the set $\mathcal{E}(u, s, u')$. The archimedean part of the same integral is (77); the two class cutoffs remain in it. Thus the integral is precisely (79). This is the unfolding carried out by `Fits.petersson_poincare` in `New/Filling/PoincareSeries.lean`. □

Set $P = (v_0\nu)^{-1/2}Q$. Then $P^*\rho(s)P$ in the coordinate basis has entries $\nu^{-1}A_{\mathrm{all}}(u, s, u')_{i,j}$. Projecting by R gives the translates of the row map after restriction to the ramified space, up to their fixed nonzero scalar. It is the normalized map P that will enter the variance estimate, whereas the integral-index theorem uses the actual unscaled map of Theorem A.

Lemma 6.10 (Identity term)**.** *The amplitude of the Bessel subgroup at $s = 1$ is exactly*

$$A_{S_B}(u, 1, u')_{i,j} = \nu\delta_{u,u'}\delta_{i,j}.$$

In particular the Bessel-subgroup contribution to P^*P is the identity operator.

Proof. The elementary reduction property used here is `Filling.eq_and_mem_classUnip_of_besselS` in `New/Filling/NearTerms.lean`: a term in S_B over $(u, 1, u')$ forces the two column classes to be equal and belongs to $U_u(g)$. This follows by projecting to H_0 and using the definition of the column classes and triviality of their automorphism groups at K_c (the precise compatibility with the modified lattice and the level is the cited formal-source input). Thus only $u = u'$ can contribute. On summing over the unipotents, covariance of W and (74) eliminate the second cutoff. The remaining integral is ν times the integral of $W_i\overline{W_j}$ on the unipotent quotient. Use the product over v of the orthonormality relation in Lemma 6.7. This proves the claimed identity (`amplitudeSum_besselS_one` in `New/Filling/AmplitudeSums.lean`). □

6.6 Energy, pairs of congruent terms, and removal of lower cells

The following notation allows the estimates for the moment to be stated with their actual normalizations. For a subset $\Omega \subseteq \mathfrak{G}$ define

$$\mathcal{E}(A_T; \Omega) = \frac{1}{\nu^2} \sum_{u,u' \in \mathcal{C}} \sum_{s \in \Omega} \|A_T(u, s, u')\|_{\mathrm{HS}}^2. \tag{80}$$

The norm here is the Euclidean norm on the $I \times I$ tensor. Put

$$D = \frac{c^2}{d_\infty \operatorname{mass}_r(E)}, \quad D_1 = |\bar{S}_B|c, \quad \tau = b^{kf/4}, \quad \mathfrak{c} = 2\tau^2(b/g)^{kf},$$

where $\bar{S}_B$ is the reduction of the integral Bessel subgroup. These are `mainEnergy`, `besselEnergy`, `tailScale`, and `codimFactor`, respectively; the formulas and inequalities in the following proof use this normalization. In particular, the admissibility assumptions give

$$\tau^{44}D_1 \le D, \qquad \tau^{18}\mathfrak{c}D \le D_1, \qquad \mathfrak{c} \le 1. \tag{81}$$

This pair of bounds is central: the Bessel-subgroup energy, despite its large number of translations, is extraordinarily small on the scale of the full expected moment. The first bound means it is inexpensive to use coarser estimates on the lower cells; the second expresses the strong saving from nontrivial finite congruence codimension. They are part of `Filling.IsAdmissible` and are verified for the data in `New/Filling/FillingAdmissible.lean`; the verification is a *Formal-source input* using the mass and unipotent measure formulas and degree bounds. The first equality in (68) shows $D/|\mathfrak{G}| = c^2/n_0$.

For two compatible collections of amplitudes A, B, any $\delta > 0$ gives

$$\mathcal{E}(A + B; \Omega) \le (1 + \delta)\mathcal{E}(A; \Omega) + (1 + \delta^{-1})\mathcal{E}(B; \Omega). \tag{82}$$

This follows from $2\Re\langle a, b\rangle \le \delta\|a\|^2 + \delta^{-1}\|b\|^2$ in each summand. By absolute convergence the amplitude for a set decomposes into amplitudes for disjoint subsets. The global truncation error follows from Proposition 6.8 by applying the triangle inequality to each tensor, squaring, and summing:

$$\mathcal{E}(A_T - A_T^{\flat}; \Omega) \le |\mathcal{C}|\,|\mathfrak{G}|\,b^{-2k^3 f}/4 \le b^{-30} D_1.$$

For the last inequality use $|\mathfrak{G}| \le (g_0 g_1)^{r^2 f} \le b^{22r^2 f}$, $c \ge |\mathcal{C}|$, $|\bar{S}_B| \ge 1$, and k much larger than h, hence $2k^3 f - 22r^2 f \ge 30$. This is `energy_sub_cut_le` in `New/Filling/SchurTest.lean`. Thus there is substantial freedom to cut off the Bessel functions. No comparable freedom is allowed in the constant of the main term.

The expansion of the squared amplitude in (80) gives pairs of rational elements *with equal finite reductions*: for each (u, s, u'), their relative positions belong to the same coset of $K(g)$ in K_f. These pairs are stratified by degeneracy of their flags and by the links between their middle spaces. We describe the stratification and record its precise resulting estimate.

Character orthogonality. Call a set of terms stable if it is invariant under left and right multiplication by rational unipotents. Write ψ_U for the additive Bessel character on rational unipotents as read in finite adelic coordinates; equivalently it is the inverse product of the infinite-place characters. For a term γ, a necessary condition to contribute is that

$$\psi_U(u) = \psi_U(\gamma^{-1} u \gamma) \quad (u \in U(E^+) \cap \gamma U(E^+)\gamma^{-1}). \tag{83}$$

Indeed the product $W(y) \otimes W(\gamma_\infty^{-1} y)^*$ transforms, on replacing y by $u_\infty y$, by the quotient of these characters. To see why cutoffs cause no obstruction, group the sum over $T \cap \mathcal{E}(u, s, u')$ into $U_u(g)$–$U_{u'}(g)$ double cosets. On a double coset, the product of the cutoffs can be replaced by any measurable weight of total one over the double action: this follows by Fubini from summing the two cutoff identities (74). The resulting integral is an orbit integral, and the mismatch character on the stabilizer makes it vanish. For a nontrivial character choose an element at which it is not one; the integral equals that nontrivial scalar times itself. In the formal proof the rational stabilizers relevant to the finite cosets and the full unipotent intersection give the same vanishing condition by the approximation theorem for the unipotent group, and the exact weight-change identity (*Formal-source inputs*: `hasSum_amplitudeSum_change_of_weight` in `New/Filling/AmplitudeSums.lean`, `amplitudeSum_eq_zero_of_not_charsAgree` in `New/Filling/CrudeEnergy.lean`, and `exists_rat_approx_unipotentChar_eq` in `NearTerms.lean`). The character must be retained until this step; replacing it by its modulus would lose the exact vanishing.

Normal forms and links. A term with degenerate discriminant $\mathfrak{d}(\gamma) = 0$ but compatible characters has a shared vector in the frame flags. The proof uses a descending family of subgroups G_j fixing the first j hyperbolic pairs ($0 \le j \le h$). For such a shared vector,

compatibility of characters forces equality of the Hermitian norms of the two leading coefficients and propagates the shared-vector relation to subsequent levels. To see the mechanism, let v be a vector orthogonal to the relevant preceding flag, and w an isotropic vector of the flag lying in that orthogonal complement. The unitary Eichler transformation is

$$x \longmapsto x + w\langle v, x\rangle - (v + \eta\langle v, v\rangle w)\langle w, x\rangle, \qquad \eta + \bar{\eta} = 1.$$

Expanding the two pairings using $\langle w, w\rangle = \langle w, v\rangle = 0$ and the relation on η shows that this transformation preserves the Hermitian form. Its simple-root coordinates against the flag are linear in the coefficients of w and in the pairings of v with the successive flag vectors. The Bessel character of the transformation tests this linear combination. Applying the transformation with the corresponding common vector on each side of the conjugation, equality of characters for all rational scalings, and nondegeneracy of the additive pairing, forces equality of the linear functionals. Elementary linear algebra then propagates the shared-vector relation up the flag. In the source these steps are `exists_mem_besselU_eq_eichlerMap` in `EichlerFlag.lean`, `hasCommonFlagVector_of_le` and `CharsAgree.exists_hasLevel` in `CrudeEnergy.lean`. The latter supplies the exact assertion used below: a term satisfying (83) with discriminant zero has a level $j \le h$ in the sense of `Filling.HasLevel`. We use this implication as a *Formal-source input*, retaining its character hypothesis.

The link of a term is a linear subspace of $X_0 \oplus X_0$. In the split basis, let $M(\gamma)$ have as columns $e_0, \dots, e_{h-1}, \gamma e_0, \dots, \gamma e_{h-1}$ and let p_{X_0} denote projection to the rank-k positive middle coordinates. The link is the image of the kernel of $M(\gamma)^* J_Y$ under the map

$$z \longmapsto (p_{X_0} z, p_{X_0}\gamma^{-1} z).$$

It describes the relation between the two positive middle coordinates after imposing orthogonality to both isotropic frames. This definition is `link` and `linkMatrix` in `New/Filling/EichlerFlag.lean`. When two terms have different links, their common part has smaller dimension; the resulting loss is reflected in the congruence codimension factor $\mathfrak{c}$. For the top cell (nonzero discriminant) the primary surviving energy is the sum of squared norms of blocks for a fixed link, not the sum of squares of individual terms. For a set $\mathcal{L}$ of links put

$$\mathcal{E}_{\mathrm{link}}(T, \Omega, \mathcal{L}) = \sum_{u,u',s\in\Omega} \sum_{\Lambda\in\mathcal{L}} \|A^{\flat}_{T\cap\{\mathrm{link}(\gamma)=\Lambda\}}(u, s, u')\|^2.$$

Unlike (80), this definition does not divide by ν^2; we keep the source normalization.

To estimate the discarded terms one first replaces the double cutoff by a partition in frame boxes of size $b^{2s_0^2}$, using a cutoff modulo the frame stabilizer U^0. Within such boxes all frame-product entries have bounded archimedean norm. Their finite congruence constraints and the product formula give lower bounds on the nonzero discriminants and on relevant lattice indices. The proof further selects near representatives of the U^0 double cosets (`IsNearRep` of `BoxPartition.lean`). For each pair of near representatives with matching finite reductions and level j, one bounds its contribution by products of explicit weights, introducing the lattice of common link vectors. The weights are inverse powers of: the discriminant of the lattice of link vectors, the index of the frame-value lattice, the saturation index of the paired frame lattice, and the factor $g^{2f\min(j,h)}$. In particular the pair invariant of `pairInvariant` is the product of the discriminant, frame-value index, and saturation index. The weight `termWeight` includes the powers $-(k+1), -k, -k$ of the three corresponding invariants (with a count normalization depending on the codimension).

The point of these inverse powers is that the total weight can be bounded by estimates for sublattices rather than a direct count of pairs. The geometry-of-numbers input (*Formal-source input*: `Filling.tsum_sublatticeDisc_inv_le` in `NearTerms.lean`) bounds sums of inverse

discriminant powers using short generators obtained by a relative Siegel lemma. Bounds on the congruence indices and lattice intersections followed by these sums give the explicit theorem `sum_pairWeight_le` in `BlockBound.lean`: the total pair weight is at most $2\tau^2$. The class amplitude size prior to summation is the function `classSize` of `SchurTest.lean`. With M_0 the normalization `countWeight0` and z the relevant index parameter it has the exact form

$$a_j(z) = \tau^2 \left(\frac{D_j}{M_0 g^{2\min(h,j)k} f\, z^{k+1}} \right)^{1/2}, \quad D_j = \begin{cases} \tau^{13-32j} D_1 & j \le h, \\ D & j = h+1. \end{cases}$$

At the last level the estimate for $a_j(z)$ is obtained from (73) and the archimedean and finite discriminants. At earlier levels the weighted Schur estimate bounds the retained common-vector blocks. The exact comparisons with class sizes and the total weights are `IsNear.le_classSize_and_cellMass_le` in `SchurTest.lean`, and `enorm_amplitudeSum_le_classSize` and `enorm_amplitudeSum_le_classSize_top` in `BlockBound.lean`. We take these as formal-source inputs. These estimates are uniform in the rank and in the field; qualitative finiteness of the double cosets or a fixed-rank Siegel lemma estimate alone would not give the stated bound.

Proposition 6.11 (Formal-source input: reduction to coherent top-cell pairs)**.** *Retain the notation D, D_1, τ. Suppose T is stable, Ω is a set of reductions with the common fixed-vector norm bound $X \ge 1$ of `Filling.FixesNormAtLeast`, and $C_0 \ge D_1$ is a constant satisfying the following extra restrictions. If a term in T over Ω is in normal form zero, require $C_0 \ge \tau^{20} D_1$; if an omitted link exists, require $C_0 \ge \tau^{12} D/X^k$. For $w \ge 0$, suppose*

$$\mathcal{E}_{\mathrm{link}}(T \cap \{\mathfrak{d} \ne 0\}, \Omega, \mathcal{L}) \le w\nu^2.$$

Then

$$\mathcal{E}(A_T; \Omega) \le (1 + b^{-4})^2 w + b^{-12} C_0. \tag{84}$$

This is `Filling.energy_le_of_linkEnergy_le` in `New/Filling/CrudeEnergy.lean`; the two subordinate assertions are `energy_le_of_level_pairs` and `energy_le_of_level_lt`. Their quantitative errors are $C_0/(b^{30} 4^{h+1-j})$ at level j and the accumulated error below level n' is at most $C_0/(b^{30} 4^{h+1-n'})$. To explain the accumulation, split off the next level by (82) (taking $\delta = 1$ here); its own error and the preceding one together cost at most four times the preceding-level budget. Thus a geometric factor 4 suffices from one level to the next. If a term has discriminant zero, the character obstruction or a lower normal form just discussed applies, so it contributes only to this error. At the last level, separate the coherent equal-link blocks, of cost $w\nu^2$, from the unequal-link blocks estimated by the common link and its codimension. Returning to the full amplitudes with (82), now using $\delta = b^{-4}$ and the extremely small truncation error, gives (84). Both the precise coefficient of w and the estimate on the omitted terms matter: choosing a fixed loss at the final step instead of $1 + b^{-4}$ would not yield a nearly spanning family. The nontrivial block estimates above, rather than their algebraic assembly, are the imported part of this proposition.

A first consequence concerns translations that locally lie in the image of the Bessel subgroup. For $i = 0, 1$, set

$$\Omega_i = \{s : \text{there exists } t \in \bar{S}_B \text{ with } s_i = t_i\}.$$

Proposition 6.12 (Formal-source input: Bessel-coset error)**.** *For each i one has*

$$\mathcal{E}(A_{G(E^+)\backslash S_B}; \Omega_i) \le b^{-12} D_1.$$

This is `energy_compl_besselS_le` in `New/Filling/Moments.lean`. For these reductions the frames satisfy the exact Bessel constraints on an entire prime's chosen factors. The finite

congruences and the character condition imply that a remaining normal-form-zero term would actually be in the Bessel subgroup, so such a term does not contribute to the complement. The common constraints supply the norm bound X for applying Proposition 6.11 with $\mathcal{L} = \varnothing$, $w = 0$, $C_0 = D_1$; the scale inequality (81) guarantees the requisite domination of the full-cell error. The exact group-theoretic constraint and norm calculation are part of the cited source proposition. This restrictive error estimate isolates the diagonal term and is used again to bound the non-Bessel part of a central projection.

6.7 The sharp coherent main term

The preceding estimates deliberately allow coarse constants, because they are used only for errors. For the coherent open-cell pairs a different argument must retain enough positivity to give the leading constant one. We record the form of this argument and isolate its three substantive inputs: local orbital kernels, rational-to-local block comparison, and the global sieve.

Kernel systems and their products. Consider a measured commutative ring A with an additive character $\psi : A \to \mathbb{C}^\times$ of modulus one, and let $A^\times$ be its group of units. For a general semilocal ring put $u(z) = z$ if z is a unit and $u(z) = 1$ otherwise. Let p, q denote sequences of units; $p[i : c]$ means the sequence obtained by replacing entry i by $u(c)$. A sieve system is a family of forms $J_i(p, q)$, $0 \le i \le h + 1$, with the following relations:

$$\int_A J_{i+1}(p[i : c], q[i : c])\, dc = J_i(p, q), \qquad |J_0(p, q)| \le \mathfrak{e},$$

$$\left| \int_A \psi(\xi c) J_{i+1}(p[i : c], q[i : c])\, dc \right| \le \beta(\xi)\Lambda_i(p)\Lambda_i(q). \tag{85}$$

The Fourier inequality is required for unit frequencies, and all the integrals in these relations are absolutely convergent. For $i < h$ one also requires the unequal-entry bound

$$|J_{i+1}(p, q)| \le \beta(\iota(p_i) - \iota(q_i))\Lambda_{i+1}(p)\Lambda_{i+1}(q)$$

when the difference of inverse-head coordinates in this expression is a unit; the map ι is inverse of a unit, with the additional factor $1/g$ at the finite places. At the top the kernel is continuous and every matrix $\left(J_{h+1}(p_a, p_b)\right)_{a,b}$ with common last entry is positive semidefinite. This is an abstraction of `SieveSystem` in `New/Filling/EichlerFlag.lean`; we have omitted the infinite tails of the sequences from the notation.

Here is a useful consequence. If p_a is any finite family of heads, for $i \le h$ define the function

$$F_i(c) = \sum_{a,b} J_{i+1}(p_a[i : c], p_b[i : c]).$$

This is nonnegative, because the top forms are positive semidefinite and lower forms are obtained by integrating a common coordinate, successively. It is lower semicontinuous wherever the head maps are continuous, by Fatou's lemma after integration over subsequent entries. Fubini and (85) imply

$$\int F_i = \sum_{a,b} J_i(p_a, p_b), \qquad |\widehat{F}_i(\xi)| \le \beta(\xi) \left(\sum_a \Lambda_i(p_a) \right)^2.$$

The same scheme applies to the product of systems at finitely many places; the top positive matrices multiply entrywise and remain positive semidefinite by the Schur product theorem,

the integration compatibility follows from Fubini, and the constants and majorants multiply. This is the elementary content of `SieveSystem.pi`, `SieveSystem.step_sum`, and `SieveSystem.nonneg_and_lowerSemicontinuous`. Positivity is retained in the sum over both heads; the individual summands need not be nonnegative, and no absolute value may be inserted before the sum without changing the argument.

Finite orbital kernels. Let Σ be a set of chosen places above g_0 and fix, for each $v \in \Sigma$, a nonzero frame vector at the place or its conjugate. Define $\Omega^\circ(\Sigma)$ to consist of the level reductions fixing the specified vectors, and for which every required leading frame minor is a unit at all chosen places above g_0 not in Σ. This is the set denoted `fixingSet` in the formal proof. The omitted-minor test at places not in Σ is indispensable. At an arbitrary finite conjugation-stable set of places containing the level primes and all denominators of a finite initial selection of rational terms, the argument works over the product of the corresponding local fields. Choose an intermediate compact open subgroup between $K(g)$ and K_f whose local factor at each place contains the appropriate fixed-vector subgroup. The ratio of the bottom orbital integral to the group measure is diminished by a factor b^{-6k} at each specified place. The precise construction is `exists_intermediateGroup_one_prime` and the product version `exists_intermediateGroup_lines` in `New/Filling/FixedLines.lean`; it uses counting on split matrices and a bound on the volume of nonunits (at all the level places this is at most b^{-9}).

The finite system incorporates the Bessel character of conductor given by $1/g$ and the different. If $p = (p_i)$ is a head over the local product of fields, let $[D_j]$ denote the generalized ring index associated to the partial product $D_j = p_0 \cdots p_j$. Write $\mathrm{den}(\xi)$ for its denominator index. The finite head majorant is

$$\Lambda_i^f(p) = \prod_{j<i} [D_j]^{s_0^6/2} \quad \text{if all } D_j \text{ are integral, and 0 otherwise.}$$

The finite frequency majorant is

$$\beta_f(\xi) = \frac{[\xi]^{(k-s_0^7)/2}}{\mathrm{den}(\xi)^6} b^{s_0^9 f},$$

where $[\xi]$ is the generalized index, not necessarily an integer. These definitions are `finHeadWeight` and `frequencyGain` in `EichlerFlag.lean` and `GlobalSieve.lean`. After dividing the top orbital kernel by the volume of a compact open additive neighborhood of zero and by the prescribed mass ratio, the bottom constant of the finite sieve system is at most

$$(1 + b^{-6}) b^{-6k|\Sigma|}. \tag{86}$$

These are the conclusions of the *Formal-source input* `Filling.IsFillingLevel.exists_finite_sieveSystem` in `New/Filling/FiniteSystem.lean`. We emphasize that the normalization of the orbital kernel and the bottom constant include the measure of the fixed-vector subgroup, rather than only a bound on its order. The proof passes down the tower of groups fixing successive hyperbolic pairs, integrating one entry of the frame Gram matrix at each step. Its key step is a local integral comparison (`Level.integral_succ_floor_eq` and `Gauge.integral_floorForm_succ`, `New/RealPlace/LocalTheorem.lean`). In the nonzero-frequency bound one uses the local character along a simple-root direction of an Eichler transformation, and averaging gives a power of the inverse module of the frequency. A parallel unequal-head calculation uses frequency $(p_i^{-1} - q_i^{-1})/g$; this is the origin of the unequal-entry bound in the definition of a sieve system. These local comparisons are substantive statements for indefinite unitary groups of growing rank, not applications of Plancherel alone.

The archimedean and finite systems together. At the infinite places use the systems of Section 6.4, and at the finite ones the systems of (86). The diagonal embedding of the rational frame entry into the local products scales it at each v by l_{vi}. Its covolume factor in Poisson summation is

$$c_i = \frac{\mathrm{vol}(\widehat{\mathcal{O}}_E)}{\prod_{v\in\Phi} l_{vi}^2} \qquad (0 \le i \le h).$$

The square in the denominator is the real Jacobian of multiplication by l_{vi} on a complex coordinate. Let A_0 be a finite set of rational units integral away from the auxiliary finite set, along with their inverses, and let $V_0 \subseteq A_0$ index the last rational entry. With the product system write $p[t]$ for the $h+1$ entries obtained from $p \in A_0^h$ and $t \in V_0$, read diagonally in the local coordinates. The following inequality is the output of the global sieve.

Proposition 6.13 (Formal-source input: the global sieve)**.** *If a finite sieve system has bottom bound $\mathfrak{e}_f$, head and frequency weights as above, then*

$$\Re \sum_{t\in V_0} \sum_{p,q\in A_0^h} J_{h+1}^f(p[t],q[t]) \prod_{v\in\Phi} J_{h+1}^v(p[t],q[t]) \le \prod_{i=0}^{h} c_i \left(\mathfrak{e}_f \prod_{v\in\Phi} \frac{1+b^{-10}}{d_v} + \frac{b^{-6kf-10}}{d_\infty} \right).$$

The exact theorem is `sieve` in `New/Filling/GlobalSieve.lean`. The proof uses `one_sided_poisson` in that file; we give its mechanism because it explains the small additive error. For the function F_i above, one-sided Poisson summation on the additive adele lattice gives a bound for the sum over a finite set of rational points by c_i times its zero-frequency integral plus c_i times an error estimated by the sum of its nonzero frequency coefficients. A one-sided version is needed for lower semicontinuous kernels; it follows by comparison with integrable smoothing approximants to a nonnegative function and the elementary finite-subset inequality. The Fourier bounds and the assumption $(\sum_p \Lambda_i(p))^2 \le \Gamma \prod_{a<i} c_a$ give an error at most $\Gamma\varepsilon \prod_{a\le i} c_a$ at this step if the sum of nonzero-frequency majorants is at most ε. When undoing a common head coordinate to descend another floor, the unequal head contributions give a second error of this size, bounded using the off-diagonal condition. Consequently induction from $i = 0$ gives a bound

$$\prod_{a\le i} c_a(\mathfrak{e} + (2i+1)\Gamma\varepsilon).$$

The source's `one_sided_poisson` states this explicitly, with rational frequencies that are integral away from the finite set. Products of the weights β_f and β_v are summable uniformly in the finite set: by the product formula the large powers of the norms cancel, leaving the damping power 6 and denominator factors. The needed elementary sum with exponent k' is the precise inequality `sum_frequencies_le`: over a finite set of nonzero T-integers y, if $|c_v|^{-1} \le |x_v| \le B_1$, $B_1 \ge 1$, then

$$\sum_y \frac{[y]_T^{k'/2}}{\mathrm{den}(y)^6} \prod_v \lambda_{k'}(|c_v v(y)|^{-1}) \le (b\,\mathrm{rd}_E)^{3f} \prod_v (|x_v|^{k'}(bB_1^2)^3).$$

This is proved in the source using a sum of inverse discriminants of rank-one sublattices of a rank-two lattice (the vector $(1,y)$), including its denominator lattice. It is a formal-source input, together with the one-sided Poisson lemma in the stated level normalization. Taking $k' = k - s_0^7$ and the actual archimedean scaling the strong inequalities of `Filling.IsAdmissible` give $(2h+1)\Gamma\varepsilon \le b^{-6kf-10}/d_\infty$. This is substantially better than a sieve bound with a factor b^η, and is the reason the second moment has the requisite leading constant.

Comparison of blocks with the orbital kernel. On the nondegenerate cell the terms are put in Weyl normal form $w_p\sigma$ with common residual middle element σ, common top entry $t = \langle z, \sigma z\rangle$, and $|v(t)| > N$ at every infinite place. The terms supported on identical finite reductions and identical link are sorted into such packets. Finite sets of initial representatives suffice by the absolute convergence of the amplitudes; subsequently enlarge the finite set of places to contain all denominators and take limits of local compact-open approximations. The exact comparison is `sum_blocks_le_topKernel` in `New/Filling/FiniteSystem.lean`; here is its normalization. Let $\mathcal{U} = |\bar{U}| = g^{fh(2k+2h+1)}$ be the order of the reduced integral unipotent subgroup, let

$$c_* = (\pi^k/((k-1)!N^k))^f\, D_E^{1/4}\, v_f^k, \qquad A_* = |\bar{S}_B|\mathcal{U}|\mathcal{C}|c_*.$$

For a finite packet of distinct quadruples (u, u', s, σ) with the given t, the squared norm of the corresponding sum of coset amplitudes, multiplied by the common normalization, is bounded by

$$(1+b^{-6})A_*\,\Re\sum_{p,q} R_{pq}\frac{J^{\mathrm{top}}(p,q)}{\mathrm{vol}(L_0)\ \mathrm{massRatio}},$$

for sufficiently small compact open additive L_0. Here R is the Gram matrix of the product of normalized archimedean correlation tensors. The exact factors in R are those of (71), and the finite module factors in the packet comparison cancel against these factors. This is a *Formal-source input*: the proof uses a count of integral points on a quadric for the last coordinate (`siegel_sum_le_lintegral_quadric`), and a local block comparison with the positive phase kernel, then the global-to-local limit identity `exists_hasSum_cosetAmplitude` in `New/Filling/GlobalToLocal.lean`. The normalization A_* incorporates the unipotent order and the same-class condition. It should not be replaced by a crude number of rational points. The factor c_* is the constant of the rank-k definite sphere sum. The factor $(|v(t)|^2 - N^2)^{k-1}$ in this count is the one that cancels against the open-cell Jacobian in the normalized real-place correlations.

The size comparison used in the compact proof is the following hypothesis of the filling datum (with our notation):

$$A_* \le (1+b^{-9})D\, d_\infty c_G^f\, v_f^{-s_0}\left(\frac{\nu\mathcal{L}}{|I|}\right)^2.$$

This formula is `packetConst_mul_mass_le` in `Filling.IsAdmissible`. It accounts for the fact that the second moment is compared with a definite genus mass, not just a count of the rows in the particular target representation. Combining the block comparison, Proposition 6.13, the scale factors, and the crude stratification bound yields the following sharp estimate. In contrast with a fixed-group spectral equidistribution statement, the constants here are uniform for the auxiliary r, k, f of this proof.

Proposition 6.14 (Formal-source input: sharp energy on the open congruence set)**.** *For every set Σ of chosen places above g_0, with specified nonzero fixed vectors, one has*

$$\mathcal{E}(A_{\mathrm{all}}; \Omega^\circ(\Sigma)) \le (1-b^{-4})(1+9b^{-4})D\, b^{-6k|\Sigma|}.$$

This is `energy_fixingSet_le` in `New/Filling/Moments.lean`. To obtain exactly this factor the coherent bound is taken with slack $(1+b^{-4})Db^{-6k|\Sigma|}$ before applying the two small error steps of Proposition 6.11. The resulting expression is at most $(1-b^{-4})(1+9b^{-4})Db^{-6k|\Sigma|}$ by the size bounds and $\tau^{20}D_1 \le Db^{-6kf}$, which follows from (81). The main analytical content of this proposition is precisely the uniform packet and local sieve inputs specified above. We have not replaced those inputs by a claimed classical trace formula theorem without quantitative estimates.

6.8 A finite-field positivity argument for singular minors

The open-set estimate of Proposition 6.14 does not yet imply a total energy estimate, since it leaves out the locus of singular minors. The next finite group argument shows that a nonnegative function of positive type cannot concentrate too heavily on such a locus unless it concentrates on fixed-vector subgroups. The distinction between the energy itself and a general positive-type function is useful because multiplying by a fixed-vector indicator preserves positive type, allowing induction.

Let G be a finite group, $\pi : G \to \mathrm{GL}_R(\mathbb{F}_q)$ a representation, and let A, A' be $R \times j$ matrices. Let $\Xi : G \to \mathbb{R}_{\geq 0}$ be of positive type; choose vectors u_g in a complex Hilbert space with $\langle u_a, u_b \rangle = \Xi(a^{-1}b)$ (the Gram representation follows directly from positive semidefiniteness). Write $S_\Xi = \sum_g \Xi(g)$. For any nonzero $x, y \in \mathbb{F}_q^j$, suppose

$$\sum_{g:,\pi(g)Ax=Ax} \Xi(g) \leq Y, \qquad \sum_{g:,\pi(g)^{-T}A'y=A'y} \Xi(g) \leq Y.$$

Lemma 6.15 (Finite-field determinant estimate)**.** *Under these assumptions,*

$$q \sum_{g:,\det(A'^T\pi(g)A)=0} \Xi(g) \leq 2S_\Xi + q^{j+R/2}Y.$$

Proof. Fix a nontrivial additive character ψ of $\mathbb{F}_q$. For nonzero y, x consider the weighted exponential sum

$$S(y,x) = \sum_g \psi(y^T A'^T \pi(g) A x)\Xi(g).$$

In $\ell^2(\mathbb{F}_q^R)$ let e_z be the delta basis and f_z its normalized Fourier basis with entries $q^{-R/2}\psi(z^T z')$. These are orthonormal bases by additive character orthogonality. For the tensors $U_1 = \sum_a f_{\pi(a)^{-T}A'y} \otimes u_a$ and $U_2 = \sum_b e_{\pi(b)Ax} \otimes u_b$, the squared norms are at most $|G|Y$: on expanding the inner products of the tensor sums, the basis inner products give exactly the indicators of the associated fixed-vector subgroups. Their mixed inner product is $|G|q^{-R/2}S(y,x)$, or its conjugate according to the Fourier convention, by reindexing $a^{-1}b$. Cauchy–Schwarz thus gives $|S(y,x)| \leq q^{R/2}Y$. If at least one of y, x vanishes then $S(y,x) = S_\Xi$.

Put $B_g = A'^T\pi(g)A$. Additive orthogonality gives $\sum_y q^{-j}\psi(y^T B_g x) = 1_{B_g x=0}$. Summing in x and weighting by $\Xi(g)$ shows

$$\sum_g \Xi(g)|\ker B_g| = q^{-j}\sum_{y,x} S(y,x) \leq 2S_\Xi + q^{j+R/2}Y.$$

On the singular locus, $|\ker B_g| \geq q$, and all summands are nonnegative. This proves the inequality. It reproduces `card_mul_sum_det_le` in `New/Heights/OneDeterminant.lean`. □

In the application $R = r$ and j is at most the size s_0 of the frame. A vector fixer indicator is of positive type, because it is the matrix coefficient of the permutation representation on the orbit of that vector. Multiplying it by Ξ gives another nonnegative function of positive type, by taking a tensor product of Gram representations. Thus one can successively impose new fixed-vector conditions. This is exactly the feature needed to derive the following result from Proposition 6.14.

Proposition 6.16 (Removing the unit-minor condition)**.** *Let $\Xi \geq 0$ be of positive type on $\mathfrak{G}$ and satisfy*

$$\sum_{s\in\Omega} \Xi(s) \leq \mathcal{E}(A_{\mathrm{all}}; \Omega)$$

for every set Ω of translations. Suppose its support consists of reductions fixing specified nonzero frame vectors over every place in a set Σ above g_0. Then

$$\sum_{s\in\mathfrak{G}} \Xi(s) \le (1+9b^{-4})D\, b^{-6k|\Sigma|}.$$

Proof. Use descending induction in $|\Sigma|$. If all places above g_0 belong to Σ, there is no additional minor condition, and the assertion follows from Proposition 6.14. In general suppose the assertion fails for Ξ at Σ. At each new place v outside Σ, intersect its support with the prescribed fixer of any nonzero frame vector x in the appropriate standard or dual frame. The restricted function remains of positive type, is dominated by the energy, and is supported on the extended set of fixers. By the induction hypothesis its total mass is at most $(1+9b^{-4})Db^{-6k(|\Sigma|+1)} < b^{-6k}S_\Xi$. Lemma 6.15 bounds the mass of each singular-minor locus by

$$\frac{2+q^{j+r/2}b^{-6k}}{q}S_\Xi \le 3b^{-10}S_\Xi,$$

where $q=g_0$, and the last inequality follows from $q\le b^{11}$, $q\ge b^{10}$, and $j+r/2\le k/2+2h+2$, which with the scale bounds gives $11(j+r/2)<6k$ (in particular $11(2h+2)<k/2$). The exact singular-minor test of `fixingSet` involves the standard and contragredient frame matrices, with one test for each of the leading minors at a chosen place; this identification is the finite matrix calculation in `sum_le_of_positive_type` of `New/Filling/Moments.lean`. Even bounding the number of tests by $2f(h+1)$, the excluded mass is at most $6f(h+1)b^{-10}S_\Xi \le b^{-4}S_\Xi$, by $f\le b/1000$, $h+1\le b$. Consequently at least $(1-b^{-4})S_\Xi$ lies in $\Omega^\circ(\Sigma)$. Domination by energy and Proposition 6.14 imply $(1-b^{-4})S_\Xi \le (1-b^{-4})(1+9b^{-4})Db^{-6k|\Sigma|}$, contradicting the supposition. This proves the claimed induction. □

Here we have used only the imported sharp estimate on the open set and finite group positivity; we have not assumed that the geometric energy is automatically of positive type. In fact the function

$$\Xi(s) = \|P^*\rho(s)P\|_{\mathrm{HS}}^2$$

is of positive type. To see this, with $T_0=PP^*$ put $u_s=\rho(s)T_0\rho(s)^*$ in the Hilbert space of Hilbert–Schmidt operators. Then $\langle u_a,u_b\rangle = \mathrm{Tr}(T_0\rho(a^{-1}b)T_0\rho(b^{-1}a)) = \Xi(a^{-1}b)$. It equals the energy per reduction by Lemma 6.9. With $\Sigma=\varnothing$ Proposition 6.16 therefore gives

$$\frac{1}{|\mathfrak{G}|}\sum_s \|P^*\rho(s)P\|_{\mathrm{HS}}^2 \le (1+9b^{-4})\frac{c^2}{n_0}. \tag{87}$$

This is where both local and global sharpness are indispensable.

6.9 First moment and bad finite-type loss

The estimates needed in addition to (87) are not general consequences of that second moment. They use the diagonal and central projections at the two different level primes. In fact

$$\sum_{i=1}^{c}\|(P^*P-I)e_i\| \le b^{-4}c, \qquad \|(I-R)P\|_{\mathrm{HS}}^2 \le 4b^{-4}c. \tag{88}$$

We explain the derivation and specify the further inputs.

For a central self-adjoint idempotent $z\in\mathbb{C}[\mathfrak{G}_i]$ and the inclusion $\jmath_i:\mathfrak{G}_i\to\mathfrak{G}$, let $M_z(s)$ be the $c\times c$ matrix with block from u to u' equal to

$$\nu^{-1}\sum_{a\in\mathfrak{G}_i} z(a)A_{G(E^+)\backslash S_B}(u,\jmath_i(a)^{-1}s,u').$$

The complement-coset energy controls this matrix at 1 after summing its row norms. Indeed convolution by the coefficients of z is an orthogonal projection on the regular representation and hence contracts the squared sum over s. To restrict to the relevant set Ω_i, note that this set is invariant under right multiplication by the embedded Bessel-compatible translations in the proof; more directly one only needs the support of the functions on the set in which their i-component is the Bessel reduction. The convolution estimate and the covariance under the Bessel subgroup give

$$|\bar{S}_B|\|M_z(1)\|_{\mathrm{HS}}^2 = \sum_{s\in\bar{S}_B} \|M_z(s)\|_{\mathrm{HS}}^2 \le \mathcal{E}(A_{G(E^+)\setminus S_B};\Omega_i).$$

The equality uses the following exact covariance calculation, a *Formal-source input* in the proof of `sum_norm_amplitudeSum_le` in `New/Filling/Moments.lean`. Each $s \in \bar{S}_B$ induces a permutation of the column classes by the rational Bessel element over (u, s, u'). The left and right covariance of the Bessel functions replaces the corresponding block by a unitary matrix acting on its coordinate space, times a scalar of modulus one. Norms are unchanged because the finite-dimensional compact types are unitary. This is derived from `amplitudeSum_move` in `New/Filling/AmplitudeSums.lean` and `exists_besselS_mem_termsOver` in `NearTerms.lean`. The convolution estimate is applied after keeping the terms whose factor-i reduction lies in the image of the Bessel subgroup; the exact group-module contraction, together with this covariance, is the assertion `sum_norm_amplitudeSum_le` cited above. In particular the resulting sum of row norms is at most $b^{-4}c$ (using $\sqrt{c}\|M_z(1)\|_{\mathrm{HS}}$ and Proposition 6.12). Taking $z = 1$ and using the identity term gives the first inequality of (88); bounding row or column norms is equivalent for the self-adjoint $P^*P - I$.

The remaining statement is the trace of a small central projection on the Bessel-subgroup packet. If z is central and $\Re(z(1)) \le b^{-8}/2$, then

$$\Re \sum_{a\in\mathfrak{G}_i} z(a) \sum_{u,j} A_{S_B}(u, \jmath_i(a)^{-1}, u)_{j,j}/\nu \le b^{-4}c. \tag{89}$$

This is the *Formal-source input* `re_trace_besselS_le` in `Moments.lean`. It is not a trace formula on arbitrary cusp forms. Its proof realizes the class sum on the finite orbital integrals over the Bessel subgroup, using invariance of z under conjugation and the local character. The ell^2 norm of the coefficients of z is exactly $z(1)$ since $z = z^* = z^2$ in the regular group algebra. The remaining orbital coefficient has squared norm at most $2c^2$ after the measure normalizations (the lower-floor estimate in `New/Filling/FixedLines.lean`); hence Cauchy–Schwarz bounds the trace by $c\sqrt{2z(1)} \le b^{-4}c$. The crucial nontrivial part is the bound on the orbital coefficient, incorporated in (89). Applying this with $z = 1 - e_i$ and Lemma 6.2, the trace of $(1-\rho(e_i))PP^*$ is at most $2b^{-4}c$, since the non-Bessel part contributes at most the same bound $b^{-4}c$ from the previous paragraph. The two projections commute, so $1 - R \le (1-\rho(e_0)) + (1-\rho(e_1))$ as positive operators. Taking their traces against PP^* proves the second inequality in (88).

The notation of the formal proof for these two inputs is deliberately separate: `re_trace_besselS_le` controls the Bessel term, whereas `sum_norm_amplitudeSum_le` controls the complement for any central projection, including the identity. Both are needed even though the discarded relative Plancherel mass itself is tiny.

6.10 From the moment bounds to the covering theorem

We now give the final argument without any further automorphic estimate; it is a general Hilbert-space inequality. First, W is nonzero. If $R = 0$ then $\|P\|_{\mathrm{HS}}^2 \le 4b^{-4}c$ by (88), but the

first bound in (88) gives $\operatorname{Tr}(P^*P) \geq (1 - b^{-4})c$. These two inequalities contradict $c > 0$ and $b \geq 1000$. Thus $n > 0$.

For $s \in \mathfrak{G}$ put $T_s = \rho(s)PP^*\rho(s)^*$ and $\overline{T} = |\mathfrak{G}|^{-1} \sum_s T_s$. Unitarity and the trace identity show

$$\begin{aligned} \|\overline{T}\|_{\mathrm{HS}}^2 &= |\mathfrak{G}|^{-2} \sum_{s,t} \operatorname{Tr}(T_s T_t) \\ &= |\mathfrak{G}|^{-1} \sum_s \|P^*\rho(s)P\|_{\mathrm{HS}}^2. \end{aligned}$$

If $J \subseteq W$ and p_J is its orthogonal projection in V, set $U_s = (R\rho(s)P)^{-1}J$. For each unit vector $u \in U_s$, the two orthogonal components of $\rho(s)Pu$ lie respectively in J and $\ker R$. Summing over an orthonormal basis of U_s and enlarging the sums for these components gives

$$\operatorname{Tr}(p_{U_s} P^*P) \leq \operatorname{Tr}(p_J T_s) + \|(I - R)P\|_{\mathrm{HS}}^2.$$

In this computation p_{U_s} is the orthogonal projection within C and R commutes with $\rho(s)$. For any orthogonal projection p_U in C, one has

$$|\operatorname{Tr}(p_U(P^*P - I))| \leq \sum_i \|(P^*P - I)e_i\|,$$

by expanding in the coordinate basis, applying Cauchy–Schwarz to each inner product, and using $\|p_U e_i\| \leq 1$. Applying this with $U = U_s$ gives

$$\dim U_s \leq 5b^{-4}c + \operatorname{Tr}(p_J T_s).$$

Averaging and using Hilbert–Schmidt Cauchy–Schwarz, $\|p_J\|_{\mathrm{HS}} = \sqrt{\dim J}$, and (87), we obtain

$$\begin{aligned} \frac{1}{|\mathfrak{G}|} \sum_s \dim U_s &\leq 5b^{-4}c + \sqrt{\dim J}\sqrt{(1 + 9b^{-4})c^2/n_0} \\ &\leq \left(5b^{-4} + \sqrt{(1 + 9b^{-4})\alpha}\right) c \\ &\leq \tfrac{1}{2}(1 + \alpha + 19b^{-4})c, \end{aligned}$$

where $\alpha = \dim J/n$. The second line uses $n \leq n_0$, and the third uses $2\sqrt{uv} \leq u + v$. Thus there exists s attaining the indicated bound. This is `Fits.exists_translate_finrank_comap_le` in `New/Filling/Moments.lean`, with the underlying Hilbert–Schmidt variance argument `finrank_sub_le_re_inner_dyadSum` in `New/Heights/OneDeterminant.lean`. The constant 19 comes from the projected and identity estimates ($2 \cdot 5$) and from the second moment (9).

Finally, ramify the columns by applying $D_N = \prod_{w|\pi}(1 - z_w)$, with z_w the generator of the finite central-unit action at a selected place above N. Its image is C^{ram}, since in a cyclic finite group the image of $1 - z_w$ is the kernel of the average over the group, and these operators at different places commute. In the source the estimate `ramifyingOp_finrank_comap_and_short` in `New/Columns/ChainValuation.lean` shows that for any subspace $U \subset C$ the dimension of its inverse image under D_N is at most $\dim U + fc/(N - 1)$: the central-unit action permutes classes freely in cycles of $N - 1$, so the kernel has relative dimension at most $1/(N - 1)$ at each place. The freeness and compatibility with the column type are the formal-source input used here; the dimension inequality also follows immediately from rank-nullity for D_N from its kernel bound. Admissibility gives $19b^{-4}/2 + f/(N - 1) \leq b^{-2}$, since $N - 1 \geq b^9$ and $f \leq b/1000$. Inserting this loss and removing the common nonzero normalization scalar proves (67). This completes the derivation of the spread property appearing in `Data.ramifiedColumns_fill_rows` in `New/Filling/FillingTheorem.lean`.

It remains to convert spread into a small number of translates. Put $\eta = b^{-2}$ and start with $J = 0$. If the current codimension $q_* = n - \dim J$ satisfies $q_* > 4\eta n$, choose s by (67). By rank-nullity applied to the quotient map from C to W/J, the increase Δ in dimension obtained by adjoining $P_s(C^{\mathrm{ram}})$ satisfies

$$\Delta \geq c\left(\frac{q_*}{2n} - \eta\right) \geq \frac{cq_*}{4n} > 0.$$

In particular this is a genuinely new translate. Define a potential on the remaining codimension by

$$\Psi(q_*) = 5n \log\left(1 + \frac{q_*}{\eta n}\right).$$

When $q_* > 4\eta n$, adjoining the new image decreases the potential by at least

$$\Psi(q_*) - \Psi(q_* - \Delta) = -5n \log\left(1 - \frac{\Delta}{q_* + \eta n}\right) \geq \frac{5n\Delta}{q_* + \eta n} \geq c,$$

since $q_* + \eta n \leq 5q_*/4$. The process terminates because dimension increases by a positive integer at every step. The potential remains nonnegative; hence the number of selected translates times c is at most $\Psi(n) = 5n \log(1 + b^2)$. The final codimension is at most $4b^{-2}n$. This proves Theorem 6.1. The greedy covering is the result `exists_few_translates` of `New/Heights/OneDeterminant.lean` (the preceding derivation is a complete mathematical proof of that deduction).

For the application, take a spanning set of short integral columns as in Theorem A, and for the above finite set T select independent vectors from its projected images after the rational rescaling. The span has the dimension supplied by Theorem 6.1, and the accumulated logarithmic index bound is at most $5n \log(1 + b^2)B$. Apply the individual squared-length estimates to these independent vectors and Hadamard's inequality to their lifted Gram matrix. The loss from the small missing fraction of dimensions is controlled in the final comparison by the crude bound for the heights. No further spanning theorem is needed. The substantive assumptions for B have thus been reduced from the theorem itself to the precise uniform Bessel system, lattice block, trace transfer and local orbital/Poisson estimates; the finite-field counting, second-moment positivity, variance, and greedy selection are explicit deductions from those estimates.

7 The determinant lower bound

We turn to the lower bound which is used in conjunction with the shortness and spanning theorems. In this part there is a second pair of mechanisms, both involving a comparison with theta lifts for an auxiliary finite-order character. The first comparison converts a nearly integral family of rows into a height estimate for the ratio of two full-dimensional determinants. The second compares the unknown determinant of a chain of lifts with determinants of single lifts, and then replaces each of these unknown determinants by a product of a special-value ratio and an integral comparison determinant. It is the last step which repays the cost of the ordinary (or idle) differentiation directions in the upper bound. The existence of adequate twists, a separate arithmetic theorem, will be considered subsequently.

We write $n = \dim_{\mathbb{C}} W > 0$, and retain $B = fa(\log b)^3$, $L = B \log b$, $H = mk\mathcal{U}$. Here the cost of one direction is the particular real number

$$\mathcal{U} = \sum_{\tau \in \Phi} (J(w_0(\tau)) - A(w_{\mathrm{nom}}(\tau), \tau)). \tag{90}$$

In particular it is not a positive-definite quadratic form or a norm. None of the inequalities below assumes its nonnegativity. Put $B_0 = B/(mk)$. The object of this section is the following precise estimate.

Theorem 7.1 (C)**.** *Let $y_1, \dots, y_q \in W$ be independent, where $q \geq (1 - 4b^{-2})n$. Suppose that* $\mathrm{IndexLe}(y, d_{\mathrm{ind}})$ *holds and that, in the coefficient-conjugate metrics of the setup,*

$$\mathrm{covol}_\sigma^2(\Theta_1 y_1, \dots, \Theta_1 y_q) \leq e^V \quad (\sigma \in \mathrm{Aut}\,\mathbb{C}).$$

Then

$$n(H - 3L) \leq V + 2d_{\mathrm{ind}}. \tag{91}$$

The statement to be reconstructed is `Data.integral_forms_long` in `New/Endgame/LongForms.lean`. The precise estimates below designated *Formal-source input* are taken from named declarations of the compact development. We give ordinary mathematical proofs of the intervening algebraic and determinant implications, and isolate the statements of the more demanding automorphic and local computations. In particular the occurrence of a theorem from the formal source as an input is not an assertion that its independent verification or its translation from a published theorem has been completed.

7.1 Conventions for norms of conjugate coefficient vectors

Fix a level on a definite unitary group. In the coefficient model of algebraic modular forms, a form is a family of vectors in weight representations indexed by a finite adelic class set, with the appropriate covariance. For the Weyl coordinate representation used in the compact code, the pairing has a rational positive-definite matrix G (the matrix denoted `weylGramOver`). Write the class-sum pairing in schematic form as

$$\langle u, v \rangle = \sum_{h \in X} a_h \, \overline{u(h)}^{\,T} G v(h), \qquad a_h = |\,\mathrm{Aut}(h)|^{-1}.$$

Here X is a choice of class representatives at the specified level; different levels give different class sets and pairings. At the levels where the class stabilizers are trivial, $a_h = 1$. At other levels one clears the stabilizer orders by a uniform integer. The normalization concerning class sums is essential: it is not the usual probability measure on the finite class set.

For $\sigma \in \mathrm{Aut}(\mathbb{C})$ the norm used in the construction is

$$\|u\|_\sigma^2 = \sum_h a_h \, \overline{\sigma(u(h))}^{\,T} G \sigma(u(h)),$$

with σ applied entrywise. If u is a tuple of forms, the associated positive Gram matrix is formed from these conjugated vectors. It is in general incorrect to identify it with σ applied to the original Gram matrix. With κ denoting complex conjugation and $\sigma^* = \kappa\sigma\kappa$ as field automorphisms, one has instead

$$\sigma\langle u, v \rangle = \sum_h a_h \, \overline{\sigma^*(u(h))}^{\,T} G \sigma(v(h)). \tag{92}$$

This follows from the rationality of the coefficient pairing and from $\sigma(\bar{z}) = \overline{\sigma^*(z)}$. Thus

$$|\sigma\langle u, v \rangle| \leq \|u\|_{\sigma^*} \|v\|_\sigma.$$

Much of the emphasis below on simultaneous estimates for *all* field conjugates is due to this distinction, not to the number of complex embeddings of a fixed coefficient field. These conventions match `LevelForm.normSqAt`, `LevelForm.pairingMatrix` and `LevelForm.covolumeSq` in `Lib/Adelic/CuspForms/LevelForms.lean`.

Lemma 7.2 (Mixed Gram determinant and completion)**.** *Let $v = (v_i)_{i=1}^{q}$ be a tuple in a finite coefficient representation as above. If its squared covolume is bounded by e^{V} at every conjugate, then*

$$|\sigma \det(\langle v_i, v_j \rangle)| \le e^{V} \quad (\sigma \in \operatorname{Aut} \mathbb{C}).$$

If $z_1, \dots, z_h$ satisfy $\|z_j\|_\sigma^2 \le e^{A}$ for every σ, the corresponding bound for the tuple (v, z) is e^{V+hA}.

Proof. Regard the scaled coefficient vectors as vectors in one ordinary finite-dimensional Hermitian space, so that the form displayed above is its Gram form. Equation (92) identifies the conjugate of the determinant with the scalar product of the exterior products of the σ^*-tuple and the σ-tuple. Cauchy–Schwarz in the exterior power bounds its modulus by the square root of the product of their squared covolumes, at most e^{V}. Equivalently, one uses

$$|\det(X^*Y)|^2 \le \det(X^*X)\det(Y^*Y)$$

for two rectangular matrices of equal column size. For the completion bound, if the conjugated v_i are dependent, the extended Gram determinant vanishes. Otherwise perform Gram–Schmidt beginning with these q vectors. Each subsequent squared length is bounded by the original squared length of the added vector. Hence

$$\det \operatorname{Gram}_\sigma(v, z) \le \det \operatorname{Gram}_\sigma(v) \prod_{j=1}^{h} \|z_j\|_\sigma^2.$$

The source proves these assertions as `norm_map_det_gram_le` and `covolumeSq_elim_le` in the same file. □

We will also use the lattice property of the coefficient pairing. Let $\phi : \mathbb{C} \to \mathbb{C}_\ell$ be a field embedding and Λ_ϕ the lattice of forms of Weyl gauge at most one at ϕ. The coordinate lattices at the individual points can be calculated in integral adelic frames, rather than in an arbitrary fixed complex coordinate basis. Changing a rational frame is compatible with the Gram pairing. Consequently, for $u \in \Lambda_{\phi\circ\kappa}$, $v \in \Lambda_\phi$ the local pairing of the two coordinate vectors is integral. At the levels under consideration the class stabilizers are trivial and this implies

$$|\phi\langle u, v\rangle|_\ell \le 1. \tag{93}$$

Indeed after passing to integral frames the pairing becomes a sum of terms $\phi(\bar{u}_i)G_{ij}\phi(v_j)$ with integral entries; the integral class-sum property then follows from the ultrametric inequality. The source identifies these frames by `exists_mul_mem_integralAbove` and its Weyl gauge change law. We use their existence and compatibility as *Formal-source inputs*: `gramForm_mem_embValRing`, `norm_inner_le_one` (`Lib/Adelic/CuspForms/LevelForms.lean`), and `Data.classAut_eq_bot_of_le_chosenLevel` (`New/Data/ChosenPlaceLevels.lean`, used in `New/Endgame/LongForms.lean`). Notice that integrality at just one embedding for both arguments would not justify (93).

7.2 Heights with varying finite-place entries

In the direct application, different primes can require different twists. The height device allows the local entries of a family to vary, while preventing the arbitrary insertion of prime-dependent normalizing factors in the archimedean term. Let

$$I = \{\iota = (\ell, \iota_\ell) : \ell \text{ a rational prime},\ \iota_\ell : \mathbb{C} \simeq \overline{\mathbb{Q}}_\ell \text{ a field isomorphism}\}.$$

The ℓ-adic absolute value is normalized by $|\ell|_\ell = \ell^{-1}$. These isomorphisms are not topological. Their use as field isomorphisms follows from the axiom of choice and cardinalities of transcendence bases: both algebraically closed fields have cardinality and transcendence degree $2^{\aleph_0}$ over a countable algebraic closure of $\mathbb{Q}$. In particular any embedding of a fixed number field in either algebraically closed field extends to the requisite field isomorphism.

Definition 7.3. For a family $u = (u_\iota)_{\iota \in I}$ of complex numbers and $c \in \mathbb{R}$ write $\mathrm{Ht}_{\le c}(u)$ if there exist $Q \neq 0$ in $\mathbb{C}$, a real number α, and a real family h_ℓ indexed by primes such that

$$\begin{aligned} |\sigma Q| &\le e^{\alpha} && (\sigma \in \operatorname{Aut}\mathbb{C}), \\ |\iota_\ell(Qu_\iota)|_\ell &\le e^{h_\ell} && (\iota \in I), \\ \alpha + \sum_{\ell < M} h_\ell &\le c && (M \text{ sufficiently large}). \end{aligned}$$

The last sum runs over primes. We allow the h_ℓ to be signed; this is an eventual bound for prime cutoffs, not an assertion of absolute summability. This is `HeightLe` in `New/Heights/Heights.lean`.

Thus Q is a single nonzero number, uniform in ι. Even if a height bound is negative, it need not contradict a product formula when the entries u_ι differ. The construction requires precisely such a negative bound for the inverse volume ratios, to which the common determinant calculation will attach their reciprocal.

Lemma 7.4 (Height calculus)**.** *The following assertions hold.*

1. *If $|\iota(u'_\iota)|_\ell \le |\iota(u_\iota)|_\ell$ for all ι, then $\mathrm{Ht}_{\le c}(u)$ implies $\mathrm{Ht}_{\le c}(u')$; the constant c can be increased.*
2. *From $\mathrm{Ht}_{\le c}(u)$ and $\mathrm{Ht}_{\le c'}(u')$ follows $\mathrm{Ht}_{\le c+c'}(uu')$. In particular estimates add under finite products, and $\mathrm{Ht}_{\le jc}(u^j)$ for $j \in \mathbb{N}$.*
3. *If $|\iota(u_\iota)|_\ell \le e^{h_\ell}$ for every ι over ℓ, and $\sum_{\ell \in F} h_\ell \le c$ for all finite sets of primes F, then $\mathrm{Ht}_{\le c}(u)$.*
4. *If $\mathrm{Ht}_{\le c}(1)$ holds for the constant family, then $c \ge 0$.*

Proof. The first assertion uses the same witness. The second uses the product witness QQ' with archimedean bound $\alpha + \alpha'$ and finite-place bounds $h_\ell + h'_\ell$. The eventual prime cutoff works by taking the maximum of the two cutoffs. The third uses witness 1. For the last, the boundedness of all field conjugates of Q already implies that Q is algebraic: if it were transcendental, isomorphisms defined by extension of transcendence bases could send Q to a transcendental element of arbitrarily large absolute value. Put $F_0 = \mathbb{Q}(Q)$, $s = [F_0 : \mathbb{Q}]$, and $z = N_{F_0/\mathbb{Q}} Q \in \mathbb{Q}^\times$. Taking products over all embeddings of F_0 into $\mathbb{C}$ and into $\overline{\mathbb{Q}}_\ell$ respectively gives

$$\frac{1}{s} \log |z| \le \alpha, \qquad \frac{1}{s} \log |z|_\ell \le h_\ell.$$

Choose the eventual cutoff M large enough that $|z|_\ell = 1$ for every $\ell \ge M$. The rational product formula gives

$$0 = \frac{1}{s} \left(\log |z| + \sum_{\ell < M} \log |z|_\ell \right) \le \alpha + \sum_{\ell < M} h_\ell \le c.$$

The named source lemmas are `HeightLe.mono`, `HeightLe.mul`, `HeightLe.prod`, `HeightLe.pow`, `HeightLe.of_le` and `HeightLe.nonneg`. □

A useful consequence, repeatedly used below, is that if Q_ι is a positive integer prime to ℓ then inserting an arbitrary integral power of Q_ι into u_ι does not change its height bound: its image has local absolute value one. No new archimedean witness is required. In the application Q_ι is the norm of the twist conductor, not a fixed normalization scalar.

7.3 The determinant lemma with index

Let $T_\theta : W \to V_\theta$ be injective linear maps of finite-dimensional complex spaces with the prescribed metrics, for example Θ_θ. Define the squared volume

$$Z_\theta = \det(T_\theta^* T_\theta) > 0.$$

For a basis $v = (v_i)_{i=1}^n$ of W write $G_\theta(v) = \mathrm{Gram}(T_\theta v)$. A change of basis from an orthonormal basis to v shows

$$\det G_\theta(v) = Z_\theta \det \mathrm{Gram}(v), \qquad \frac{Z_\theta}{Z_1} = \frac{\det G_\theta(v)}{\det G_1(v)}. \tag{94}$$

Here it is important to use the same basis of the row space for every twist. The proof of Theorem C uses the following form of the determinant argument; its applicability does not require an integral basis of all row forms at a fixed finite place independent of the twist.

Lemma 7.5 (One determinant)**.** *Suppose v is a full basis and $|\sigma \det G_1(v)| \le e^\alpha$ for every σ. Fix a family θ_ι of twists and a family of nonzero scalars γ_ι such that $\mathrm{Ht}_{\le g_0}((\gamma_\iota\bar\gamma_\iota)^{-1})$. Let h be a nonnegative integer. Suppose the following holds for every ι over ℓ and every embedding $\phi : \mathbb{C} \to \mathbb{C}_\ell$: a matrix M_ϕ combines $T_{\theta_\iota} v$ into coefficient-integral rows at ϕ, with*

$$|\phi(\det M_\phi)|_\ell \ge e^{-e_\ell} |\phi(\gamma_\iota)|_\ell^h, \qquad \sum_{\ell \in F} e_\ell \le d_{\mathrm{ind}}$$

for every finite prime set F. Assume that the integral lattices pair as in (93). Then

$$\mathrm{Ht}_{\le \alpha + 2d_{\mathrm{ind}} + hg_0}\left(\frac{Z_{\theta_\iota}}{Z_1}\right). \tag{95}$$

Proof. Take $\phi = \iota_\ell$ and its composition with κ; the two matrices are not required to be the same. After applying ι_ℓ, the matrix of mixed pairings of the new rows is integral by hypothesis, and in the complex model it is

$$\overline{M_{\phi\circ\kappa}} G_{\theta_\iota}(v) M_\phi^T.$$

All its entries have image of absolute value at most one, hence so does its determinant by the Leibniz formula. Evaluating the determinant product and using the two determinant lower bounds gives

$$|\iota_\ell(\det G_{\theta_\iota}(v)(\gamma_\iota\bar\gamma_\iota)^h)|_\ell \le e^{2e_\ell}.$$

In view of (94), using $\det G_1(v) \ne 0$ as witness shows that $(Z_{\theta_\iota}/Z_1)(\gamma_\iota\bar\gamma_\iota)^h$ has height at most $\alpha + 2d_{\mathrm{ind}}$. Multiply by the hth power of the assumed inverse-square family, and apply Lemma 7.4. This is `one_determinant` in `New/Heights/OneDeterminant.lean`. □

The coefficient 2 in front of the index is therefore forced by the Hermitian determinant and is not a generous error allowance. The source treats the general situation in which the chosen lattices vary with ι, θ_ι, and ϕ. The quantifier over *all* embeddings ϕ here is intentional. In particular, an index bound at ι_ℓ only, in the absence of one at $\iota_\ell \circ \kappa$, is insufficient to reach (95).

7.4 Special-value data and fit twists

It will be helpful to give the exact analytic quantities to which the determinants are compared. Let M denote the set of chosen split places above $g_0 g_1$ (not the analytic cutoff on primes), and for a twist ϑ let $Q(\vartheta)$ denote its support, disjoint from M and its conjugate in the sense of the tame level datum; write $N(Q) = \prod_{w \in Q} Nw$. A *pair* $p = (\nu, z)$ of the proof consists of a tame unitary character ν of the norm-one torus of weight $-c_\tau$, unramified away from M, and an integer z with $1 \le z \le t$ and $z \le c_\tau \le 4b$. Tame part is denoted ψ. Define

$$\rho_p(\vartheta, s) = \frac{|L^{M \cup Q}(z, (\nu \vartheta^s)^{-1})|}{|L^M(z, \nu^{-1})|}, \qquad s \in \{-1, 1\}. \tag{96}$$

The removed local factors are to be understood in the source convention, including the conjugate places for the Hecke L-functions on the norm-one torus. The lower point p_v is the same pair with z replaced by v, for $1 \le v \le z$. For an integer $e \ge 0$ put

$$\Pi_S(\eta, e) = \left(\prod_{w \in S} Nw \right)^{e^2} \prod_{v=1}^{e} |L^S(v, \eta)|^2.$$

This is precisely `lValueProd` of `Lib/SiegelWeil/Tower/LargeSideConstantTerm.lean`; the ratio of two products at e satisfies

$$\frac{\Pi_{M \cup Q}((\nu \vartheta^s)^{-1}, z)}{\Pi_M(\nu^{-1}, z)} = N(Q)^{z^2} \prod_{v=1}^{z} \rho_{p_v}(\vartheta, s)^2. \tag{97}$$

Indeed each factor agrees and the first prefactor is a ratio of products of residue cardinalities over disjoint sets. The definitions are `Data.Pair.ratio`, `Pair.lowerPoint` and the identity `Pair.lValueProd_eq` in `New/Exchange/LValueRatio.lean` and `New/Exchange/TriangleLevelExponent.lean`.

The needed nonvanishing of these special values uses the weights of the pairs and their particular Hecke characters. We do not use GRH for them. As a *Formal-source input*, the declaration `ratio_ne_zero_and_real` of `New/Exchange/TriangleLevelExponent.lean` gives that every ratio here is nonzero and real; this follows in the code from `AutChar.HasWeight.L_ne_zero`. This property allows reciprocal families of ratios to enter the height calculus without exceptional zero entries.

The local projection condition. For an auxiliary pair $O = (W_j, V_j)$ with $\dim_E W_j = r$, let V_O denote its algebraic form space on the definite group. At each split place w outside $M \cup Q$ and its conjugate, there is a normalized Hecke polynomial operator $H_w(X)$ on V_O. The test point is

$$X_w = (a_O(w)\, Nw^{\mathrm{rankDiff}(O)-1})^{-1},$$

where $a_O(w)$ denotes the Frobenius character value in the normalization of `AuxPair.a.frobValue` for rank r. We use the precise operator in `AuxPair.heckePolyOnForms`. The definition of X_w is `AuxPair.testValue` in `New/Rows/StackedDeterminants.lean`. If e_0 denotes the chosen central projector and $Z \subset V_O$ the space of lifts, set

$$F_w = \operatorname{im} e_0 \cap \ker H_w(X_w) \cap Z^{\perp}.$$

The points X_w are designed so that genuine lifts belong to the first two factors. A vector is *tested* if it is orthogonal to F_w at every such place; the false-lift subspaces depend on w. This is `AuxPair.IsTested` in `New/Selmer/FitTwists.lean`. The orthogonal complement $Z^{\perp}$ is computed with the ordinary complex pairing.

For a fixed isomorphism ι_ℓ, Z has denominator bound β when the following holds for $d_Z = \dim Z$: for every $u_1, \dots, u_{d_Z}$ integral at every ℓ-adic embedding, and every $v_1, \dots, v_{d_Z}$ integral at ι_ℓ and tested,

$$\left|\iota_\ell \det(\langle u_i, p_Z v_j\rangle)\right|_\ell \le e^{d_Z r \beta}. \tag{98}$$

Here p_Z is the complex orthogonal projection. This is the actual declaration `DenominatorLe` in the same file. The projection is only applied to the second family; equality with the pairing of both projections may be used at the ordinary complex metric. In this definition it would be substantially stronger to ask for an entrywise denominator bound for every integral vector, or to drop the word tested. Neither strengthening is assumed.

An m-fold twist θ has first m powers still nontrivial at every prime of its support, each power with the same level support. The family θ_ι is fit if:

1. $\ell \nmid N(Q(\theta_\iota))$ for ι over ℓ;
2. for every $j \in [1, m]$ there exist numbers β_ℓ (allowed to depend on j) such that $\sum_{\ell \in F} \beta_\ell \le B_0$ for all finite prime sets F, and the good single-lift subspace at twist θ_ι^j has denominator bound (98); the same β_ℓ bounds the rank-one lift lines of all pairs of rank-one type at this index. For the latter the rank in (98) is one;
3. for all j, both signs s, and all pairs p ramified exactly at M,

$$\mathrm{Ht}_{\le x^{999}}\left(\rho_p(\theta_\iota^j, s)^{-1}\right).$$

Here rank-one type means $m < z \le r$, tame part ω^j, and $w_j + r - t \le c \le w_j$ at all infinite places (with the rank index condition in `Data.Pair.IsRankOne`); note that the term rank-one refers to a lift *from* $U(1)$ and not the dimension of its algebraic coefficient representation. At a fixed j the same finite-place costs must handle all these auxiliary pairs. These conditions are precisely `Data.IsFitFamily` in `New/Selmer/FitTwists.lean`. The existence result `exists_isFitFamily` in `New/Selmer/FitTwistsExist.lean` is a *Formal-source input* here; its substantial arguments are developed separately later in the paper.

Definition 7.6 (Cost of an exchange factor)**.** For a complex scalar $c(\theta)$ depending on the m-fold twist, say it has cost s squares if for every fit family one has

$$\mathrm{Ht}_{\le 2stx^{999}}\left(|c(\theta_\iota)/c(1)|^{-2}\right).$$

This description is used below only for nonzero factors. It is `Data.CostsSquares` in `New/Exchange/CostOfSquares.lean`.

Lemma 7.7. *Costs add under multiplication (and may be increased). A factor whose squared modulus equals a fixed positive number times an integral power of $|\mathcal{O}_E/\mathfrak{m}(\theta)|$, for $\mathfrak{m} = M^a Q(\theta)$ with fixed a, costs zero squares. If a factor c satisfies*

$$|c(\theta)|^2 = C\Pi_{M \cup Q(\theta)}((\nu\theta^j)^{-1}, z),$$

with $C > 0$ independent of θ, and $p = (\nu, z)$ ramified exactly at M, then c costs one square.

Proof. The first assertion is Lemma 7.4. The cardinality of the residue ring is a fixed integer times $N(Q(\theta))$, so in the second assertion the squared modulus ratio is a local unit at the relevant ℓ; the constant height witness 1 suffices. For the last assertion apply (97) and the fit-family condition to each of the z inverse ratios, twice. The factor $N(Q)^{-z^2}$ is a local unit and does not alter the bound. Since $z \le t$, the resulting constant is at most $2tx^{999}$. These are the rules expressed in `CostsSquares.mul`, `CostsSquares.of_norm_sq_eq_mul_card_zpow` and `exists_isTameFor_costsSquares_one`. □

7.5 The correspondence of complementary rank and the finite square

There are two uses of an exchange identity: finding a spanning family of crude rows, and factoring the volume on the good row space. We give their common input in some detail. For Hermitian spaces V, V', A let $I_A^V(\xi, \Phi)$ be the theta integral over $[U(V)]$, with real datum ξ and finite datum given by Φ on the matrix residue ring. The source makes this precise as a mixed real–finite datum `modDatum`, with splitting characters on the two members having product the prescribed rank twist. The input character ν has weights μ_τ and tame finite character χ modulo a level ideal $\mathfrak{m}$. The normalized character on the large space of rank N is denoted $\mathrm{rankTwist}(\nu^{-1}, N)$; it includes a power $\lfloor N/2 \rfloor$ of the fixed splitting character. We do not suppress this normalization in comparing the special values.

Proposition 7.8 (Complementary-rank exchange; formal-source input)**.** *Let $N = \dim_E A = \dim_E V + \dim_E V'$, $\dim_E V = \dim_E V' + 2e$, $e \geq 0$. At the infinite places the negative signs of V, V' add up to those of A; write this number as q_A. Assume $2q_A + 2 \leq N$ and $N \leq 2\mu_\tau + 2q_A$ at every infinite place. Fix the index identification of the complementary blocks and a rapid real datum v. Then there are a rapid real datum x and a constant C independent of the varying modulus and character such that, for every admissible tame character $(\mathfrak{m}, \nu, \chi)$, there is $c \neq 0$ with*

$$|c|^2 = C\Pi_{\mathfrak{m}}(\nu, e), \qquad I_A^V(x, \Phi) = cI_A^{V'}(v, \Psi)$$

whenever the finite data Φ, Ψ are partners. Here $\Pi_{\mathfrak{m}}$ means the product defined above with the places of $\mathfrak{m}$.

This is `exchange_theorem` in `New/Exchange/ExchangeTheorem.lean`. Its proof in the formal development reduces the nonzero constant at equal rank to the calculation `exists_equalRankThetaRatio_value` in `New/Exchange/EqualRankValue.lean`, and then applies `exchange_of_partners` in the same file. Along a hyperbolic plane the surviving constant term of the lifted character has a scalar given by local Gauss sums and a Tate integral; the size of the total scalar under repeated reduction is precisely the product Π. Its factors have the conductor exponent e^2 since adjoining successive isotropic lines adds exponents $1, 3, \ldots, 2e-1$. The actual matching of these integrals and nonconstant terms is not established by this counting identity. The tower and constant-term results in `Lib/SiegelWeil/Tower/LargeSideConstantTerm.lean`, `New/Exchange/ExchangeDifference.lean` and `New/Exchange/EqualRankValue.lean` are essential background to the displayed theorem. We take the complementary-rank theorem with its stated hypotheses as an input, rather than simply appeal to a general Siegel–Weil identity without verifying normalization or ramified data.

The passage from such an equality of integrals to an exchange of operators is important. With two exterior Hermitian spaces U_1, U_2, combine them to play the role of A. The two intermediate spaces have ranks differing by $2e$. Suppose the characters on all four edges are matched so that their products are the rank twist on $U_1 \oplus U_2$. Opposite choices of signs are used in the descent edges. Pairing the composition of a lift and a descent with a functional on the exterior target gives an integral on the product of the exterior groups, with a theta integral over the intermediate group. This is the situation of Proposition 7.8. Inserting the complementary-rank equality changes the order to descent then lift through the other intermediate group.

The source takes special care about integrability in this argument, because some of the involved groups are indefinite. Its lemma `exists_integrable_kernel_bound` bounds, for rapid vectors on the exterior groups, the integral of their products against a theta kernel by a constant times a weighted norm of the real datum. The permitted real datum on the other side is obtained as a limit in that norm and approximated by finite sums of elementary tensors.

The closedness of a one-dimensional span in the resulting space of periods then gives a common scalar. The approximating datum is not necessarily a single product of two edge vectors. This is why the resulting scalar independent of the twist is not asserted nonzero in the general square law. The exact square assertion is `one_square` in `New/Exchange/OneSquare.lean`; we use it, including this potentially vanishing scalar, as a *Formal-source input*.

Here is the finite-ring part of this assertion. Let $R = \mathcal{O}_E/\mathfrak{m}$. For a prescribed pair of rectangular matrix index sets and an identification σ_0 of complementary dimensions, partners are the finite data in `IsPartner`; they are defined with the character twist and complementary full-rank blocks. The function of the edge datum on one side of a square is $F(X)$, and on the opposite side $F'(Y)$ (these are complex numbers after applying a specified functional). In the source a square law of scalar c_0 consists of the following three conditions:

$$\begin{aligned} F(kX) &= \chi(\det k)F(X), \\ F'(Yl) &= \chi(\det l)F'(Y), \\ \sum_Y \Psi(Y)F'(Y) &= c_0 \sum_X \Phi(X)F(X) \quad ((\Phi,\Psi) \text{ partners}). \end{aligned}$$

Here k, l are invertible matrices of the appropriate sizes. In the application $c_0 = \kappa_0/c(\theta)$ with $c(\theta) \neq 0$ and κ_0 independent of θ. This definition is `SquareLaw` in `New/Exchange/SquareLaw Fourier.lean`.

The actual periods to be moved use point data and inclusion matrices. The square law is applied not to a single arbitrary point on one side and a single arbitrary point on the other, but to appropriate averages of points. For example the finite transformation called `nextDatum` sends a matrix function p to the function which, on block matrices

$$Y = \begin{pmatrix} y & * \\ 0 & I \end{pmatrix},$$

takes value $p(-y)/|\operatorname{Mat}_{n'\times b'}(R)|$, and is zero on other matrices; here n', b' are just the relevant block sizes. For this particular change the identity `SquareLaw.sum_eq_nextDatum` reads

$$|\operatorname{GL}_{a'}(R)| \sum_y p(y) F'\begin{pmatrix} y \\ I_0 \end{pmatrix} = c_0 |\operatorname{GL}_{n'+b'}(R)| \sum_x p^+(x) F\begin{pmatrix} I_0 & x \end{pmatrix}.$$

The inclusion matrices I_0 have the sizes prescribed by the block decomposition. This is a *Formal-source input* from `New/Exchange/SquareLawFourier.lean`. Its shape reflects invariance under elementary block changes and averaging over the unconstrained $*$ block. Crucially, a determinant character is evaluated only on invertible blocks; their values are roots of unity.

To make the transport of a finite family of data precise, consider a grid of spaces and put one residue matrix on each oriented edge $e = (P, P')$. If X_e is the matrix on that edge, a family of vertex matrices k_P acts by

$$X_e \longmapsto k_{P'}^{-1} X_e k_P.$$

A functional Λ on tuples of residue matrices is called tidy of order v on a set of edges $\mathcal{E}$ if it satisfies character covariance for admissible families k of vertex matrices,

$$\Lambda(f \circ k) = \left(\prod_{e\in\mathcal{E}} \chi_e\big(\det k_{P'} / \det k_P\big) \right) \Lambda(f),$$

and if it is a sum of at most Q_*^v point evaluations with nonzero coefficients roots of unity, supported on tuples whose matrices on the distinguished ramified edges have left inverses. Here

$$Q_* = |R|^{(2n_m)^2}, \qquad n_m = k + m(t+1)$$

is the bound for the number of matrices on any one edge. The precise admissible vertex families are those satisfying the endpoint level conditions and the prescribed diagonal block relation between the two grids. This is `IsTidy`, whose short-sum component is `IsShortSum`, in `New/Exchange/GridMove.lean`. The distinction between the full tidy datum and a merely level-invariant datum matters at the next square.

One elementary observation in the formal source is that orders add: if, for each point S, its evaluation is replaced by at most Q_*^w evaluations with root-of-unity coefficients supported on valid new tuples, then a tidy short-sum of order v becomes a short-sum of order $v + w$, provided covariance is checked for the new tuple action. Indeed expand the original sum and then the replacement sum. The number of terms is at most $Q_*^v Q_*^w$, and all nonzero new coefficients are products of the permitted roots of unity. The covariance must be verified separately; it does not follow from this count. The counting statement is `IsShortSum.bind` in the same file.

The actual replacement at a square is `IsTidy.exists_square_move` (`New/Exchange/GridMove.lean`), a *Formal-source input*. Its hypotheses are the square law, the matching identifications on the edge blocks, and the compatibility of an admissible vertex family with those block embeddings. Its proof uses the finite kernel δ of `graphKernel` in `New/Exchange/SquareLawFourier.lean`. For a matrix $k \in \mathrm{GL}_N(R)$ and a decomposition $N = a' + b'$, write x_k for the last a' rows of k^{-1} and y_k for the first b' columns of k. The kernel supported on such pairs is

$$\delta(x_k, y_k) = \frac{\chi(\det k)}{|\,\mathrm{GL}_{b'}(R)|}.$$

It has the following useful property: if $\Psi(yl) = \chi(\det l)^{-1}\Psi(y)$ for $l \in \mathrm{GL}_{b'}(R)$, then

$$\sum_y \delta(x_k, y)\Psi(y) = \chi(\det k)\Psi(y_k).$$

In fact for fixed x_k its complementary columns span $\ker x_k$, a free direct summand, and its full-rank column frames are $y_k l$ as l runs through $\mathrm{GL}_{b'}(R)$. Replacing k by $k\,\mathrm{diag}(l, I)$ keeps x_k fixed, transforms y_k by l, and multiplies the determinant character by $\chi(\det l)$. The two character factors cancel term by term, proving the formula. Changes in the other complementary block that preserve both data change k by a block unipotent, so the definition is unambiguous. This is `sum_graphKernel_mul` in the source. It explains the cancellation of the averaging denominator: without the covariance, individual coefficients $1/|\,\mathrm{GL}_{b'}(R)|$ would eventually destroy the lattice estimate. The square move uses further block support and equivariance identities to normalize the graph kernel so that nonzero coefficients are roots of unity. This last assertion is included in the formal-source square-move input.

7.6 The two-grid exchange and the crude rows

We now describe the endpoint transport used for the first of the two height estimates. The grid has two families of vertices; in the notation of the source a vertex is (z, i, j) with $z = 0, 1$ and $0 \le i, j \le m$. Their ranks are

$$N(z, i, j) = r + z + i(t + z) + j(t - z). \tag{99}$$

The grid with $z = 0$ has equal rank increments t on both axes, whereas in the second grid the increments are $t + 1$ and $t - 1$. Thus the two diagonal vertices $(0, j, j)$ and $(1, j, j)$

have ranks differing by one. The sign count at offset zero is $\lfloor (m-j+z)/2 \rfloor$; in particular the endpoint $j = m$ is definite in both grids. The initial and final diagonal pairs differ by a line of sign $(-1)^j$. At the real places, the Harish–Chandra parameters are the initial r-block together with blocks built from the weights w_l of the construction and their negatives. The spacing $h = 2r^8$ ensures the required gap conditions and the terminal parameters interlace. These rank, sign, gap and interlacing assertions are the *Formal-source inputs* `card_gridIndex`, `card_neg_gridSigns`, `gridDiagEquiv_spec`, `hasGap_gridParam` and `gridParam_interlace` in `New/Exchange/GridParameters.lean`; (99) itself follows by counting the disjoint blocks.

In grid z, the path of stage j starts at $(z, 0, m)$, descends through the vertical edges to $(z, 0, j)$, then lifts along horizontal edges to (z, j, j). The union of these two paths is the path of stage j. A pair of forms on the two starting definite groups and a tidy datum on the path gives a period at the two diagonal endpoints: one of the forms is restricted through the codimension-one Hermitian embedding, with the real translation and line multiplier specified by the grid, and then multiplied by the other and integrated.

The reason for using the second grid is the following separation result.

Proposition 7.9 (Starting periods; formal-source input)**.** *There exist tidy data Λ_θ of order $4m$ on the stage-zero path, a fixed real translation, a scalar α, and a linear surjection R_0 from the specified endpoint form space on the second grid onto the full initial principal-level row form space, such that, writing ϕ for a form in the domain and Ξ for a form used to test the first endpoint,*

$$[K_f : J_m(\theta)] \int_{[U(Y_m)]} \Xi(u)\ \mathrm{lift}_\theta(R_0\phi)(u)\, du = \alpha\ \mathrm{gridPeriod}_0(\Lambda_\theta; \Xi, \phi).$$

Here $J_m(\theta)$ in the index is the level `Data.levelY` *at m as in the period statement; the integral uses exactly the source's bilinear, not Hermitian, period normalization. The map R_0 and real translation are independent of θ.*

The declaration is `Data.exists_isTidy_integral_eq_gridPeriod` (`New/Exchange/CrudeRows.lean`). The surjectivity input in its proof is stronger than existence of nonzero periods for some forms: by `Data.eq_zero_of_orthogonal_gridLevelForms` in `New/Exchange/RowsReached.lean`, there is one real translation for which a form orthogonal to all the restricted forms on the second grid is zero. It employs the real gap construction `Data.exists_isSteep` and the restriction theorem `eq_zero_of_petersson_restrict_eq_zero`. The descent from the definite endpoint onto that grid's level forms is onto by `Data.gridLevelForms_eq_map_iotaPlus`; one uses injectivity of the adjoint weighted lift chain to prove surjectivity. Thus a fixed linear surjection, rather than a different map for each row, is available.

For $j < m$, transporting the two paths from stage j to stage $j+1$ requires commuting a descent past j lifts in each grid and moving across the remaining diagonal configurations. The source packages this precisely in `Data.hasRoute_succ` in `New/Exchange/CrudeRows.lean`, a *Formal-source input*: the cost and increase in the tidy order are at most $4j+2$, and the initial period equals a scalar independent of the twist times the inverse of a nonzero factor of that cost times the new period (in the direction fixed by the source's `HasRoute`). Its proof uses the row ladder `Data.descent_past_gridPathMap` in `New/Exchange/RowLadder.lean` and the elementary square move. Iterating the statement is elementary: the sums of costs and of orders are bounded by

$$\sum_{j=0}^{m-1} (4j+2) = 2m^2.$$

Products of scalar factors respect the bound by Lemma 7.7. Therefore the start period can be transferred to the terminal stage with one factor $c(\theta)$ costing $2m^2$ squares, and with tidy endpoint data of order $4m + 2m^2$. This is `Data.gridPeriod_climb`; in particular the switch to the terminal endpoint is not achieved with exponentially many independent normalization factors.

At the terminal diagonal pair (z, m, m) both unitary groups are definite. The interlacing noted after (99) gives a bilinear branching pairing of their Weyl coefficient spaces, represented by an integer matrix. The terminal composite consists of two definite chains, each beginning at $(z, 0, m)$ and lifting to (z, m, m). Let $\mathrm{endPeriod}(\Lambda, y)$ be the resulting class-integral pairing, linear in the tidy datum and bilinear in the endpoint forms $y = (y_0, y_1)$, as defined in `Data.endPeriod` (`New/Exchange/GridMove.lean`). The terminal levels are chosen so that the coefficient integration can be computed as a finite class sum with the branching matrix. For later use the following are the precise uniform estimates.

Proposition 7.10 (Terminal periods; formal-source input)**.** *There is a positive integer D_*, independent of the twist, with*

$$D_* \le \exp(x^{99} b^2)$$

such that $D_* \,\mathrm{endPeriod}(\Lambda, y)$ is integral at every $\ell \nmid N(Q)$ whenever Λ is tidy on the terminal path and both endpoint forms have gauge at most one and are fixed by the terminal endpoint levels. For the trivial twist and tidy data of order $4m+2m^2$, if every coefficient-conjugate value of the two input forms has squared norm at most S_0, S_1, respectively, then at any $\sigma \in \mathrm{Aut}\,\mathbb{C}$,

$$|\sigma(\mathrm{endPeriod}(\Lambda, y))| \le \exp(x^{200} b^2)\sqrt{S_0 S_1}.$$

These are `Data.exists_mul_endPeriod_mem_embValRing` and `Data.endPeriod_le` in `New/Exchange/CrudeRows.lean`. The first clears the denominators of the successive theta kernels with the standard vacuum integrality statements, and the final branching factor with an integer involving powers of the level and N_1. The bound `boxes_endWeightData_le` bounds the algebraic weights and their box counts. For the second estimate the same finite branch weights bound the size of all derivatives and the branching pairing; the class numbers at each endpoint are bounded using `card_stdClasses_le_exp_mul_log`, and the tidy-data bound counts the residue terms. Although the trivial twist is used for this archimedean estimate, every nontrivial twist is covered by the finite-place estimate away from its own conductor. Thus one does not need an archimedean norm bound growing uniformly in all the permissible twist conductors.

We record the result of the terminal-period calculation with its precise quantifiers.

Proposition 7.11 (Crude rows; formal-source input)**.** *There is a nonzero function $c(\theta)$ of the twists costing $2m^2$ squares, and a spanning set $\mathcal{C}_0$ in the full principal-level row space (before good projection), such that for $y \in \mathcal{C}_0$:*

1. *$(c(\theta)/c(1))\,\mathrm{lift}_\theta(y)$ is coefficient-integral at every embedding over every $\ell \nmid N(Q(\theta))$;*
2. *every coefficient-conjugate value of the untwisted lift has squared Weyl norm $\le \exp(x_1^2 b^2)$.*

The declaration is `Data.crude_rows` (`New/Exchange/CrudeRows.lean`). Its construction combines Proposition 7.9 and the route calculation with Proposition 7.10, choosing spanning short integral forms at the second definite endpoint. For the first definite endpoint use delta forms as coefficient tests: if the pairing with every integral coefficient test of a vector lies in the valuation ring, the dual Weyl lattice criterion gives its integrality. The branching matrix and denominators in the dual pairing are accounted for in the common rescaling of the row forms.

The source obtains a common normalization independent of θ besides the ratio $c(\theta)/c(1)$; this uniformity is part of the proposition and indispensable for the one-determinant argument.

To apply the proposition, project the spanning set to good rows with the integral normalized projector. Its integrality assertion is `goodProjTranslate_mem_integralForms` and its metric bound is `normSqAt_goodProjTranslate_le` in `New/Rows/MonomialCoefficientVector.lean` (both are used in `New/Endgame/LongForms.lean`). These are *Formal-source inputs* concerning the given finite levels. Select $n-q$ projected members extending $y_1,\dots,y_q$ to a basis; this is possible by choosing a basis of the quotient of W by their span from the projected spanning set. Their integral normalization is $\gamma_\iota = c(\theta_\iota)/c(1)$. Its inverse squared norm admits height bound

$$g_0 = 4m^2 t x^{999}.$$

Complete the local recombining matrix for the original q rows by a diagonal block $\gamma_\iota I_{n-q}$ to make the new basis integral for the same twist. The determinant condition of Lemma 7.5 holds, with $h = n-q$.

The initial archimedean upper bound for the extended full determinant is of the form $V + (n-q)A_*$, obtained from Lemma 7.2 after conversion of the crude value bound to the class-sum norm. The exact size estimate, which is a calculation with the class-count bound, the normalized projector and the admissible magnitudes, gives

$$A_* + 4m^2 t x^{999} \le \tfrac{1}{4} b^3.$$

This is the final per-vector numerical input in `Data.heightLe_volume_rows` (`New/Endgame/LongForms.lean`); it is a *Formal-source input* as regards the bounds for the coefficient models. Its scale is natural: the leading crude exponent $x_1^2 b^2$ is $o(b^3)$ since $x_1^{10^6} \le b$, and the remaining polynomial losses and class logarithms are much smaller than b^3. The final step does not require these terms separately to be smaller than the sharp error L.

Write from now on $Z(\theta) = \det(\Theta_\theta^* \Theta_\theta)$; injectivity, used for the determinant denominator, is proved below. Applying Lemma 7.5 and using $n - q \le 4n/b^2$ gives

Proposition 7.12 (Forward ratio)**.** *For the indicated family of short rows and fit family θ_ι, one has*

$$\mathrm{Ht}_{\le V + 2d_{\mathrm{ind}} + nb}\left(\frac{Z(\theta_\iota)}{Z(1)}\right). \tag{100}$$

Proof. The preceding determinant calculation gives a bound $V + 2d_{\mathrm{ind}} + (n-q)(A_* + g_0)$. Bound its last term by $(4n/b^2)(b^3/4) = nb$. This is the implication from the precise crude-row theorem to the forward-ratio estimate. □

7.7 The triangle of weighted lifts

The chain determinant is not a multiplicative determinant of ordinary square matrices: the Hermitian spaces increase in dimension and the lift maps take their values in proper subspaces. We need an exchange computation to make the determinant on the initial row space factor. The formal source uses an abstract triangle which we now describe. Fix a twist and let $F_{i,v}$ be spaces on the triangular array $i + v \le m$. Use arrows

$$a_{i,v} : F_{i,v} \longrightarrow F_{i,v+1}, \qquad d_{i,v} : F_{i,v+1} \longrightarrow F_{i+1,v}.$$

The $v = 0$ spaces contain isometric copies $j_i(U)$ of a common Hilbert space U of forms. Let $w_v > 0$ be the weights (integral finite group orders in the application). Define

$$A_{i,v} = (w_{v-1} a_{i,v-1}) \cdots (w_0 a_{i,0}), \qquad A_{i,0} = 1.$$

The first-arrow adjoint law, on $j_i(U)$ and all target vectors z, is

$$\langle z, a_{i,0} j_i(u)\rangle = \langle d_{i,0} z, j_{i+1}(u)\rangle.$$

Moreover $d_{i,0}a_{i,0}j_i(u) = j_{i+1}(B_i u)$ for an injective $B_i : U \to U$. For the other squares the hypotheses contain a scalar $\beta_{i,v}$ with

$$w_v a_{i+1,v} d_{i,v} = \overline{\beta_{i,v}}^{-1} w_{v+2} d_{i,v+1} a_{i,v+1}. \tag{101}$$

The square scalar is potentially zero in the initial exchange statement. For norms, the source has an additional adjoint version of the square, including an adjoint vector for a horizontal arrow, and a factor of cost zero. All these conditions together are precisely `IsTriangle` in `New/Exchange/TriangleNorms.lean`; (101) is interpreted with the source's zero-inverse convention when necessary. The induction below on injectivity does not divide an unknown zero scalar.

The weighted chain satisfies the commuted-descent identity

$$A_{i+1,v} d_{i,0} a_{i,0} = \frac{w_{v+1}}{w_1 w_0 \overline{\beta}}\, d_{i,v} A_{i,v+1} \tag{102}$$

for a product of square factors (with the normalizations of the source). Here at $v = 0$ the factor is $1/w_0$. This follows by applying (101) as one moves the descent past each arrow in order; the finite-rank weight ratios telescope. Successive adjoint laws in the norm case give the compatible identity

$$\langle z, a_{i,v} A_{i,v} j_i u\rangle = \kappa \langle d_{i,v} z, A_{i+1,v} j_{i+1} u\rangle \tag{103}$$

with the product $\kappa\beta$ costing at most v squares. In the general injectivity case only (102) is used. These are the two assertions of `IsTriangle.weightedChain_descent`. The source proves their simultaneous induction beginning with the first-arrow adjoint law; each adjoint square supplies one factor, and the extra zero-cost factor leaves the prescribed budget unchanged. We take the full normalized adjoint identities as *Formal-source inputs* from that declaration.

Lemma 7.13 (Injectivity of the weighted chain)**.** *For $i + v \le m$, the restriction of $A_{i,v}$ to $j_i(U)$ is injective.*

Proof. Use induction on v. This is clear for $v = 0$. If $A_{i,v+1} j_i u = 0$, apply (102): $A_{i+1,v} d_{i,0} a_{i,0} j_i u = 0$. Since its argument is $j_{i+1}(B_i u)$, the induction hypothesis and injectivity of B_i give $u = 0$. This proves `IsTriangle.eq_zero_of_weightedChain`. □

Lemma 7.14 (Gram identity of the triangular array)**.** *Suppose a subspace $W \subset U$ is preserved by the first-arrow Gram operators S_i characterized on W by*

$$w_0 \langle a_{i,0} j_i u, a_{i,0} j_i v\rangle = \langle u, S_i v\rangle.$$

In the adjoint case of `IsTriangle`, for $i + v = m$ there is a scalar η of cost at most v^2 such that the Gram form of $A_{i,v} j_i$ on W equals

$$\eta \frac{w_0 w_1^v}{w_v} \langle u, T_i v'\rangle, \qquad T_i = S_i S_{i+1} \cdots S_{m-1},$$

where v' is the second vector, and the product is the identity for $v = 0$. In particular $\det T_i = \prod_{h=i}^{m-1} \det S_h$.

Proof. The first-arrow adjoint law implies $w_0 B_i = S_i$ on W when paired against vectors of W (equivalently one uses its preserving operator on the full common space). For the inductive step take the Hermitian pairing of the last arrows, use (103) on one side and (102) on the other. The factors of w_v, w_0, w_1 combine to give $w_0 w_1^{v+1}/w_{v+1}$, with scalar multiplier contributed by $\kappa\beta$ at cost at most v. Apply the induction hypothesis in row $i+1$ to S_i and the remaining chain, using self-adjointness of its Gram form to put S_i in the first argument of the endomorphism product. The determinant is independent of the resulting product order, so even without commutativity the conclusion about determinants holds. Starting with cost zero at length zero and increasing the bound by no more than $(v+1)^2 - v^2$ proves the stated cost. This is the induction in `IsTriangle.inner_weightedChain_eq`; the adjoint normalizations and the exact scalar in this calculation are the ones in (102)–(103). □

For the application one uses the triangle on the left vertical strip of the preceding grids, with a varying offset i. At offset i and stage v the space of forms is that of the real parameter `gridParam` and signature `gridSigns`; the initial forms are identified by negating signs and using the isometry `formNegEquiv`. The weights are $w_v(\theta) = \mathrm{triangleOrder}(\theta, v)$, computed by the order of the finite group of the relevant modulus. The prescribed real vectors harmonics and local inclusion data define the arrows `triLift`. The first-arrow adjoint law follows from the usual pairing of lift and descent; the finite normalizing order is important for the square identity. The first-arrow injectivity uses the Rallis-type norm calculation of `hasSum_inner_triLift`, with the chosen large rank. The theorem `exists_isTriangle` in `New/Exchange/TriangleArrows.lean` verifying these requirements is a *Formal-source input*. The additional adjoint condition for the Gram identity holds in the $z = 0$ grid.

To relate these operators to the single lifts we need the Hecke transfer between groups with different signatures. For $1 \le j \le m$, let U_j denote the good form space on the auxiliary group of rank r with $m - j$ negative signs, and let $V_j(\vartheta)$ denote the definite algebraic form space of the auxiliary pair of ranks $(r, r+t)$. Write

$$u_{j,\vartheta} : U_j \to V_j(\vartheta)$$

for the good lift `Data.goodLift`, which is the adjoint of a normalized theta descent. We reserve the letter W for the good rows on Y_0, not for the Hermitian space of the same symbol in the code. The following precise transfer statements are used.

Proposition 7.15 (Rallis pairing and transfer; formal-source inputs)**.** *For every j, the map $u_{j,\vartheta}$ is injective for every allowed twist ϑ, and $\dim_{\mathbb{C}} U_j = n$. The pairings defining the squared norms of the first arrow in row $j-1$ of the triangle and the auxiliary lift u_{j,θ^j} are, up to the explicit constants and finite group orders, sums over the same family of Hecke terms. In particular they are represented by one endomorphism in the common off-level Hecke algebra acting on the good form spaces. For every element P of that algebra, the characteristic polynomials of its actions on W and U_j agree.*

The first-arrow formula is `hasSum_inner_triLift` (`New/Exchange/TriangleArrows.lean`). The auxiliary-lift formula is `Data.auxLift_injective_and_hasSum_inner` (`New/Exchange/CharacterLiftNorm.lean`): there is $A_j \neq 0$ independent of ϑ with

$$\langle u_{j,\vartheta} x, u_{j,\vartheta} y \rangle = A_j \delta_{j,\vartheta} \sum_{\gamma} R_{j,\vartheta}(x, y; \gamma).$$

Here $\delta_{j,\vartheta}$ is `AuxPair.gramFactor`, and the terms `Data.rallisTerm` are explicitly the Hecke coset matrix coefficients, with their character factors and divided by the number of $(r+t) \times r$ matrices over the level residue ring. The Rallis identity has a stable-range hypothesis and uses

a chosen sharp local datum; the usual abstract nonvanishing of a theta lift alone would give neither this sum nor its transformation under level. We rely on the source's precise identity.

The fact that the series of Hecke terms represents one element simultaneously on the two spaces is `exists_common_hecke_rallis`, in `New/Trace/GoodFormTransfer.lean`. This follows from an elementary finite-dimensional closure step once the local coset expression is known: the images of the common algebra in the product of the two endomorphism spaces, and their corresponding pairings, are finite dimensional and closed. Each common coset term is in this image after averaging over the principal-level subgroup. The convergent sum stays in the image. It need not define a convergent sum in an infinite-dimensional topology on the Hecke algebra itself.

For the characteristic polynomials the essential theorem is `trace_heckeOnGood_eq` in the same file. Its proof isolates each occurring eigensystem by a projector in the common algebra, and uses primitive label rigidity (`New/Trace/LabelRigidity.lean`) to rewrite the good idempotent, on the eigensystem, in terms of functions supported at classes where the simple trace formula applies. The trace comparison is a transfer result on the specific inner forms, with the cuspidal type and gap assumptions; it is not implied by the same rank of two groups. Once the traces agree for all P, however, the rest of the conclusion is ordinary algebra. Taking $P = 1$ shows equality of dimensions. The traces of all powers P^h agree and determine its characteristic polynomial by Newton's identities. This proves the determinant comparison `det_heckeOnGood_eq` and the rank identity `finrank_rows_eq_finrank_goodForms`. The assertions of trace comparison and primitive label rigidity here are formal-source inputs; the elementary polynomial implication was just proved.

Let

$$\lambda_j(\theta) = \det(u^*_{j,\theta^j} u_{j,\theta^j}) > 0.$$

The positive quantities λ_j cannot be assumed algebraic merely because they are determinants of spaces of forms; the height inequalities are statements about chosen ratios with uniform normalization scalars. The inner product on the final form space of the main chain is a class-sum inner product. Relative to the Petersson metric of the triangular chain there is a conversion factor

$$J(\theta)\, D_m, quad J(\theta) = [K_0 : J_m^{\mathrm{res}}(\theta)], \quad D_m = \#\,\mathrm{WeylIndex}(Y_m, \kappa_m).$$

The identity `stdLevelIndex_lastResLevel_eq` in `New/Exchange/ChainNorms.lean` expresses $J(\theta)$ as

$$CN(Q)^{k_0} \delta_{1,\theta}^m w_m / (w_0 w_1^m).$$

Here k_0 denotes the (integral) exponent called k in that declaration; it is unrelated to the rank of columns. The constants C and k_0 do not vary with the twist. This count compares block parabolic stabilizers in the quotient and is a *Formal-source input*; it cancels the triangular Gram factor $w_0 w_1^m / w_m$ exactly. The formula `inner_iotaAtLevel` in the same file gives the metric conversion. Injectivity follows from Lemma 7.13. Taking determinants of the resulting Gram comparison and applying Proposition 7.15 stage by stage yields the following identity.

Proposition 7.16 (Volume factorization)**.** *For every allowed m-fold twist, Θ_θ is injective; there exists a function $c(\theta)$ of cost m^2 squares such that*

$$Z(\theta) = |c(\theta)|^n \prod_{j=1}^{m} \lambda_j(\theta) > 0. \tag{104}$$

This is `Data.volume_iota_eq_prod_liftVolume` (`New/Rows/RowVolume.lean`). To check the scalar exponent, multiplying a Hermitian Gram form by a scalar multiplies its determinant

by the nth power of that scalar; the factor in (104) is $|c|^n$, not $|c|^{2n}$. The residual factors resulting from the metric conversion, finite group orders and *delta* factors have squared norm equal to a fixed constant times an integral power of $N(Q)$, so they cost zero by Lemma 7.7. All that remains is the triangular cost of m^2. This explains both the injectivity presupposed in the forward-ratio calculation and the factorization required for the reverse ratio. The normalized scalar identities in this calculation are precisely those of the cited formal-source theorem.

7.8 Comparison forms for one lift

We next establish the height estimate for $\lambda_j(1)/\lambda_j(\theta_\iota)$ for each j. Fix j and abbreviate $u_\vartheta = u_{j,\vartheta}$, $Z_\vartheta = \operatorname{im} u_\vartheta$. In the next few paragraphs Z_ϑ is a subspace, not the scalar $Z(\theta)$. The strategy is to use prescribed algebraic forms in the definite target, not to estimate the unknown Petersson norms of every automorphic representation separately.

Three character pairs are required for this particular j. Let $e_+ = m$, $e_- = m/2$, and $e_0 = 7(m/2) + k/2$ (the admissible parities make these integers). Write $p_i = (\nu_i, e_i)$ for $i \in \{0, +, -\}$. The pairs $p_\pm$ have character $\upsilon^{-e_\pm}\chi^{\pm j}$, where υ is the fixed rank-normalizing character and χ is the chosen tame character. The character in p_0 is $\mu_j^0 \upsilon^{j+\lfloor (t-11m-1)/2 \rfloor}$; here μ_j^0 is the untwisted character determining the jth lift. Use the twist ϑ positively for p_0, p_+ and negatively for p_-. Put

$$\Gamma_j(\vartheta) = \prod_{i=0,+,-} \Pi_{M\cup Q}((\nu_i \vartheta^{s_i})^{-1}, e_i), \quad s_0 = s_+ = 1, \quad s_- = -1. \tag{105}$$

This is `Data.heckeValueProd`, from `New/Rows/Separation.lean`; that file gives the precise pair definitions as `Data.linePair`. The long string p_0 will yield the term $k\mathcal{U}$, whereas the two other strings contribute to the small error.

Proposition 7.17 (Comparison maps; formal-source input)**.** *There is a finite-dimensional complex table space T, independent of ϑ, a surjection $A : T \to U_j$, and for every twist a linear map $g_\vartheta : T \to V_j(\vartheta)$ and $\gamma_\vartheta \neq 0$ such that*

$$u_\vartheta^* g_\vartheta = \gamma_\vartheta A, \ |\gamma_\vartheta/\gamma_1|^2 = \Gamma_j(\vartheta)/\Gamma_j(1). \tag{106}$$

The vectors $g_\vartheta A^* v$ are tested relative to $Z_\vartheta^\perp$ for all $v \in U_j$. At every $\ell \nmid N(Q)$ and embedding ϕ, $g_\vartheta\zeta$ is integral whenever the conjugate tables $\bar\zeta$ are integral at $\phi\circ\kappa$ with the prescribed Weyl gauges of their classes. Finally

$$h_j \|g_1\zeta\|_\sigma^2 \le e^{2rB_0} \sum_\alpha \|\sigma\zeta_\alpha\|_{\mathrm{Weyl}}^2,$$

where $h_j > 0$ is the hook number of the small algebraic weight.

This is exactly `exists_comparison_forms` in `New/Rows/ComparisonForms.lean`. We describe the detailed construction to specify its content. Split the definite Hermitian space in the auxiliary pair into $V'' \oplus V^\circ$ (the `DividedSplitting` data), and take tables indexed by a finite member index and by representatives of the small principal level $\bar\pi_g^2$ on V°. Their entries lie in the small Weyl module of weight given by `colsMinus`; thus h_j is the hook number of this module. To a member index corresponds a split lattice obtained by gluing the two blocks in the split model along the residue matrix. The same glued lattice can be chosen for every twist. The group in the large block which fixes the specified rows modulo $\bar\pi_g^2 Q$ has free class action; this is `classAut_rowFixingLevel_eq_bot`. If $y_{\vartheta,i}(h)$ is the line multiplier for member i, the comparison map is the algebraic form

$$g_\vartheta\zeta = g^{\mathrm{loss}} \sum_{(i,c)} \sum_{h\in X_i} \mathrm{productForm}_{L_i}((h,c), \overline{\beta_1 y_{\vartheta,i}(h)}, \zeta_{i,c}).$$

Here L_i is the glued lattice, X_i is the class set for the row-fixing level on the large block, β_1 is one nonzero common line scale, and the recipe called `productForm` applies the Cartan product of the representations in the divided splitting. This is the actual `comparisonForm`. The exponent loss is the one in the divided splitting; it clears the gluing loss from off-diagonal blocks. The source proves integral and small bounds for the line multiplier after applying β_1 in `Data.lineMultiplier_integral_and_small`, and controls the Cartan product in `New/Rows/CartanIntegrality.lean`. Multiplying a delta form by this product and summing over a free class set preserves finite-place integrality. The archimedean estimates use the number of tables, classes and weights, and multiply the line bound with the squared Weyl norm of the input table. The output estimate is `normSqAt_comparisonForm_le` and the input integrality is `comparisonForm_mem_integralForms`. These assertions concerning the glued finite and infinite kernels are formal-source inputs.

The resulting inner product with a lift of $F \in U_j$ has a common period expression independent of the twist up to scalar: it is a weighted sum over members i of

$$\int_{[U(W_j)]} \overline{F(u)F_i^\natural(u)}\,\theta^\circ(\zeta_{i,\bullet})(u)\,du.$$

Here $F_i^\natural$ is a fixed natural form, and θ° the theta kernel from the small block. The scalar is determined by `Data.cosetKernelSum_eq_fNatural`, which gives $F_i^\natural$ as the coset kernel sum involving the line multiplier at every twist times κ_ϑ satisfying

$$|\kappa_\vartheta|^2 = C/\Gamma_j(\vartheta).$$

This is initially an assertion potentially with $C = 0$. The equality of period expressions is `inner_adjoint_comparisonForm`; the nonvanishing and surjectivity use the separation theorem `Data.eq_zero_of_integral_fNatural_tableTheta`. This states that if a good F gives zero in every displayed period, then $F = 0$. In fact if $\kappa_\vartheta = 0$, all $F_i^\natural = 0$ and the separation theorem forces $U_j = 0$, contrary to $n > 0$. If the range of $u_1^* g_1$ were not all of U_j, a nonzero orthogonal vector would annihilate all the periods, again contradicting separation. Set $A = u_1^* g_1$ and $\gamma_\vartheta = \overline{\kappa_1/\kappa_\vartheta}$. The period identity gives the first equation of Proposition 7.17; the quotient of the two norm identities for κ gives (106). This explains the proof of nonzero scalars and the fact that a *single* common surjection works for every twist. The equality of coset kernel sums and the separation theorem are *Formal-source inputs* in `New/Rows/Separation.lean`.

The tested condition in the proposition is equally substantive. Off level, the Hecke polynomial of a genuine lift has a product of t known roots and an affine rescaling of the original Hecke polynomial. Explicitly, if the auxiliary ranks are $r, r+t$, the polynomial at a split place w is

$$\prod_{v=0}^{t-1}(1 - q_w^{v+r/2}a_O(w)X)P(q_w^{t/2}a_O(w)b_O(w)^{-1}X)$$

in the character convention of `Data.lifts_auxLift` (`New/Exchange/CharacterLiftNorm.lean`); $a_O(w), b_O(w)$ in that display indicate the source's rank-normalized Frobenius values, and P is the source Hecke polynomial. In particular the test operator vanishes on the genuine lift, as expressed by `heckePolyOnForms_testValue_eq_zero`; one must use the test point in its precise source normalization, not insert a point from a different convention. A false lift satisfies this Hecke test and lies in the good part but is orthogonal to every actual lift. Its theta descent to the smaller group has zero good part (by adjointness and equivariance).

The crucial vanishing of the comparison periods for a false lift comes from enlarging the descent group by hyperbolic planes. The good polynomial labels imply that the reflection sums

associated with the chosen tame character act by zero. Together with the vanishing of the initial descent and nontriviality of the character at the twist places, this kills the relevant boundary terms in the constant-term expansion. At a split place off level, a bounded descended form killed by a Hecke polynomial at a sufficiently small test point must vanish: expand the Hecke polynomial, use the sum of norms of the nonconstant terms strictly below one, and compare the sup norm. The precise statements `AuxPair.descentsVanish_of_reflectionSum`, `eq_zero_of_heckePolyOp_eq_zero` and `descentsVanish_add_planes` in `New/SiegelWeil/AuxPairVanishing.lean` and `New/SiegelWeil/DescentsVanishPlanes.lean` formalize these steps. The period which results from the glued product is zero by `DividedSplitting.stdClassSum_inner_productForm_eq_zero` in `New/Rows/MemberOrthogonality.lean`. Applying that statement to the comparison construction gives `tableTheta_adjoint_eq_zero`: for a false lift Ξ, each small form $\theta^\circ((g_\vartheta^*\Xi)_{i,\bullet})$ vanishes. Consequently the common period formula gives

$$\langle \Xi, g_\vartheta A^* v\rangle = \langle A g_\vartheta^* \Xi, v\rangle = 0$$

for each v and every false-lift space. These special vanishing assertions are *Formal-source inputs*; the last implication is merely the adjunction identity. The source does not prove that arbitrary integral vectors satisfy the tested condition.

7.9 Compressed Gram determinants

Let p_ϑ be projection onto Z_ϑ, with respect to the class-sum inner product at the ordinary embedding. Define on Z_ϑ

$$D_\vartheta = \det(h_j p_\vartheta g_\vartheta g_\vartheta^*|_{Z_\vartheta}).$$

The key determinant calculation is independent of automorphic theory. Regarding u_ϑ as an isomorphism onto Z_ϑ, adjunction and Proposition 7.17 give

$$h_j u_\vartheta^* g_\vartheta g_\vartheta^* u_\vartheta = h_j|\gamma_\vartheta|^2 AA^*.$$

Taking determinants and using $u_\vartheta^* p_\vartheta = u_\vartheta^*$ gives

$$\det(u_\vartheta^* u_\vartheta) D_\vartheta = h_j^n |\gamma_\vartheta|^{2n} \det(AA^*). \tag{107}$$

Since A is onto, both sides are nonzero and g_ϑ^* restricted to Z_ϑ is injective. Dividing the identities at the trivial and arbitrary twist yields

$$\frac{\lambda_j(1)}{\lambda_j(\theta_\iota)} = \frac{D_{\theta_\iota^j}}{D_1}\left(\frac{\Gamma_j(1)}{\Gamma_j(\theta_\iota^j)}\right)^n. \tag{108}$$

This elementary identity is `det_compressedGramOp`, `adjoint_comp_eq_conj_smul` and their use in `heightLe_liftVolume_ratio` (`New/Rows/LiftVolumeHeight.lean`).

The next two estimates for the compressed determinant must be proved without assuming that projection to a complex automorphic summand is integral or that it is an orthogonal projection at all coefficient conjugates. We give the determinant-level argument.

Lemma 7.18 (Finite norm-adapted basis)**.** *Let V be a complex vector space of dimension d, let $(\ell_a)_{a\in A}$ be a finite separating family of complex linear functionals on V, and let ϕ be a field embedding into a non-Archimedean valued field. Then some d of these functionals are the dual basis of vectors v_i with $|\phi\ell_a(v_i)| \le 1$ for all a, i.*

Proof. Take any basis of V. Among all d-tuples of the given functionals, choose one maximizing the absolute value under ϕ of their determinant in this basis. The maximum is positive by

separation. Use the chosen tuple as a dual basis. Express any other functional in this dual basis by Cramer's rule. Each coordinate is a quotient of a determinant obtained by replacing one of the selected rows by the new functional, by the maximizing determinant. Its *phi*-absolute value is at most one. This proves the assertion. The finite case of `exists_auerbach` in `Lib/General/LinearAlgebra/CommonEigenvectors.lean` is all that is needed here. □

At a fixed place ι, choose rational integral adelic frames for the small-block classes. Denote by f_a special table vectors given by the coordinate duals in those frames, scaled by the hook-number convention. Thus the functional on Z_ϑ is

$$\ell_a(z) = \langle f_a, g_\vartheta^* z \rangle.$$

The duals are integral in the Weyl gauge convention because the Fock Gram matrix times the relevant Weyl Gram matrix is the hook number times the identity (the Fock pairing of standard tableaux supplies the integral dual). In these frames the values $\ell_a(z)$ are exactly h_j times the conjugate-coordinate test of the table $g_\vartheta^* z$. Consequently, selecting a dual basis as in Lemma 7.18, each $h_j g_\vartheta^* z_i$ satisfies the requisite integral table test at the chosen place. The forms $g_\vartheta f_{a_i}$ are integral at all the embeddings used in the determinant test, and

$$v_i = h_j g_\vartheta g_\vartheta^* z_i$$

are integral at ι and tested: adjunction gives $g_\vartheta^* u_\vartheta = \bar{\gamma}_\vartheta A^*$, so the tested condition follows from the comparison theorem for $z_i \in Z_\vartheta$. In this step the assertion on the coordinate duals (including their gauge after conjugation) is the concrete integral-linear-algebra input of `norm_det_compressedGramOp_padic_le`; it relies on the integral frame lemma and `weylGram_mul_fockPairing` rather than on orthonormal coordinates having integral entries.

Apply (98) to these two families. The matrix of their pairings

$$(\langle g_\vartheta f_{a_i}, p_\vartheta (h_j g_\vartheta g_\vartheta^* z_h) \rangle)_{i,h}$$

is the matrix of the compressed operator in the chosen basis, by duality of ℓ_{a_i}. Its determinant is D_ϑ. Hence the determinant bound at twist $\vartheta = \theta_\iota^j$ is

$$|\iota D_\vartheta|_\ell \le e^{nr\beta_\ell}. \tag{109}$$

This derives the implication of the denominator test in `norm_det_compressedGramOp_padic_le`; it is precisely why determinant-level denominators suffice.

For the archimedean estimate, use the following property for all conjugate coefficient metrics: if y is tested and Z_ϑ lies in the good kernel of the Hecke polynomial at the appropriately chosen off-level place, then

$$\|p_\vartheta y\|_\sigma \le \|y\|_\sigma. \tag{110}$$

This is `AuxPair.normSqAt_starProjection_le` in `New/Selmer/FitTwists.lean`. Here is the mechanism. Write $y = z + x$ with $z \in Z_\vartheta$, x orthogonal to Z_ϑ. For a tested vector, x lies in the sum of the kernel of the good projector and the range of the Hecke polynomial (taking the complementary conjugate-place normalized operator as appropriate). This follows by taking orthogonal complements of the intersection defining the false-lift test, using the normality of the Hecke operators. Under application of a field automorphism σ, the adjoint of $H_w(X)$ for the coefficient-conjugate pairing is $H_{\bar{w}}(\sigma^{-1}(\overline{\sigma X}))$. This is `LevelPlaces.LevelForm.innerAt_heckePolyOnForms`, and the commuting Hecke operators and definiteness of the twisted metric imply that their relevant kernels are stable under taking the adjoint. In addition the good projector remains self-adjoint for every conjugate pairing because its underlying finite character action is unitary (character scalars are roots

of unity). Hence z, x remain orthogonal for the σ-pairing. Positivity gives (110). The adjoint and projector laws with their precise local hypotheses are *Formal-source inputs* in `New/Selmer/FitTwists.lean`.

Apply this contraction to the vectors $g_1 g_1^* z$ occurring in the compression. One cannot simply apply σ to the spectral bound at the ordinary embedding. Instead, take an eigenvector z with eigenvalue η, set $\zeta = g_1^* z \neq 0$, and pair the eigenvalue equation with the conjugate-compatible table obtained from ζ by the map $\kappa \circ \sigma^{-1} \circ \kappa \circ \sigma$ on coefficients. The mixed Cauchy–Schwarz inequality of (92), contraction (110), and the two table norm estimates in Proposition 7.17 then give

$$|\sigma\eta| \sum_{\alpha} \|\sigma\zeta_\alpha\|_{\mathrm{Weyl}}^2 \leq e^{2rB_0} \sum_{\alpha} \|\sigma\zeta_\alpha\|_{\mathrm{Weyl}}^2.$$

The sum is positive by injectivity of $g_1^*|_{Z_1}$. This is the eigenvalue calculation in `norm_det_compressedGramOp_arch_le` (`New/Rows/LiftVolumeHeight.lean`). Thus every conjugate of each eigenvalue has modulus at most e^{2rB_0}, and the determinant, the product of the n eigenvalues with multiplicity, satisfies

$$|\sigma D_1| \leq e^{2nrB_0}. \tag{111}$$

Using witness D_1 in the height definition, the local estimate (109), the global sum of β_ℓ, and (111), we obtain

$$\mathrm{Ht}_{\leq 3nrB_0}\left(\frac{D_{\theta_\iota^j}}{D_1}\right). \tag{112}$$

7.10 The height saving in a rank-one special value

It remains to show how derivative divisibility enters the lower bound, in the estimate for $\Gamma_j(1)/\Gamma_j(\theta_\iota^j)$. Fix a pair $p = (\nu, z)$ of rank-one type at j. The definite target of the lift of the trivial character of $U(1)$ has even rank $2z$ and its one-row coefficient weight at τ has degree $d_\tau = c_\tau - z$. Let $\Xi_{p,\vartheta}$ denote this lift, with the CM vacuum and jet scaling, namely `Data.Pair.trivialLift` in `New/Exchange/AlgebraicPartDefs.lean`. It is a scalar multiple of a sum of kernel forms over the classes of the norm-one torus, at the prescribed level. Put $d_{\mathrm{tot}} = \sum_\tau d_\tau$. The scaling of the theta kernel in this sum is

$$\kappa_0^{2z} (\overline{\pi_g}\,\mathrm{per})^{d_{\mathrm{tot}}},$$

where κ_0 is the common vacuum scale, π_g is the chosen level generator and per the jet scale. The integer

$$\mathrm{den}_p = g^{4fz} (gd)^{d_{\mathrm{tot}}}$$

clears its denominators outside its conductor. Let $M_* = \mathrm{chosenGroupOrder}$ be the fixed finite group order bounding the relevant class stabilizers. The normalized integral value used in the source is

$$\mathcal{A}_p(\vartheta) = M_* \,\mathrm{den}_p^2 \,\frac{\|\Xi_{p,\vartheta}\|^2}{N(Q)^{2z-1}}.$$

Nonzero proportionality of this value to $L^{M\cup Q}(z, (\nu\vartheta)^{-1})$ with a common constant independent of ϑ is the rank-one Rallis norm theorem `LValueData.inner_trivialLift_self_eq_L` (`New/Exchange/CharacterLiftNorm.lean`) and `intAlgPart_eq_mul_L` of `New/Exchange/TriangleLevelExponent.lean`. The factorial in the norm is exactly $\prod_\tau d_\tau!$, because a one-row standard monomial has Fock norm its factorial. The assertion that the residual normalization

is common, including the indicated conductor power, is a *Formal-source input*; the ensuing ratios satisfy

$$\rho_p(\vartheta,1)=\frac{\mathcal{A}_p(\vartheta)}{\mathcal{A}_p(1)}=N(Q)^{-(2z-1)}\frac{\|\Xi_{p,\vartheta}\|^2}{\|\Xi_{p,1}\|^2}. \tag{113}$$

In particular the required lift is nonzero. That normalized values are values of one group-ring sum with coefficients integral away from Q, when the twist varies at fixed modulus, is `isIntegralMeasure_intAlgPart` in the latter file. It follows by expanding the squared class sum of kernels into the two torus class indices and pulling the character factor out of their two translated kernels; the remaining coefficients are integral by the norm-one lift denominator estimate and clearing the stabilizers by M_*. These precise integral-measure and norm assertions are formal-source inputs, not appeals to a general interpolation theorem for L-functions.

To estimate the untwisted norm, compare with a flat pair p^0 of the same rank, i.e. of all degrees zero, and with a permissible weight-zero character after rank normalization. Define

$$Q_0=\mathcal{A}_{p^0}(1)/\|\Xi_{p,1}\|^2.$$

The ordinary rank-one norm itself need not be integral, but $Q_0\|\Xi_{p,1}\|^2=\mathcal{A}_{p^0}(1)$ is integral at all finite embeddings. The precise archimedean comparison is

Proposition 7.19 (Rank-one conjugate norms; formal-source input)**.** *For fixed rank $2z$, at every automorphism fixing K there is a common c_σ independent of the pair such that*

$$\log|\sigma(\|\Xi_{p,1}\|^2)|=c_\sigma+e_{p,\sigma}+\sum_\tau J(d_\tau),\qquad e_{p,\sigma}\ge -4f.$$

For a flat pair $e_{p,\sigma}\le 4f$ as well.

This is `Data.Pair.log_norm_map_norm_sq_trivialLift` in `New/Exchange/AlgebraicPartConjugates.lean`, obtained from the permutation and translation of the CM theta coefficients under field conjugation (`LValueData.map_trivialLift`). For automorphisms not fixing K, compose with complex conjugation to reduce the required modulus estimate to this case. Compare the upper bound for the flat numerator with the lower bound for the nonflat denominator. Because the weight of a pair of rank-one type has $d_\tau\ge w_0(\tau)$, the jet-size estimates compare $J(d_\tau)$ with $J(w_0(\tau))$. The extra logarithm of the fixed integer $M_*\,\mathrm{den}_{p^0}^2$ and the two losses of $4f$ have combined size at most B_0 (the exact bound is `D.le_of_hasLogBound` in `heightLe_norm_trivialLift_sq`). We deduce

$$\mathrm{Ht}_{\le B_0-\sum_\tau J(w_0(\tau))}(\|\Xi_{p,1}\|^2). \tag{114}$$

This is the declaration `heightLe_norm_trivialLift_sq` (`New/Rows/Separation.lean`). It is an estimate for the constant family with entries the specified squared norm, not for the absolute size of the norm at every place. The common unknown c_σ disappears on taking the ratio.

We need the finite-place lower bound on the pairing with a test form that restores the forced divisibility term. Put $F_\ell=\sum_\tau A_\ell(w_{\mathrm{nom}}(\tau),\tau)$, the quantity `Data.forcedPart`, so that the elementary prime-power formula gives

$$2\sum_\ell F_\ell\log\ell=\sum_\tau A(w_{\mathrm{nom}}(\tau),\tau).$$

All terms have finite support. The relevant depth indicator is `Data.depth_indicator_jet` in `New/Rows/GluedFormTheta.lean`. It is a *Formal-source input*: for the degree d_τ in the

allowed interval of that declaration, one common nonzero scale of archimedean conjugates bounded by its prescribed small budget serves all places; locally there is a jet of the prescribed total degrees (distributed among twelve coordinates) with valuation at most the forced depth loss, after multiplication by the scale. The proof in the source uses tight frames and the inverse-theta normal form to realize a large coefficient of the needed jet. An earlier mere upper bound for the norms of all jets cannot replace this attainability statement.

In obtaining an actual test form, the twelve coordinates sit next to flat blocks of ranks $2z-20, 4, 4$. They permit the coefficient of the jet to be detected by a genuine coefficient pairing and absorb the tame character choices through the flat forms. The test-form existence formula is `exists_test_form_inner_kernelForm` (`New/Rows/HeckeValueHeight.lean`). It constructs a form by three class sums weighted by the flat theta lifts in these three blocks. Its pairing with the kernel is an explicit product of their three normalized values, a multinomial coefficient

$$\prod_{\tau}(d_\tau!)/\prod_{\tau,i}\alpha_{\tau,i}!$$

and the remaining rank-twelve monomial jet kernel (with the matching character). The left side contains the class-number normalization and the vacuum-scale power. It asserts the resulting form is integral at every embedding over ℓ and is tested for the rank-one false lifts. The tested assertion relies on the same extra-plane descent vanishing as the comparison form theorem, with the rank-one label projector. This is a separate *Formal-source input*, since the existence of an integral kernel coefficient alone is not enough to produce a class-sum pairing against an integral tested vector.

For clarity there are two elementary finite Fourier steps in passing from these pointwise test-form formulas to the uniform height witness.

Lemma 7.20 (Fourier selection and multiplicative selection)**.** *Let G be a finite abelian group and $s : G \to \mathbb{C}$ any function. For each embedding ϕ in a non-Archimedean field there is a character χ such that*

$$\left|\phi\left(\frac{1}{|G|}\sum_{g\in G}\chi(g)s(g)\right)\right| \geq |\phi(s(1))|.$$

Furthermore, let $a(\chi), b(\chi)$ be two families indexed by a finite character group $G^\vee$, and let Q satisfy $Q^{|G^\vee|} = \prod_\chi a(\chi)b(\chi)$. For every prescribed $\psi \in G^\vee$ there are χ_1, χ_2 with $\chi_1\chi_2 = \psi$ such that

$$|\phi(a(\chi_1)b(\chi_2))| \geq |\phi Q|.$$

The second assertion also holds for Archimedean embeddings.

Proof. Fourier inversion expresses $s(1)$ as the sum of the normalized Fourier coefficients. The ultrametric inequality bounds its absolute value by their maximum. For the second assertion, take the product of the numbers $a(\psi\chi^{-1})b(\chi)$ as χ ranges over all characters. It is $Q^{|G^\vee|}$ by the permutation of the character group, so one absolute value is at least $|\phi Q|$. □

In `exists_twisted_classSum_ge`, the finite group is the group of torus classes after translating the chosen base class to the identity; its stabilizers are trivial and its characters are realized by the permitted tame weight-zero characters. The second step is `exists_factorisation_le` in the same file. These two selections allow the flat class sums to be paired with the depth indicator at every local embedding while a single product of flat values is used globally as the height witness.

Here is the precise quantitative interface obtained by applying these constructions. Fix the rank-one pair p, the index j and the small flat pair p_2 (rank four, tame part the chosen

primitive character ω). Let N_c be the torus class number at the principal chosen level. The combination of `depth_indicator_jet`, `exists_test_form_inner_kernelForm`, `exists_twisted_classSum_ge` and `exists_factorisation_le` in the proof of `heightLe_inv_ratio` (`New/Rows/HeckeValueHeight.lean`) yields a common $Q \neq 0$ such that $|\sigma Q| \leq e^{B_0}$ for every σ, and for each ℓ and each embedding ϕ there exists a tested, everywhere ℓ-integral vector u_ϕ in the rank-one target satisfying, at $\vartheta = \theta_\iota^j$,

$$|\phi Q|_\ell \, \ell^{-F_\ell} \leq \left| \phi \left(\frac{N_c^2 \langle u_\phi, \Xi_{p,\vartheta} \rangle}{\rho_{p_2}(\vartheta, 1)} \right) \right|_\ell . \tag{115}$$

The witness Q is an hth root of the product of the flat normalized values times the depth scale over the h tame characters, which gives the archimedean bound using the flat-conjugate estimates. For the fixed twist and a different finite embedding the test vector may change, but Q does not. Equation (115), including its conjugate-compatible version when one uses the field embedding composed with complex conjugation, is the local assertion called `K` in the proof of that declaration. We take this simultaneous assertion, which uses the attainable jet and all the class and gluing integrality facts just discussed, as a formal-source input. The following calculation reconstructs its numerical consequence.

For an arbitrary nonzero vector v in a complex Hilbert space, projection onto its line is

$$p_{\mathbb{C}v} w = v \langle v, w \rangle / \|v\|^2 .$$

The fit denominator condition for the rank-one lift line therefore bounds the pairings of two integral tested vectors with this projection by e^{β_ℓ}. In the source the strengthened form `DenominatorLe.rank_one_bound` in `New/Selmer/RankOneDenominator.lean` gives

$$|\iota(N_c \langle u, p_{\mathbb{C}\Xi} v \rangle)|_\ell \leq e^{\beta_\ell}$$

for such vectors. The extra class number N_c is retained by replacing the second vector by a sum of its twisted central translates, which are still integral and tested, whereas the line of Ξ is fixed under this central action; this is again a *Formal-source input*. Apply (115) at the two embeddings giving the two factors of the rank-one projection pairing. The lower bound for their product and the projection upper bound give

$$\left| \iota \left(Q \overline{Q} \, \frac{\rho_{p_2}(\vartheta, 1)^2}{N_c^3 \|\Xi_{p,\vartheta}\|^2} \right) \right|_\ell \leq e^{\beta_\ell + 2F_\ell \log \ell} .$$

The exponent 3 of N_c is obtained from the two factors of N_c^2 and the single factor of N_c in the strengthened projection denominator. Since $|\sigma(Q\bar{Q})| \leq e^{2B_0}$, summing the finite-place bound gives

$$\mathrm{Ht}_{\leq 3B_0 + 2\sum_\ell F_\ell \log \ell} \left(\frac{\rho_{p_2}(\vartheta, 1)^2}{N_c^3 \|\Xi_{p,\vartheta}\|^2} \right) .$$

Multiply by the estimate (114) and by the square of the inverse ratio of the ramified flat pair p_2 (height at most $2x^{999}$ from fit). Using (113), the product is $\rho_p(\vartheta, 1)^{-1} / (N_c^3 N(Q)^{2z-1})$. The integer factor N_c^3 costs nothing when multiplying back (its local absolute value is at most one), and the conductor power is a local unit. The saving is exactly minus the cost (90). Since $x^{999} \leq B_0$ by the admissible magnitudes, we conclude

$$\mathrm{Ht}_{\leq 6B_0 - \mathcal{U}}(\rho_p(\theta_\iota^j, 1)^{-1}) . \tag{116}$$

This is `heightLe_inv_ratio`.

7.11 Combining the special values and closing the lower bound

We now have precisely the estimate for the long string in (105). For p_0 consider lower points z with $7m/2 < z \le e_0$. There are $k/2$ of them, all of rank-one type at j; this follows from the specified weight of ν_0 and the numerical relations between r, k, t, m (the range verification occurs in `heightLe_inv_heckeValueProd_div`). Thus these $k/2$ inverse ratios have height at most $6B_0 - \mathcal{U}$ each, by (116). The remaining factors in the three products have $7m/2 + e_+ + e_- = 5m$ lower points in all; their underlying pairs are ramified at exactly M and their inverse ratios are bounded by the fit hypothesis with cost x^{999} each. All ratios enter squared, and the conductor powers in (97) are local units. Consequently

$$\mathrm{Ht}_{\le 10mx^{999}+k(6B_0-\mathcal{U})}\left(\frac{\Gamma_j(1)}{\Gamma_j(\theta_\iota^j)}\right). \tag{117}$$

This proves the deduction in `heightLe_inv_heckeValueProd_div` (`New/Rows/HeckeValueHeight.lean`). One should notice the count $k/2$ before squaring; counting k factors before the square would yield an incorrect leading coefficient.

Combine (117), the compressed determinant estimate (112), and identity (108). The total height of the one-lift inverse volume ratio is at most

$$3nrB_0 + n(10mx^{999} + k(6B_0 - \mathcal{U})) \le nk((\log b)B_0 - \mathcal{U}).$$

The scale inequality on the right is the one in `heightLe_liftVolume_ratio` (`New/Rows/LiftVolumeHeight.lean`): r/k is bounded, $mx^{999}/(kB_0)$ negligible and $\log b$ very large. Thus all the positive errors are absorbed by the $k(\log b)B_0$ allowance. Multiplying over $j = 1, \dots, m$ gives height bound $n(L - H)$ for the product of the ratios.

By Proposition 7.16 the full-chain ratio squared is the product of the squared single-lift ratios multiplied by the nth power of the inverse squared exchange factor. Its cost is m^2 squares, so its nth power has height at most $2nm^2tx^{999} \le 2nL$. We obtain

Proposition 7.21 (Reverse ratio)**.** *For every fit family,*

$$\mathrm{Ht}_{\le 2n(2L-H)}\left(\frac{Z(1)^2}{Z(\theta_\iota)^2}\right). \tag{118}$$

This is `Data.heightLe_volume_ratio_sq` (`New/Endgame/LongForms.lean`). Finally choose a fit family, apply Proposition 7.12 and square its ratio. The product of its ratio squared and that of (118) is identically one by injectivity. Lemma 7.4 gives

$$0 \le 2(V + 2d_{\mathrm{ind}} + nb) + 2n(2L - H).$$

As $b \le L$, rearrangement yields $n(H - 3L) \le V + 2d_{\mathrm{ind}}$, proving Theorem 7.1 from the explicitly identified inputs.

The arithmetic content of this argument is not an estimate of the size of individual L-values. It consists of finding comparison vectors for all good forms, extracting the rank-one ratios from their common periods, recovering adequate local integral tests from the geometrically constructed theta coefficients, and ensuring that projections of the tests have small determinant denominator. The same prime-power sum that appears in the divisibility cost in Theorem A is recovered, with precisely the opposite sign, in (116). The remaining issue, necessary for both occurrences of the height argument, is the existence of the simultaneous fit family. This calls for the congruence and local cohomological arguments in the next section.

8 Congruences, auxiliary primes, and the fit family

We now address the integrality issue in the determinant calculation. The notation for ranks and form spaces is as above. In this section a row of a Galois representation and a column of such a representation refer to the distinguished block and the entire representation, respectively, not to vectors in W and C. Let $B_0 = B/(mk)$, where $B = fa(\log b)^3$. We explain the reduction of the existence of a fit family to a global twisting theorem, a first-order congruence theorem, and the local crystalline estimates. The distinction between established linear algebra and the remaining specialized number-theoretic statements is important. Assertions identified as *formal-source inputs* below are precise imports from the accompanying Lean proof, not results independently proved in this manuscript. Their proofs require the auxiliary constructions indicated by the source references.

8.1 The integral projection problem

Consider an auxiliary unitary pair of odd source rank $r' \le r$ and odd rank difference $t' \le t$; its definite target has rank $r' + t'$. Write $\mathfrak{A}$ for its finite dimensional space of algebraic forms, with the class-sum pairing, and $Z \subseteq \mathfrak{A}$ for an image of theta lifting. Let P_Z be the orthogonal projection. The auxiliary pair has a finite level group, a chosen group-algebra idempotent e, and commuting split-place Hecke polynomials $H_w(X)$. As in Section 7, their test value is denoted X_w. The condition imposed on a form v is

$$v \perp \left(Z^{\perp} \cap \operatorname{im} e \cap \ker H_w(X_w)\right) \quad \text{for every split } w \text{ outside the level and its conjugate.} \tag{119}$$

It is essential that the subspace in (119) involves a kernel at each such place, rather than merely the space of theta lifts. These kernels can contain other forms with a Galois representation resembling that of a lift. We call a form satisfying (119) tested. The source definitions are `AuxPair.IsTested` and `DenominatorLe` in `New/Selmer/FitTwists.lean`.

Fix a prime ℓ and an abstract field isomorphism $\iota : \mathbb{C} \simeq \overline{\mathbb{Q}}_\ell$. The absolute value on the latter is normalized by $|\ell|_\ell = \ell^{-1}$; put $v_\ell(z) = -\log|z|_\ell / \log \ell$. For an integral coefficient ring $\mathcal{O}$ in a finite extension of $\mathbb{Q}_\ell$, let $\lambda_{\mathcal{O}}$ be its ordinary module length and $e_{\mathcal{O}} = \lambda_{\mathcal{O}}(\mathcal{O}/\ell)$. We use $\lambda = \lambda_{\mathcal{O}}/e_{\mathcal{O}}$ for normalized length. Thus $\lambda(\mathcal{O}/(a)) = v_\ell(a)$ for nonzero $a \in \mathcal{O}$. The conventions extend to the value $+\infty$ when necessary. Integral forms carry the lattice defined by their Weyl-coordinate values; integrality at all ℓ-adic embeddings is distinguished from integrality at one fixed embedding.

If $q = \dim_{\mathbb{C}} Z$, the denominator condition at ι with allowance β is

$$\left|\iota\left(\det\left(\langle u_a, P_Z v_b\rangle\right)_{a,b=1}^{q}\right)\right|_\ell \le \exp(qr'\beta) \tag{120}$$

for all $u_a \in \mathfrak{A}$ integral at every ℓ-adic embedding and $v_b \in \mathfrak{A}$ integral at ι and tested as in (119). The conditions concern the determinant of Hermitian pairings, not the norm of an arbitrary matrix of coefficients. In particular, r' is the rank of the source unitary group, distinct from q; the factor $r'\beta$ represents the loss allowed per form in the determinant. If $q = 0$ the empty determinant is one. For the good lifts one uses $e = e_0$ and Z the image of the auxiliary lift on good forms. For the rank-one comparison pairs one uses the tuple of labels $X - 1$, $r' = 1$, and Z the line generated by the particular trivial lift. The subspace in (119) is then the corresponding kernel of descent.

The height condition on the special values concerns ratios of twisted to untwisted Hecke values. Denote by $\rho_p(\vartheta, s)$ the ratio `Data.Pair.ratio` for a ramified pair p and a sign datum s. The letter p in a pair-label is not necessarily a rational prime. We require ℓ-adic bounds on the inverse ratios for all such pairs. Recall that `Data.Pair.IsRamified` specifies that the tame character of a ramified pair is nontrivial exactly at the chosen places above $g_0 g_1$.

Definition 8.1. For an admissible datum $\mathcal{D}$ a fit family consists, for every ι (including the choice of ℓ), of an m-fold twist θ_ι with conductor prime to ℓ. For every $1 \le j \le m$ there exists a family of real allowances $\beta_{j,\ell}$ such that

$$\sum_{\ell \in T} \beta_{j,\ell} \le B_0 \quad \text{for every finite set } T \text{ of rational primes,}$$

and (120) holds at the character θ_ι^j both for the good auxiliary lift at step j and for all rank-one pairs of index j. For every j, p, s as above, the family of inverse ratios $\rho_p(\theta_\iota^j, s)^{-1}$ has height at most x^{999} in the sense of the height predicate of the determinant argument.

In `Data.IsFitFamily`, the index j is a lift index everywhere. The rank-one pairs in `Data.Pair.IsRankOne` have integer parameter z with $m < z \le r$, prescribed tame part ω^j, and weights between $w_j + r - t$ and w_j. All conditions are imposed for the same character and its powers. The definition does not require a single twist working for all primes, nor does it allow independent choices for the arrows of one chain.

Theorem 8.2 (Fit family)**.** *Every admissible datum has a fit family.*

The intended proof is `exists_isFitFamily` in `New/Selmer/FitTwistsExist.lean`. We give its quantitative deduction from the other inputs in this section. An integrality estimate of this strength is necessary: a bound on a projection over $\mathbb{C}$ says nothing about its image of an integral form at ℓ, and a global estimate with an uncontrolled contribution from a new auxiliary conductor need not be usable in the final height sum.

8.2 Labels and Galois representations of lifted eigensystems

We summarize the normalization used for the attached Galois representations. Put $G_E = \mathrm{Gal}(\overline{E}/E)$, let c denote complex conjugation in the ambient absolute Galois group, write $g^c = cgc^{-1}$, and put $\epsilon_\ell = \chi_{\mathrm{cyc},\ell}$. Thus ϵ_ℓ^{-e} is the character of $\mathcal{O}(-e)$. A character of $U(1)$ is transferred to an anticyclotomic character of G_E by base change and class field theory; ψ denotes the avatar of its inverse. For the auxiliary pair the source rescales the representation attached to a form by a particular algebraic character, `chiOneInv`. If A is a place over a split unramified $w \nmid \ell$ and g a geometric Frobenius normalized as in the source, the scalar of this character is $\iota(a_O(w))^{-1}$, where $a_O(w)$ is the Frobenius character value of the pair (distinct from the Hecke test point X_w). This normalization is designed so that the string in the theta-lift Hecke relation becomes the fixed Tate string $1, \epsilon_\ell^{-1}, \dots, \epsilon_\ell^{1-t'}$.

The key input is not merely an assignment of semisimple Galois representations to eigensystems. It is the simultaneous assignment to all the lifted systems of a row representation T of rank r', independent of the twist, with prescribed inertia and jumps, and an identity of traces for the whole normalized representation U:

$$\operatorname{tr} U(g) = \psi(g) \operatorname{tr} T(g) + \sum_{e=0}^{t'-1} \epsilon_\ell(g)^{-e} \qquad (g \in G_E). \tag{121}$$

For the first block the field of coefficients and lattice can depend on ℓ and ι; the row T is the one associated to the untwisted system. It must be possible to work with integral lattices that satisfy the balance and matrix-span hypotheses below.

The labels at a chosen place v are monic squarefree polynomials c_v of degree r' over its residue field. The roots label residue-field characters, and twisting by the inverse tame character scales the roots by the corresponding residue-field element. The label thus modified is denoted $\widetilde{c}_v$. The aptness conditions include $1 \notin \mathrm{Roots}(\widetilde{c}_v)$; for $r' = r$, at least one primitive

label occurs above each of g_0, g_1, while for $r' = 1$ the unmodified label used is $X - 1$. The ranks r' and t' are odd. The other conditions specify the chosen places, the extra tame level, and the tame part $\omega^j\vartheta$ of the auxiliary character. This is `IsApt` in `New/Selmer/AptTupleDefs.lean`; aptness is checked for the good and rank-one pairs by `isApt_auxPair` and `isApt_rank_one` in `New/Selmer/HeckeStableSubspaces.lean`.

Here are the row conditions that enter the congruence argument. Let M_g denote the integer `Data.chosenGroupOrder`, the order of the finite product of general linear groups used to bound automorphism and level denominators; in particular

$$g_i^a - 1 \mid M_g \quad (i = 0, 1,\ 1 \le a \le rt). \tag{122}$$

Define

$$c_\ell = v_\ell(M_g) + 2^{-\ell}.$$

A row is continuous, trace-nondegenerate, and conjugate self-dual with the normalizing Tate shift: in the rank-difference convention of the source the trace identity has the factor $\epsilon_\ell(g)^{1-t'}$. It is unramified off ℓg and crystalline at primes above ℓ prime to g, potentially crystalline at the other stipulated places, with the indicated multiset of jumps. The row type requires, for each complex embedding, jumps either in a low interval bounded by h_2^2 in absolute value (and bounded below by $h_2/2$), or within t of the appropriately signed w_j. The number of jumps with absolute value at most $2h_2^2$ is bounded by $r'm/k$, and the number of negative jumps at conjugate embeddings together is at most r'. Further:

1. At each place above g not above ℓ, some tame inertia generator γ (the specified one at the chosen places) has determinant bound

   $$v_\ell(\det(T(\gamma) - 1)),\ v_\ell(\det(T((\gamma^c)^{-1}) - 1)) \le r'c_\ell.$$

   The inertia action is killed by an integer prime to the residue characteristic at that place.

2. In all sufficiently large coefficient rings, each lattice representation $T_{\mathcal{O}}$ spans the full algebra $M_{r'}(\mathcal{O})$ as an $\mathcal{O}$-module by its group elements. There is a Galois element having scalar action and cyclotomic character $g_i^{r'}$ for one of the two chosen primes.

3. There is also a scalar element in inertia above g off ℓ, with scalar s such that all powers $(s^{\ell^a})^j$ for $1 \le j \le 4r' + 2$ differ from one by a unit, for all sufficiently large a as in the source. In particular one can arrange $s - 1$ and $s + 1$ to be units.

See `IsColumn`, `IsRow` and `HasRow` in `New/Selmer/AptTupleDefs.lean`, and the row-type definition in `New/Selmer/DeterminantBound.lean`. The precise phrase "for all sufficiently large coefficient rings" is the eventual assertion over the directed collection of ℓ-adic coefficient rings; it permits further finite extension after choosing the twist or an auxiliary level. We do not identify the algebraic closure with one finite coefficient ring.

Proposition 8.3 (Formal-source input: rows for all twists)**.** *For every nonzero Hecke eigensystem occurring in the good forms at the final single-pair source W_m, and every auxiliary lift index j, there is a row T as above satisfying (121) for each twist of the auxiliary pair and every attached normalized representation of a lifted eigensystem. The corresponding statement holds for every rank-one pair, with the untwisted rank-one system. All relevant row representations can consequently be put in one finite list for a fixed $\ell, \iota, \mathcal{D}$.*

This is `exists_hasRow_auxPair` and `exists_row_for_all_twists` in `New/Selmer/RowsExist.lean`, applied using the apt labels and rank-one statements. The passage from an automorphic eigensystem to a labelled column uses the local-global compatibility packages in `Lib/GaloisRep` and `New/Selmer/LabelledColumns.lean`. The latter obtains the conjugate self-dual inertia polynomials and proves `IsLabelledColumn.isRow`. At a chosen place, for inertial s, there are $r' + t'$ roots y annihilated by $y^{g_i^d - 1} = 1$ for some $1 \le d \le r$, and the product of $X - y$ kills s and the conjugate inverse on every labelled column of the specified type. For an untwisted row no root is 1 at the distinguished generator. This is `exists_inertia_eigenvalues`. The proof uses the primitive label above a prime different from ℓ to enforce nondegeneracy and the matrix-span condition. The row-type verification uses the explicit highest weights; the bound on the number of low jumps is what controls the crystalline defect in the local part of the argument.

The finiteness assertion deserves emphasis. The space of good forms is finite dimensional, and the appropriate Hecke operators are simultaneously diagonalizable. For a finite-dimensional space the commuting normal Hecke operators admit a common eigenbasis, and therefore only finitely many characters occur. The equality of relevant multiplicities across signatures identifies the list used in constructing the auxiliary pairs. There are only finitely many pairs with the prescribed weights and tame levels: they are characters in finite sets at fixed conductor and prescribed infinity type, and the construction ranges over finitely many integer weights. Thus the row list may be formed before choosing the finite-order twist. The same is not true of a putative bound obtained directly from the infinitely variable twists of all eigensystems.

We record two algebraic details used in this reduction. First, if the target of a one-dimensional polynomial label is a value z of finite order and neither of z^a, z^b becomes equal after raising to any ℓ-power, then $|z^a - z^b|_\ell = 1$. Indeed write $z = z_{\ell'} z_\ell$ as commuting roots of unity of prime-to-ℓ and ℓ-power order. Reduction modulo the maximal ideal kills z_ℓ and is injective on roots of unity of prime-to-ℓ order (by the separability of $X^n - 1$ and the uniqueness of a lift of the root 1 in the principal unit group). The premise says $z_{\ell'}^a \neq z_{\ell'}^b$, so their difference is nonzero on reduction. Second, if traces of two representations agree on every group element in characteristic zero and each has nondegenerate trace pairing on its image algebra, their representations are conjugate after matching multiplicities; this is the semisimple trace argument used to turn (121) into block diagonal lattices for the distinguished systems. To see the role of trace nondegeneracy, consider finite sums in the group algebra. If such a sum acts as zero in one representation, pairing its product with every group element gives trace zero in the other; nondegeneracy implies that it acts as zero in the other as well. Thus the image algebras identify with their trace functionals. Over an algebraically closed field the multiplicity of each simple module is determined by this functional, giving equivalence. For the source proof the relevant finiteness and trace-nondegeneracy properties of the attached representations are part of the input, not automatic from comparing unramified Euler factors alone.

8.3 Selmer modules over finite coefficient rings

For a stable lattice $T_{\mathcal{O}}$ with row action $R : G_E \to \mathrm{GL}_{r'}(\mathcal{O})$ and a finite-order character ψ with values in $\mathcal{O}^\times$, let V_M be the twist of $(\mathcal{O}/\ell^M)^{r'}$ by $\psi \otimes R$. Its continuous cocycles obey

$$z(gh) = z(g) + (\psi R)(g) z(h).$$

For a finite set of places S stable under conjugation, and a prescribed family $\mathcal{L}$ of finitely supported linear relations on functions on G_E, write

$$\mathfrak{S}_M(\psi R; S, \mathcal{L}) = \left\{ z \in Z^1(G_E, V_M) : \ z(I_v) = 0 \ (v \notin S), \ z \text{ obeys } \mathcal{L} \right\}.$$

At primes over ℓ the family $\mathcal{L}$ is the collection of linear relations obeyed by the characteristic-zero crystalline cocycles Z^1_f of the chosen integral lattice, pulled from the decomposition groups to G_E; it is not the assertion that each torsion cocycle itself is the reduction of a crystalline cocycle. This weaker formulation is important in the local comparison; it is `crysRels` in `New/Selmer/DeterminantBound.lean`. The ambient space is a module of actual cocycles, not cohomology classes. The source `SelmerLengthLe` normalizes by one evaluation $z(\gamma) = 0$, where $\psi R(\gamma) = sI$ and $s - 1$ is a unit, $\epsilon_\ell(\gamma) = 1$.

The normalization has a simple conceptual meaning. If any module of local conditions is stable under adding coboundaries, the evaluation at γ maps its coboundary subgroup isomorphically onto V_M: the coboundary defined by $v \in V_M$ is $\delta v(g) = ((\psi R)(g) - 1)v$, and $\delta v(\gamma) = (s-1)v$. Thus for every cocycle z there is a unique $v = (s-1)^{-1}z(\gamma)$ such that $z - \delta v$ vanishes at γ. The normalized cocycles then represent the corresponding Selmer classes uniquely. This argument explains why an actual cocycle-module bound can stand in for a more familiar cohomological bound. The source includes the scalar and the normalization in its definitions because the use of both at the integral level is essential later.

Several times we shall use the following elementary length facts. Over a DVR, if $A \in M_N(\mathcal{O})$ has nonzero determinant, the kernel of A on $(\mathcal{O}/(a))^N$ has length at most $\lambda_{\mathcal{O}}(\mathcal{O}/(\det A))$. This follows by writing A in Smith normal form with diagonal entries π^{e_i} up to units; the kernel on each factor has length $\min(e_i, v_\pi(a))$. If $Y, Z \subseteq V$ are submodules of a finite-length module, the length lost by intersection with Z equals $\lambda(Y/(Y \cap Z))$, and such relative lengths are subadditive under successive intersections. Thus if each new linear condition takes values in a torsion module of length at most c_j, imposing all of them decreases length by at most $\sum c_j$.

Definition 8.4 (Balanced lattices)**.** A representation R over a coefficient ring $\mathcal{O}$ is balanced up to a constant β_0 at depth M for a ramification set S, relations $\mathcal{L}$, and normalizer γ if it has the following properties. Its reduction modulo ℓ^M is locally constant; it is unramified off S; its group matrices span $M_{r'}(\mathcal{O})$; the residue field has more than $4r' + 2$ elements. The distinguished inertial element γ lies above S off ℓ, with $R(\gamma) = R((\gamma^c)^{-1}) = sI$ and $s \pm 1$ units. The coboundaries obey $\mathcal{L}$ modulo every divisor of ℓ^M. Finally there are $g \in G_E$ and scalars α, α' with $R(g) = \alpha I$, $R((g^c)^{-1}) = \alpha' I$, such that for each $a \mid \ell^M$ the local cocycle conditions satisfy

$$\sum_{v \in S} \inf_{A|v} \lambda_{\mathcal{O}}(Z^1_v(\mathcal{L}, \mathcal{O}/(a))) + r' v_\pi(\alpha - \epsilon_\ell(g^c)\alpha') \\ \leq (|S| + f)\lambda_{\mathcal{O}}((\mathcal{O}/(a))^{r'}) + \beta_0. \tag{123}$$

Here v_π is the valuation equal to one on a uniformizer of $\mathcal{O}$ and Z^1_v denotes the cocycles on the local decomposition group obeying the induced relations. The local fields and closures are chosen as in the global representation. This is the predicate `IsBalanced` of `New/Selmer/Corepresentable.lean`.

Proposition 8.5 (Formal-source input: balanced row lattices)**.** *Fix a row T of rank r' and a chosen place w above g not above ℓ, with tame generator z. For every sufficiently large coefficient ring and depth M there is a lattice representation R of T, a normalizer γ and β_0 making it balanced for ramification at the specified level and the local crystalline relations. They satisfy*

$$\frac{\beta_0}{e_{\mathcal{O}}} + v_\ell \det(R(z) - 1) + v_\ell \det(R((z^c)^{-1}) - 1) \leq 10 f r' c_\ell. \tag{124}$$

This is `IsRow.eventually_isBalanced` in `New/Selmer/RowBalanced.lean`; the finite ramification set may be taken as the places over ℓg embedded in the common exceptional set

of primes. The proof relies on a local-global comparison of dimensions of crystalline cocycles and the scalar difference. The arithmetic point is the factor $10fr'c_\ell$ rather than a bound in terms of the depth M: the depth will be much larger than this and is not an available contribution to the determinant allowance. We retain (123) to identify exactly where local conditions enter the subsequent use of global duality.

8.4 New level characters and tied cocycle conditions

Let S_0 be a fixed finite exceptional set containing the original ramification places, the places over ℓ, and the places of the prescribed levels that must not be used as auxiliary places. A finite set Q is auxiliary if each place $P \in Q$ has prime norm, both P and P^c lie off S_0, and $P^c \notin Q$. Thus the rational place below P is split in the relevant quadratic extension, and the auxiliary conditions at the two conjugate sides can be specified compatibly. The formal definition is `IsAuxiliary`. We reserve Q for the modifying conductor of an actual twist, and J initially for a set of auxiliary places carrying further linear conditions. The subsequent Taylor–Wiles places used in the projection argument are different sets.

For a depth N write $q_N = \ell^N$. Auxiliary prime data consist of a place P with $\mathrm{N}P \equiv 1 \pmod{q_N}$, a chosen geometric place above it, a tame inertia generator τ_P, an arithmetic Frobenius φ_P, and an anticyclotomic homomorphism

$$h_P : G_E \longrightarrow \mathbb{Z}/q_N\mathbb{Z}, \qquad h_P(\tau_P) = 1,$$

trivial on the prohibited local subgroups and of conductor supported on P, P^c. The condition of coming from $U(1)$ includes the local triviality and anticyclotomic condition needed for the subsequent passage to tame characters. Frobenius and inertia nomenclature follows the conventions of `AuxPrimeData` and `IsAnticyclotomicAt` in `New/Selmer/Corepresentable.lean`.

Proposition 8.6 (Formal-source input: one anticyclotomic prime)**.** *Let z lie in the normal closure of the inertia groups above a finite set R outside the original prescribed set (and its conjugate) and the places over ℓ. Given a neighborhood U of z, and any N, there is a new prime-norm place P outside $R \cup R^c$ with auxiliary prime data at depth N, with $\varphi_P \in U$, $h_P(\varphi_P) = 0$, and h_P trivial at the specified old local places and above ℓ.*

This is the precise auxiliary-prime mechanism of `exists_isAnticyclotomicAt` in `New/Selmer/Corepresentable.lean`. Its proof uses the finite generation of the group of S-units. A Kummer extension is formed by adjoining q_N-th roots of a finite generating system and the q_N-th roots of unity. Inertia at allowed places fixes these roots as required. Chebotarev then produces a prime-norm place with Frobenius close to z and whose residue field satisfies the splitting conditions. The local residue-field unit logarithm, sending a suitable tame unit to 1 modulo q_N, is trivial on the image of the prescribed global units, so it descends to the idelic quotient after extending along the chosen ideal classes. Taking the quotient by complex conjugation realizes an anticyclotomic class-field character. One finally modifies the representative of Frobenius by an inertia element to obtain $h_P(\varphi_P) = 0$ (in additive notation). The compatibility of the global unit condition and the Frobenius prescription is the point of the Kummer construction; it cannot be obtained by selecting an arbitrary prime congruent to 1 modulo ℓ^N.

A pencil of auxiliary places consists of an auxiliary finite set J, a helper $P_0 \in J$, the data just described for every $P \in J$, and elements $t_{P,i} \in \mathcal{O}$ for the different row representations R_i. Besides the auxiliary-place conditions, the following hold:

$$h_{P'}(\varphi_P) = \begin{cases} 1, & P' = P_0, \\ 0, & P' \neq P_0, \end{cases} \quad (P \neq P_0), \tag{125}$$

φ_{P_0} lies in a prescribed neighborhood of a chosen element z, and for every i with $t_{P,i} \neq 0$ both φ_P and $(\varphi_P^c)^{-1}$ act trivially on R_i modulo ℓ^M. The tied conditions for row i are

$$\begin{aligned} c(\tau_P) &= t_{P,i}c(\varphi_P), \\ c((\tau_P^c)^{-1}) &= t_{P,i}c((\varphi_P^c)^{-1}). \end{aligned} \tag{126}$$

They are linear relations on cocycles unramified outside $S \cup J \cup J^c$. Notice the special role of P_0: although each row uses different slopes, the chosen auxiliary places and their characters will be common to all rows. This permits one actual twisting character at the end. The source definitions are `IsPencil` and `tieRelations`.

A basic property of these pencils explains how imposing and then removing ties can be controlled. For a tuple $y = (y_P)$ of q_N-th roots of unity, viewed as units modulo some torsion modulus, define

$$\psi_y(g) = \prod_{P \in J} y_P^{h_P(g)}. \tag{127}$$

It follows from the unramified condition at the other primes that $h_{P'}(\tau_P) = \delta_{PP'}$, and anticyclotomicity gives the same values on $(\tau_P^c)^{-1}$. Thus ψ_y has value y_P at the two tame generators at P, P^c, and by (125) has the value y_{P_0} at both of the appropriate Frobenius elements when $P \neq P_0$. This is `IsPencil.pointChar_idOrConjInv` in `New/Selmer/TiltingLemma.lean`. The point of (125) is that one can choose different points on a line $y_P = 1 + t_{P,i}(u-1)$, with helper coordinate u, for each row, and still use polynomial interpolation to find one common character.

Lemma 8.7 (Removal of a tie; elementary calculation)**.** *Let V be a finite-length module of cocycles over a DVR, with values in $(\mathcal{O}/(a))^{r'}$. Suppose that at one pair of generators (τ, φ) the induced affine actions of τ and φ commute. Suppose also that for the twisted action R', one has $R'(\varphi) \equiv uS$ modulo a and*

$$R'(\tau) - 1 = t(R'(\varphi) - 1)$$

on the coefficient module. Then the length lost by imposing the relation $c(\tau) = tc(\varphi)$ is at most $v_\pi(\det(uS - 1))$ (with the infinite case interpreted as vacuous).

Proof. Write the affine action as $v \mapsto R'(g)v + c(g)$. Evaluating commutativity at zero gives

$$(R'(\varphi) - 1)c(\tau) = (R'(\tau) - 1)c(\varphi).$$

Hence the defect $c(\tau) - tc(\varphi)$ belongs to the kernel of $uS - 1$. The quotient of V by the tied submodule identifies with the image of the defect map, and the kernel estimate from Smith normal form bounds its length by the indicated valuation. □

In the application, tame local relations say that Frobenius conjugates a tame inertia element by the power $\mathrm{N}P$. With $\mathrm{N}P \equiv 1 \pmod{\ell^N}$ and N larger than the depth of the coefficients and of the twist, the corresponding affine actions commute modulo ℓ^M. The relation between $R'(\tau)$ and $R'(\varphi)$ on the line follows from (127): if the slope is nonzero the original row matrix at Frobenius is the identity modulo ℓ^M, whereas if the slope is zero the twisted inertia action is trivial. At the helper, removing ties uses $u = 1$ instead. Summing Lemma 8.7 over both conjugate generators and the subset of places whose ties are removed gives the exact determinant-loss term in `IsPencil.length_le_of_drop_ties`, in `New/Selmer/TiltingLemma.lean`. The full source statement permits specified replacements for the Frobenius elements as long as their matrices agree modulo ℓ^M. The ramification and affine-action assertions just invoked are hypotheses verified there; in particular one does not divide by ℓ or by residue cardinalities inside the torsion module.

8.5 Shrinking several tied Selmer modules

There are two logically separate ingredients in the existence of a pencil that shrinks every module. One is finite-dimensional linear algebra over a DVR; the other is a reciprocity lemma ensuring a new place on which a long module produces a nonzero detection. We record both, because the arithmetic conclusion does not follow from mere genericity of slopes.

Lemma 8.8 (Tilting a pair of linear maps)**.** *Let $\mathcal{O}$ be a DVR with residue field κ, let $A = \mathcal{O}/I$ have finite length, and let $\Phi_0, \Phi_1 : V \to A^u$ be $\mathcal{O}$-linear. For a finite-length submodule $Y \subset V$, all except at most u residue classes $s \pmod{\mathfrak{m}_{\mathcal{O}}}$ satisfy*

$$\lambda_{\mathcal{O}}(Y \cap \ker(s\Phi_0 - \Phi_1)) \le \lambda_{\mathcal{O}}(Y \cap \ker \Phi_0 \cap \Phi_1^{-1}(\Phi_0 Y)). \tag{128}$$

Proof. Let Y_1 be the sum of the kernels on the left as s varies. Then $Y_1 \subseteq Y$, and $\Phi_1 Y_1 \subseteq \Phi_0 Y_1$ because this holds on each summand. Let $L \subseteq \mathcal{O}^u$ be the inverse image of $\Phi_0 Y_1$ under reduction modulo I. Since L is free of rank at most u, lift the quotient map $\pi : L \to \Phi_0 Y_1$ to a map $\sigma : L \to Y_1$ with $\Phi_0 \sigma = \pi$. Lift $\Phi_1 \sigma$ to an endomorphism T of L using the same projectivity. For every s whose reduction is not a root of the characteristic polynomial of the reduction of T, the endomorphism $sI - T$ of L is invertible. Therefore

$$(s\Phi_0 - \Phi_1)\sigma(L) = \pi((sI - T)L) = \Phi_0 Y_1.$$

The image of $s\Phi_0 - \Phi_1$ on Y_1 is contained in $\Phi_0 Y_1$, hence equal to it. Its kernel on Y_1 has length $\lambda_{\mathcal{O}}(Y_1) - \lambda_{\mathcal{O}}(\Phi_0 Y_1) = \lambda_{\mathcal{O}}(Y_1 \cap \ker \Phi_0)$. The kernel on Y is included in Y_1 by definition. Finally $Y_1 \cap \ker \Phi_0 \subseteq Y \cap \ker \Phi_0 \cap \Phi_1^{-1}(\Phi_0 Y)$, giving (128). □

This is the lemma `tilting`, in `New/Selmer/TiltingLemma.lean`. In the arithmetic application $u = 2r'$, because both conjugate sides are included in each evaluation. Avoiding the exceptional slopes at two different moduli, and also the slope zero, requires at most $4r' + 1$ exceptional residue classes. This explains the residue-field requirement in `IsBalanced`.

The reciprocity statement used in conjunction with this lemma is the following special form of a global-duality obstruction.

Proposition 8.9 (Formal-source input: the mirror criterion)**.** *Let R and the pencil be as above, with R balanced up to β_0, and $a \mid \ell^M$, $a \neq 0$. Let $X(J, a)$ be the corresponding normalized module of cocycles obeying the ties modulo a. Suppose some element of $X(J, a)$ has a unit coordinate. If for every admissible enlargement by a new prime P, compatible with the pencil for every choice of its new slope, the image of $X(J, a)$ under evaluation at the two Frobenius elements lies in the image under tame-inertia evaluation of the relaxed module (in which the new ramification is allowed but its ties not imposed), then*

$$\lambda_{\mathcal{O}}(X(J, a)) \le \beta_0.$$

This is `IsPencil.length_le_of_unseen_class` in `New/Selmer/TiltingLemma.lean`. A unit-valued coordinate is used to detect a nontrivial row in the full matrix algebra. The detailed proof constructs dual cocycles via global reciprocity, with isotropic pairings against the class in question. The dual classes, evaluated on a suitable open subgroup on which the row and cyclotomic character act trivially modulo the torsion modulus, are bounded by the size of the torsion subgroup for the scalar difference $\alpha - \epsilon_\ell(g^c)\alpha'$ from (123). The "mirror" estimates use the full span of the matrix algebra and the units $s \pm 1$ to constrain that image. The algebraic assertion is `card_le_torsion_of_isotropic` in `New/Selmer/LinePointPowers.lean`; the reciprocity construction and its local restrictions are developed in the proof of the mirror criterion. The determinant or torsion size term is absorbed by the balance bound, not by a global bound on all cohomology classes.

Proposition 8.10 (Boundedly many ties; formal-source input)**.** *Fix finitely many balanced representations R_i at depth M, a neighborhood of an allowed inertial element z, and a divisor a of ℓ^M. There is an integer J_0 such that for every depth N of the auxiliary characters there is a pencil of at most J_0 places with normalized cocycle modules satisfying*

$$\lambda_{\mathcal{O}}(X_i(J,a)) \le \beta_{0,i} \qquad \text{for every } i.$$

This is `exists_isPencil` in `New/Selmer/TwistTheorem.lean`. The important quantifier is that J_0 is chosen before N. To explain its proof, one inducts on the length of $\mathcal{O}/(a)$. Lifting a good pencil for a smaller divisor a to πa gives a finite a priori bound on the length of the larger-depth modules, by evaluating on the finitely many tame generators. For the larger divisor minimize the total length of all the tied modules subject to preserving the smaller-depth inequalities and an upper bound on the sum of the number of places and this total length. If a row has normalized module longer than the allowed bound, and there is no unit-valued element, it reduces to the smaller depth: multiplication by π identifies the modules killed by the shorter divisor, preserving all the linear relations. If there is a unit-valued element the mirror criterion applies. Unless its conclusion already gives the desired bound, it supplies a new place where the image under Frobenius evaluation is not contained in that of the relaxed module under tame-inertia evaluation.

At such a place, write Y for the relaxed module, Φ_0 for tame-inertia evaluation, and Φ_1 for Frobenius evaluation. Imposing $\Phi_0 = 0$ recovers the original unramified module. The failure of the mirror inclusion gives a strict decrease when imposing the additional condition $\Phi_1(c) \in \Phi_0 Y$. Lemma 8.8 yields a slope giving at least this strict decrease at the larger depth, and by avoiding the exceptional classes at the smaller depth as well it preserves the smaller-depth inequality. The new slopes for all other rows can be set to zero, so their modules do not increase. Consequently the minimized total length strictly decreases, while the sum of length and number of primes does not increase. This is the contradiction with minimality. The induction bounds the required number of new places independently of the auxiliary-character depth. We have described the induction's length mechanism in order to identify where the key reciprocity input is used; the actual selection of primes satisfying all neighborhood constraints together is part of the cited formal-source assertion.

8.6 Finite presentations and interpolation of lengths

To convert the linear pencils into common characters, one must avoid choosing a different character for every row. The device is to bound the length of each Selmer module by a polynomial in the tame-character parameters. The polynomial need not have a degree bounded in terms of the original rank. A special feature of the root-of-unity argument below is that it handles arbitrary polynomial degree uniformly, after a common finite bound on the number of variables is fixed.

Here is the underlying finite-presentation lemma. Fix a modulus $m_0 \neq 0$, a representation given by matrices over a coefficient ring Λ, a finite ramification set, a subfamily of linear relations L, and an open neighborhood of the identity on which all the relevant actions on the m_0-torsion coefficient modules are trivial. Continuous cocycles of these types can be described by finitely many coordinates (the values at chosen group elements) obeying a submodule of relations in a finite free Λ-module. One standard justification is to choose an open normal subgroup on which the action is trivial, restrict the cocycle there, and observe that its restriction is a homomorphism to a module killed by m_0. The maximal abelian exponent-m_0 quotient of the subgroup unramified off the prescribed finite set is finite, by global class field theory (or the corresponding finite-character finiteness theorem). Thus there is one common finite quotient through which all such cocycles factor, independently of their coefficient module.

More explicitly, the restriction on the open subgroup is a homomorphism with the prescribed ramification; it factors through the common quotient because any nonzero value in a torsion module is detected by a homomorphism to $\mathbb{Q}/\mathbb{Z}$ (and hence to the subgroup killed by m_0). Taking normal cores yields the desired finite quotient of the original group. On its finite set of elements, the cocycle relations, inertia-zero relations, and the imposed linear relations are simply linear equations. This is the principle of `exists_corepresenting_coords` in `New/Selmer/Corepresentable.lean`. Its source proof also handles the universal coefficient ring by constructing the relation module directly.

Let the character parameters be $Y = (Y_1, \dots, Y_q)$, with $h : G_E \to (\mathbb{Z}/\ell^N)^q$ a finite family of homomorphisms. Use the polynomial ring $\Lambda[Y]$ for the pointwise operators with coefficients $\prod_j Y_j^{h_j(g)}$ (least nonnegative lifts of the exponents). At an actual tuple of roots of unity, or modulo a depth at which the tuple is a root of unity, these operators represent the twist. The cocycle condition and all the other local conditions give a submodule of relations $R(Y) \subseteq \mathcal{O}[Y]^u$ which contains $\ell^M \mathcal{O}[Y]^u$. This also makes sense at arbitrary tuples: even if the formal operators fail to be a representation there, the algebraic equations define a submodule, and we only need their specialization at true or approximate characters.

Lemma 8.11 (Sharp Fitting polynomial)**.** *Let $R \subseteq \mathcal{O}[Y]^u$ contain $\ell^M \mathcal{O}[Y]^u$. For $Y = y$, write R_y for the specialization and define*

$$X_y = \{\xi \in (\mathcal{O}/\ell^M)^u : \langle r, \xi \rangle = 0 \text{ for every } r \in R_y\}.$$

Fix y_0 over a DVR $\mathcal{O}$. There is $F \in \mathcal{O}[Y]$ with

$$v_\pi(F(y_0)) = \lambda_{\mathcal{O}}(X_{y_0}), \qquad \lambda_{\mathcal{O}'}(X_y) \le v_{\pi'}(F(y))$$

for every specialization over a DVR $\mathcal{O}'$ above $\mathcal{O}$. Here F is mapped to the appropriate coefficient ring in the second formula.

Proof. Over $\mathcal{O}$, choose a basis of the full-rank sublattice $R_{y_0} \subseteq \mathcal{O}^u$. Each basis element is a finite $\mathcal{O}$-linear combination of evaluations of elements of R, so we can lift the basis to u polynomial rows of R. Take F to be their determinant. At y_0, Smith normal form gives elementary divisors π^{s_i}, all dividing ℓ^M because R_{y_0} contains $\ell^M \mathcal{O}^u$. The annihilator modulo ℓ^M has length $\sum_i s_i = v_\pi(F(y_0))$. At any other specialization X_y lies in the kernel modulo ℓ^M of the same matrix of specialized rows. If its determinant is nonzero, the kernel bound by Smith normal form gives the inequality; if zero, the inequality is vacuous. This proof also works if R was presented by an infinite set of relations, since only the finite chosen rows enter the determinant. □

This is the algebraic mechanism of `exists_sharp_fitting_poly` in `New/Selmer/LinePointPowers.lean`, where R is obtained by the preceding corepresentability lemma. The equality rather than a mere upper bound at y_0 is essential: the tied Selmer theorem controls the actual module length, not an arbitrary chosen determinant. The lemma applies also after imposing a selected subset of ties; its sharp specialization can be different for every row.

Given a root of unity $u^{\ell^k} = 1$ and $t \in \mathcal{O}$, the point $y = 1 + t(u-1)$ is an approximate root of unity:

$$\ell \mid (y-1)^{\ell^k}, \qquad \ell^M \mid y^{\ell^N} - 1 \quad (N \ge k + M). \tag{129}$$

Indeed in characteristic ℓ, $(u-1)^{\ell^k} = u^{\ell^k} - 1 = 0$, so $y^{\ell^k} \equiv 1 \pmod{\ell}$. If $v \equiv 1 \pmod{\ell^j}$ with $j \ge 1$, the binomial expansion gives $v^\ell \equiv 1 \pmod{\ell^{j+1}}$; iterate this argument. This is `dvd_one_add_mul_pow_sub_one`. In particular, the rows can each be specialized on their own line of approximate characters without giving up the interpretation of the matrices modulo ℓ^M.

8.7 Degree-free selection of roots of unity

We give the polynomial selection argument, since it supplies the main transition from the tied linear conditions to a single actual anticyclotomic character. It is also an example of a quantitative lemma whose conclusion is uniform in the degrees of arbitrarily many high-degree determinants, an essential feature here.

For a polynomial F over the integers of a finite ℓ-adic extension, and $0 < R \le 1$, write $|F|_R = \max_i |a_i|_\ell R^i$ for its Gauss norm, where $F(X) = \sum a_i X^i$. It is multiplicative and nondecreasing in R. We use this multiplicativity over the finite extensions containing the roots of unity in the following proof.

Lemma 8.12 (Simultaneous one-variable avoidance)**.** *Fix $0 < R < 1$ with $|\ell|_\ell < R^{\ell-1}$, and finitely many positive numbers u_i. For all sufficiently large N, in every sufficiently large coefficient ring containing the ℓ^N-th roots of unity, the following holds: if F_i are integral polynomials, of arbitrary degrees, and $|F_i|_R \ge u_i$, then there is a single ℓ^N-th root of unity ζ with*

$$|F_i(\zeta - 1)|_\ell \ge R u_i \qquad \text{for every } i.$$

The required depth is independent of the degrees.

Proof. Put $F = \prod_i F_i$ and $Q_i = \prod_{h \ne i} F_h$. Suppose, for contradiction, that at every ℓ^N-th root of unity at least one inequality fails. Take an integer $a \ge 1$ so large that $R^a < |\ell|_\ell$. For $0 \le j < N$ set $R_j = R^{1/\ell^j}$ and $n = \ell^{j+1}$. The roots α of $P_j(X) = (1+X)^n - 1$ all have absolute value less than R_j, since a primitive root of order ℓ^{j+1} satisfies

$$|\alpha|_\ell = |\ell|_\ell^{1/(\ell^j(\ell-1))} < R_j.$$

For each such root choose $i(\alpha)$ with $|F_{i(\alpha)}(\alpha)|_\ell < R u_{i(\alpha)}$. Let $b = a(j+1)$ and form

$$S_0(X) = \sum_{\alpha : P_j(\alpha) = 0} \frac{P_j(X)}{X - \alpha} F_{i(\alpha)}(\alpha)^b Q_{i(\alpha)}(X)^b.$$

The roots of P_j are simple. At each root, S_0 agrees with $P_j' F^b$, so their difference is divisible by P_j. Write $P_j' F^b - S_0 = P_j S$. For $T = R_j$ and $T = R_{j+1}$ we have $|P_j|_T = T^n$ and $|P_j'|_T = |n|_\ell$ (the latter because $P_j' = n(1+X)^{n-1}$). Moreover $u_i \le |F_i|_T$, so

$$|S_0|_T \le T^{n-1} R^b |F|_T^b < |n|_\ell |F|_T^b,$$

using $T \le 1$ and $R^{a(j+1)} < |\ell|_\ell^{j+1}$. The ultrametric inequality and multiplicativity give

$$|S|_T = \frac{|n|_\ell |F|_T^b}{T^n}.$$

Since $|S|_{R_j} \le |S|_{R_{j+1}}$, it follows that

$$\log |F|_{R_{j+1}} - \log |F|_{R_j} \ge \frac{n}{a(j+1)} \log(R_{j+1}/R_j) = \frac{-(\ell-1)\log R}{a(j+1)}.$$

Summing these inequalities over $j < N$ produces a contradiction for large N: the left hand side is at most $-\log |F|_R \le -\sum_i \log u_i$, whereas the harmonic sum on the right is unbounded. If a lower bound $u_i > 1$ makes the hypotheses impossible there is of course nothing to prove. This establishes the uniform assertion. □

This is the interpolation proof of `exists_root_of_unity_eval_large` in `New/Selmer/RootsOfUnityPoints.lean`.

Lemma 8.13 (Simultaneous selection in several variables)**.** *Fix a positive real δ, positive integers m, a, q, a nonnegative integer L_0, and finitely many nonnegative numbers B_i. For every sufficiently large N and every sufficiently large integral coefficient ring, suppose that $F_{i,j}$ $(1 \le j \le m)$ are integral polynomials in at most q variables. Suppose there exist tuples $u_{i,j}$ with $\ell \mid (u_{i,j,k} - 1)^a$ in each coordinate and*

$$v_\ell(F_{i,j}(u_{i,j})) \le B_i.$$

Let also D_h be monic one-variable polynomials with total degree at most L_0. Then there is one tuple ζ of ℓ^N-th roots of unity such that simultaneously

$$v_\ell(F_{i,j}(\zeta^j)) \le B_i + \delta, \qquad v_\ell(D_h(\zeta_k)) \le \delta \quad \text{for every } k, h. \tag{130}$$

Here the depth can be chosen before the polynomials or tuples are specified.

Proof. Choose $R < 1$ so close to one that $R^q, R^{L_0+1} > \ell^{-\delta}$ and $R^{\max(ma,\ell-1)} > |\ell|_\ell$. Assume that some coordinates have already been replaced by ζ_k^j in the points for each pair (i, j), and that the norm bound has deteriorated by at most a factor R per replacement. For the next coordinate, all other coordinates stay fixed (possibly different for different pairs). The polynomial in X given by substituting $(1 + X)^j$ in this coordinate has Gauss norm of radius R at least the currently established lower bound. To justify that assertion, the equation $(1 + X)^j = u_{i,j,k}$ has a root with $|X|_\ell \le R$: the monic polynomial $((1 + X)^j - u_{i,j,k})^a$ has constant term of absolute value at most $|\ell|_\ell < R^{ja}$; if all its roots had absolute value larger than R, their product would violate this bound. The value at such a root is exactly the previously known value of $F_{i,j}$, bounded above by the Gauss norm. We may use the algebraic closure to compare the norm and do not need the root itself in the coefficient ring.

Include in the one-variable family the polynomial $\prod_h D_h(1+X)$, whose monic leading term gives Gauss norm at least R^{L_0}. Apply Lemma 8.12 simultaneously. This selects one ζ_k for the next coordinate, with loss at most a further factor R, and also gives $|\prod_h D_h(\zeta_k)|_\ell \ge R^{L_0+1}$. Because each factor has norm at most one, this lower bound implies the same lower bound for every factor. After at most q steps, the accumulated factor is at least $R^q > \ell^{-\delta}$. The one-variable lemma is applied with the finite list of lower bounds $\ell^{-B_i} R^u$, $0 \le u \le q$, and R^{L_0}, so its threshold is uniform in the polynomial coefficients and their degrees. This yields (130). □

The source declaration is `exists_roots_of_unity_ellVal_le`, in the same file. An important application of the monic polynomials D_h is to $Y^j - 1$: at a selected tuple all the first m powers have nontrivial tame values, with total valuation as small as desired after accounting for the bounded number of variables.

8.8 The global twisting theorem and value families

The twisting theorem controls more than the Selmer modules. It also preserves the valuation of the special Hecke values needed in the norm comparison. The relevant *value family* (`IsFamily` in `New/Selmer/Corepresentable.lean`) is given by continuous unit-valued characters μ_i and a function $V(Q, \psi)$, defined for auxiliary conductors Q and characters. Its untwisted value $V(\varnothing, 1)$ is nonzero; at each fixed Q it is an integral finite linear combination of values of ψ, say

$$V(Q, \psi) = \sum_{a \in A_Q} c_a \psi(g_a), \qquad |c_a|_\ell \le 1. \tag{131}$$

In particular, on a character given by (127), this specializes an integral polynomial in Y_P. The family satisfies a norm relation in the form of an estimate: if $P \notin Q$ and ψ is anticyclotomic of conductor Q, the inequality

$$v_\ell(1 - \mu_i(\varphi_P)\psi(\varphi_P)^{\pm 1}) \le c \quad \text{for all } i$$

implies

$$v_\ell V(Q \cup \{P\}, \psi) \le v_\ell V(Q, \psi) + 8c. \tag{132}$$

For a ramified pair p the source constructs these families by an integral special-value measure, including the factors at the local level places. More precisely, the integral measure F satisfies a distribution relation upon adding v to the modulus: its value is multiplied by a product of at most four factors

$$\prod_{u \in K_0} \prod_{\epsilon \in \{1,-1\}} (1 - (\nu\vartheta)_{\mathrm{bc}}(\varpi_v)^\epsilon q_v^{u-z}), \quad K_0 \subseteq \{0,1\},$$

where z is the integer of the pair and ν its underlying character. The value on a twisted character is $F(1)\,\mathrm{N}(Q)^h \rho_p(\vartheta)^{h+1}$ for some nonnegative integer h. This is the formal-source input `Data.Pair.exists_isIntegralMeasure` in `New/Selmer/ValueFamilies.lean`; its derivation uses the algebraic CM Hecke special values and integral interpolation. The passage to the fixed-sign ratios and to the Galois-valued functions in (131) is `exists_isFamily` in that file. Because the character of a ramified pair is nontrivial at the chosen place w off ℓ, the listed auxiliary characters satisfy

$$v_\ell(\mu_i(z) - 1) \le v_\ell(q_w - 1) \le c_\ell \tag{133}$$

for its tame generator z. The last inequality follows from (122). Inverting the special-value ratio changes the valuation inequality to an upper bound on its ℓ-adic norm. No assumption of nondivisibility of every original Hecke value is needed.

Proposition 8.14 (Formal-source input: simultaneous anticyclotomic twisting)**.** *Let R_i be a finite collection of balanced representations with local bounds $\beta_{0,i}$ at depth M, and fix an inertial element z off ℓ as in Proposition 8.6. Suppose $v_\ell(\mu_i(z) - 1) \le c$ for the characters of finitely many value families. If*

$$\frac{\beta_{0,i}}{e_{\mathcal{O}}} + v_\ell \det(R_i(z) - 1) + v_\ell \det(R_i((z^c)^{-1}) - 1) \le B_i < \infty,$$

then for every $\eta > 0$ and every m there are an auxiliary set Q and an anticyclotomic character ψ such that, for each $1 \le j \le m$:

1. *the character ψ^j comes from $U(1)$ with conductor Q and is locally trivial at the prescribed old places and above ℓ, and*

$$\sum_{P \in Q} v_\ell(\psi^j(\tau_P) - 1) \le \eta;$$

2. *the normalized lengths of the cocycle modules of $\psi^j R_i$ at depth M, with the original crystalline relations and with value zero at the corresponding γ_i, are at most $B_i + \eta$, over all sufficiently large coefficient rings;*

3. *for each value family, $v_\ell V(Q, \psi^j) \le v_\ell V(\varnothing, 1) + 8c + \eta$.*

This is `exists_anticyclotomic_twist` in `New/Selmer/TwistTheorem.lean`, with the normalization of lengths as above. To see the use of the preceding lemmas, choose a pencil as in Proposition 8.10, using at most J_0 places. Choose a small allowance $\delta > 0$ with $(8J_0+2)\delta < \eta$. The construction is in two specialization stages. The first uses an ℓ-power root of unity u and the line $y_P = 1 + t_{P,i}(u-1)$ for the ith row, with $y_{P_0} = u$. By (129) all such points can be used modulo ℓ^M after the auxiliary depth has been chosen sufficiently large.

Use the sharp Fitting polynomial of Lemma 8.11 for the tied module, with the helper tie omitted. At $u = 1$ the polynomial has valuation bounded by that of the tied Selmer module plus the determinant terms for removing the helper tie. The one-variable avoidance lemma, applied also to the finitely many characteristic polynomials of the matrices at the nonhelper Frobenius places and the Euler factors for the value families, chooses a common u for which all these valuations increase by at most δ. This is why the pencil was required not merely to have suitable slopes but also to have the Frobenii of all nonzero slopes trivial on the corresponding row modulo ℓ^M. Along the chosen line the extra cost of dropping the other ties can then be bounded by Lemma 8.7: these are determinants of $uR_i(\varphi_P) - 1$ and their conjugate counterparts. These are precisely among the polynomials selected for the small-valuation conditions.

For each value family one may instead take the line with only the helper coordinate equal to u and the others equal to 1. Its initial valuation costs at most $8c$ by (132), using φ_{P_0} close to z and the bound (133). On inserting the other primes successively, their Frobenius values are the helper value by (125). By the monic Euler-factor avoidance in the first stage the remaining cost is at most 8δ per insertion. This bounds the value family on this new polynomial point. There is no need to find one tuple lying on the lines constructed for all different rows and value families.

At the second stage use the sharp Fitting polynomials for the unrestricted modules, now evaluated at the good row-specific points just constructed; also include the polynomials for all the value families. The number of variables is at most J_0, and by (129) all these specializing points satisfy the near-1 power condition with one common exponent. Lemma 8.13 therefore gives one tuple of ℓ^N-th roots of unity working for all polynomials, and simultaneously for their evaluations at the first m powers of that tuple. The monic factors $Y^j - 1$ ensure nontriviality and bound the valuations at the tame generators. The total error from these selection stages and from inserting the Euler factors is bounded by $(8J_0+2)\delta < \eta$. Formula (127) gives the character ψ with common conductor, and the finite coefficient rings can be enlarged without changing the normalized length bounds. This describes the proof's organization and supplies the polynomial and length arguments, while the exact global compatibility of the pencil and integral special-value measures is retained in the stated formal-source input.

Corollary 8.15 (Formal-source consequence: twisting all rows)**.** *For a finite collection of rows T_i of ranks r_i of admissible data, fixed ℓ, ι and any depth M, there is an m-fold twist θ such that for all $1 \le j \le m$, writing $\vartheta = \theta^j$ and ψ for its inverse avatar, one has*

$$\sum_{P \in Q} v_\ell(\psi(\tau_P) - 1) \le c_\ell, \tag{134}$$

$$\lambda\big(\mathfrak{S}_M(\psi T_i; S \cup Q \cup Q^c, \mathcal{L}) \cap \ker \mathrm{ev}_{\gamma_i}\big) \le (10 f r_i + 1) c_\ell, \tag{135}$$

$$\left|\iota(\rho_p(\vartheta, s)^{-1})\right|_\ell \le \exp(9 c_\ell \log \ell) \quad \textit{for every ramified } p, s. \tag{136}$$

The normalizer satisfies the scalar and unit conditions of SelmerLengthLe*.*

Deduction from the stated inputs. Take a chosen place w above one of g_0, g_1 different from ℓ. Proposition 8.5 applies for all rows after one common finite enlargement of the coefficient ring.

Choose $B_i = 10fr_i c_\ell$ and $\eta = c = c_\ell > 0$ in Proposition 8.14, using (133) for all the finite value families. The resulting character is the inverse avatar of a tame character, by the class-field identification `exists_tameChar_invAvatar_eq` of `New/Selmer/ValueFamilies.lean`. The nontriviality at the tame generators for all powers makes it an m-fold twist, while being trivial at the original places leaves the normalizers and crystalline conditions unchanged. The bounds on the normalized Selmer length and local ramification give the first two assertions. Using (132) and the ramified-pair norm relations in that file gives $v_\ell(\iota\rho_p(\vartheta, s)) \le 9c_\ell$, equivalent to (136). This is exactly `exists_mFoldTwist_selmerLengthLe` in `New/Selmer/TwistForRows.lean`. □

8.9 The Tate characters and triangular Taylor–Wiles places

The bound on the distinguished row block obtained above is not a bound on all congruences of the normalized column. The extra characters $\xi_e = \epsilon_\ell^{-e}$ also give torsion cocycles. It is not possible simply to replace their modules by a conjectural bound on the even cyclotomic class groups. The argument uses instead a finite collection of auxiliary primes that depends on the torsion depth but whose cardinality is bounded independently of that depth. These primes are distinct from the conductor of θ in the twisting theorem.

For a Tate character ξ_e, define $\mathfrak{E}_{\nu,e}(S)$ to consist of the locally constant cocycles z in $(\mathcal{O}/\ell^\nu)(-e)$ unramified outside the places above ℓ and $S \cup S^c$, with $z(g^c) = z(g)$ and vanishing on the prescribed subgroup of the local decomposition groups in the kernel of the cyclotomic character. For $e > 0$ the relevant odd module $\mathfrak{H}^-_{\nu,e}$ is given by the same unramified and local conditions without the added primes and with the odd symmetry condition. At $e = 0$ the odd-class package of the source uses a different convention, namely the unramified cocycles with its local condition without the odd restriction. These are the definitions `evenClasses` in `New/Selmer/Corepresentable.lean` and `shaOdd` in the odd-class bound. We do not implicitly decompose into two direct eigenspaces at $\ell = 2$; the use of sums and differences, and factors of 2, is important at that prime.

Let N_E be the cyclotomic conductor specified for E in the datum, so $E \subset \mathbb{Q}(\zeta_{N_E})$. The following background result is one of the few global class-number terms that survive in the final bound.

Proposition 8.16 (Formal-source input: one integer for all odd classes)**.** *There exists an integer $X > 0$ with*

$$X \le 2^{4(N_E+t+3)^{27}} h_E, \tag{137}$$

where h_E is the class number of E, such that for every ℓ, every $0 \le e < t$ and every ν, the normalized length of $\mathfrak{H}^-_{\nu,e}$ is at most $v_\ell(X)$. The coefficient rings are assumed, in addition to the integral DVR hypotheses, to contain enough roots of unity of order the exponent of $(\mathbb{Z}/N_E)^\times$.

This is `exists_length_shaOdd_le` in `Lib/Cyclotomic/OddSha/OddShaBound.lean`. It is a finite-level statement uniform in ν, not a generic assertion of Iwasawa nonvanishing. The proof uses the cyclotomic embedding of E, annihilating elements of Stickelberger type, the class-number bound in the trivial twist and explicit estimates of their sizes; it constructs the integer X by multiplying the corresponding annihilator factors. We retain this input as stated, since the bound (137), not just finiteness for every fixed (ℓ, e, ν), is needed for the global sum.

In contrast, the size of $\mathfrak{E}_{\nu,e}$ will not be included. To see the substitute, let $a \in G_E$ separate the distinct Tate characters up to ℓ-adic loss. A triangular system of depth ν consists of a finite ordered set J', a nondecreasing type function $\mathrm{ty} : J' \to \{1, \dots, t' - 1\}$, and distinct Taylor–Wiles places w_j with tame generators τ_j and Frobenius elements ϕ_j fixing the ℓ^ν-th

roots of unity. The places have prime norm and are conjugate-disjoint. Given bounds H_ℓ and C_ℓ, the `IsPrimeTriangle` predicate includes the following conditions:

1. There is a finite detecting set s of group elements such that evaluation at s detects all unramified-outside-ℓ Tate cocycles modulo ℓ^ν in the indicated range, and $(|J'|+|s|)H_\ell \le C_\ell$.
2. For each $0 \le e < t'$, the product $\prod_{k\neq e}(\xi_e(a) - \xi_k(a))$ divides ℓ^{H_ℓ} in the integral ring.
3. For every $z \in \mathfrak{E}_{\nu,\mathrm{ty}(j)}(\{w_i\})$, if $z(\tau_i) = 0$ for $i < j$ and $z(\phi_i) = 0$ for every earlier i of the same type, then $\ell^{H_\ell} z(\tau_j) = 0$.
4. Write $d_e(g) = \xi_e(g) - 1$. For $e > 0$, $4d_e(a)^2 \mid \ell^{H_\ell}$, and if an even cocycle z with no new ramification is zero at the ϕ_j of type e, then
$$d_e(a)(2z(g)) = d_e(g)(2z(a)) \qquad (g \in G_E).$$
Thus a controlled multiple of such a cocycle is a coboundary.

The precise predicate is in `New/Selmer/Corepresentable.lean`; writing a as in that predicate accounts for the inverse in the expression of the Tate characters.

Proposition 8.17 (Formal-source input: existence of triangular places)**.** *Suppose $4 \le f$, $N_E + t' < y$, $\epsilon_\ell(\gamma)^a - 1 \mid \ell^h$ for $0 < a < t'$, and $\ell \le y^2$ or $h = 0$. Assume the desired Frobenius conditions U hold on the generic Frobenius data near some element of the form $\delta^{\ell^\nu} g$ for each ν, g, where $\epsilon_\ell(\delta) = 1$. Then there exist an ordered index set J' and a type function, both independent of ν, such that at every depth $\nu > 0$ one can choose Taylor–Wiles places satisfying the triangle conditions with $a = \gamma^{-1}$, $H_\ell = yh$ and $C_\ell = y^{22}h$, and with all their Frobenius data satisfying U.*

This is `exists_isPrimeTriangle` in `New/Selmer/TrianglePrimes.lean`. The source constructs the detecting set uniformly and indexes the triangle by the product of its detecting set and the nontrivial types. The unusual order on this product, increasing by type, is necessary in the subsequent polynomial-support argument. One ingredient is a reciprocity functional z_0 on G_E, locally constant modulo ℓ^ν, which on the kernel of the cyclotomic character modulo ℓ^ν is an odd additive map transforming by cyclotomic weight $e+1$. At split places outside a finite exceptional set, if the functional is zero at all other Frobenius elements in a finite collection, its value on the remaining Frobenius, multiplied by 2, annihilates the corresponding even inertia value, by reciprocity. The source proves existence of such a functional, vanishing at a prescribed set of no more points than were previously used, and taking a unit value on some element of the cyclotomic kernel: `exists_isLaw` in `New/Selmer/ReciprocityLaw.lean`. The even cocycles are tested through all cyclic additive quotients, which detect zero by duality for finite abelian groups. By taking a linear combination of elements of the cyclotomic kernel and applying the products of differences of the cyclotomic weights, the construction makes the new functional vanish on all previously chosen Frobenius elements of other types. It then applies Chebotarev with the generic Frobenius conditions simultaneously. The annihilating error in this selection is the controlled product of differences above, not the length of the even module. This is the mechanism of the induction in Proposition 8.17; the integral construction of the reciprocity functionals and uniform detecting points is retained as part of the formal-source input.

For our datum the elementary separation of the characters is available with $h \le v_\ell(M_g)$ and $y = h_2 = r^8$. Indeed for ℓ larger than h_2^2 the field conductor is prime to ℓ and an inertial cyclotomic value whose reduction generates $\mathbb{F}_\ell^\times$ separates all exponents less than t

by units; take $h = 0$. For the smaller primes, choose a degree-one place above the chosen prime $g_i \neq \ell$. Its Frobenius has cyclotomic value g_i, and each $g_i^d - 1$ with $0 < d < t$ divides M_g by (122); take $h = v_\ell(M_g)$. This is `exists_cycChar_pow_sub_one_dvd` in `New/Selmer/DenominatorBound.lean`. The admissible size conditions give $N_E + t < h_2$, and the field contains sufficiently many infinite places. The generic Frobenius conditions needed in Proposition 8.17 arise from the Hecke string lemma discussed below. In particular the global odd term in (137) and the bounded valuations of the separating products are sufficient even if the even-class modules are very large.

8.10 The first-order Ribet construction

The algebraic argument works with a finite set Y of normalized column representations U_y over $\mathcal{O}$. A subset $\mathfrak{X} \subseteq Y$ consists of the distinguished lifts; for $y \in \mathfrak{X}$ the representation has block form $R_y \oplus \bigoplus_e \xi_e$ by (121), where R_y denotes the already twisted row block. For the projection we distinguish one index $y_0 \in \mathfrak{X}$. Use the group algebra $\mathcal{O}[G_E]$, and let a separate its Tate characters. Choose a normalizer γ with $R_{y_0}(\gamma) = sI$, $s - 1$ a unit, and $\xi_e(\gamma) = 1$. Put

$$\mathcal{N} = [\gamma a] - s[a] + p_0 D \sum_g n_g[g], \qquad \beta_y = \det\left(\sum_g \mathcal{N}_g(U_y(g) - 1)\right), \tag{138}$$

where p_0 is the row-cocycle modulus, D a separating factor, and the n_g are finitely supported nonnegative integers. This is `ribetNormaliser` in `New/Selmer/RibetCramer.lean`. For distinguished systems $\beta_y = 0$ since the block $\xi_0 = 1$ is annihilated by every $U_y(g) - 1$. On other systems one can arrange $\beta_y \neq 0$ whenever some $U_y(g) - 1$ is invertible by the following finite avoidance observation: each determinant is a polynomial in the finitely many n_g, nonzero as a polynomial if at least one leading matrix $U_y(g) - 1$ is invertible. Finitely many nonzero polynomials over a characteristic-zero field do not vanish on all tuples of nonnegative integers (induct on the number of variables). This is the reason the normalizer is allowed the final summand. One must still treat the systems on which every $U_y(g)$ has eigenvalue 1, rather than assuming those systems are precisely the lifts.

Let $A = \mathcal{O}[X_j : j \in J']$. We use a subalgebra $\mathfrak{S}$ of functions from Y to polynomials over the coefficient field, containing all scalar variables and characteristic polynomial coefficients of the $U_y(g)$. Its trace ideal $\mathfrak{J}$ contains, for all g, $p \operatorname{tr} U_y(g)$, with $p = \ell^\nu$, and the trace ties (for the two copies of j)

$$\operatorname{tr}((U_y(\tau_j) - 1)U_y(g)) - X_j \operatorname{tr}((U_y(\phi_j) - 1)e_j(U_y(\phi_j))U_y(g)), \tag{139}$$

where e_j is a polynomial with integral coefficients and $e_j(1) = 1$; conjugate inverse generators are included with the same variable. The precise abstract predicate `IsTraceIdeal` is in `New/Selmer/RibetFirstOrder.lean`. For that file the two conjugate copies may be prescribed independently but share X_j.

The Ribet element assertion is stronger than a mere membership of a determinant in an annihilator. It asserts the existence of $F \in \mathfrak{S}$ with

$$F\beta \in \mathfrak{J}, \qquad \beta_y = 0 \ (y \in \mathfrak{X}), \qquad [X^d]F_y = 0 \ (y \in \mathfrak{X},\ d < \mathbf{1}), \qquad [X^{\mathbf{1}}]F_{y_0} = c_0, \tag{140}$$

where $X^{\mathbf{1}} = \prod_{j \in J'} X_j$ and $d < \mathbf{1}$ means a proper coordinatewise divisor of this monomial. In addition $c_0 \neq 0$ has a quantitative upper valuation bound, given below. This is `IsRibetElement`. Crucially the vanishing condition is not that all terms of smaller total degree vanish; the source works with the lower order ideal of exponents coordinatewise at most $\mathbf{1}$. It is precisely this vanishing condition that permits the later first-order coefficient extraction at several primes.

Here are some details of the algebra of producing such an element. Because the distinguished block has full matrix span, one can construct group-algebra expressions isolating the Tate blocks. For each $y \in \mathfrak{X}$, write $R_y(\gamma_y) = s_y I$, $s_y - 1$ a unit, and $\xi_e(\gamma_y) = 1$. The element

$$k = \prod_{y \in \mathfrak{X}} (s_y - [\gamma_y])/(s_y - 1)$$

is integral, acts by zero on every distinguished row block, and by one on all Tate blocks. The factors need not commute: each row representation sends its own factor to zero, which annihilates the entire product in that representation. For each index e, Lagrange interpolation supplies an element q_e with $\xi_i(q_e) = D\delta_{ie}$, where D is a common multiple of the products of differences. Explicitly take a multiple of $\prod_{i \neq e}([a] - \xi_i(a))$ by integral coefficients. Both constructions remain valid in characteristic ℓ after retaining the multiplier D, whereas dividing by the differences would not. These are the two block-separation operations at the start of `ribet_first_order`.

To explain the Fitting module in this proof, let $\mathfrak{M}$ be the finite $\mathcal{O}$-module spanned, inside the family of matrices, by the elements $(U_y(g) - 1)_{y \in Y}$. It is finitely generated since Y is finite. A linear map from $\mathfrak{M}$ into a torsion row module produces a cocycle when it respects left multiplication by the $U_y(g)$, because

$$U_y(gh) - 1 = (U_y(g) - 1) + U_y(g)(U_y(h) - 1).$$

Conversely a cocycle obeying all linear relations between the family $(U_y(g) - 1)_y$ can be evaluated by a linear map on $\mathfrak{M}$ with the same equivariance. This motivates `obeyingCocycles` in `New/Selmer/BlocksAndCocycles.lean`. To compare with local crystalline classes it is useful to extend this graph map slightly: if the target is killed by a, Artin–Rees gives q with $\mathfrak{M} \cap a^q\mathfrak{V} \subseteq a\mathfrak{M}$, for the ambient matrix lattice $\mathfrak{V}$. Its evaluation therefore extends by zero on $a^q\mathfrak{V}$ to $\mathfrak{M} + a^q\mathfrak{V}$, and its graph is a subrepresentation with injective first projection. The map is well defined because differences in the intersection are killed. This is the lemma `obeyingCocycles.exists_graph_subrepresentation`; it also shows why the extracted cocycles are continuous and have the necessary unramifiedness, even though the definition by linear relations alone did not initially impose continuity. Where all U_y act trivially on inertia, the corresponding matrix differences are zero, so the cocycle is zero there. The graph description and continuity of the integral matrices show that a finite reduction of them determines the cocycle.

After imposing the trace ties and the relations of (138), one works in the polynomial extension of this finite presentation. At each distinguished block, a bound on the torsion length supplies a determinant (a maximal minor of a relation matrix) with small valuation. For the Tate block one must use a first-order variant: the relations between evaluation at a pair of conjugate inertia elements and evaluation at the Frobenius elements, multiplied by X_j, make the corresponding determinant divisible by the products of X_j of that type. To state its precise arithmetic input, for a module Z of Tate cocycles put $I_e = \{j : \mathrm{ty}(j) = e\}$. The triangle cost with multiplier D and allowance b_e consists of a set L of evaluations such that, for cocycles zero on L, agreement at earlier pairs of tame generators and at earlier Frobenius pairs of the indicated type forces

$$D\big(z(\tau_j) - z((\tau_j^c)^{-1})\big) = 0 \qquad (j \in I_e).$$

For submodules of these cocycles vanishing at all tame generators and agreeing at the Frobenius pairs of type e, their length modulo D-torsion plus the length cost of the initial evaluations on L is at most b_e. This is `HasTriangleCost`.

Proposition 8.18 (Formal-source input: first-order Ribet bound)**.** *Use the preceding block setup and trace ideal. Suppose the normalized row cocycle module, modulo p_0, obeying the linear matrix-difference relations and vanishing at the new tame generators and its normalizer, has ordinary length at most $b_0 < v_\pi(p_0)$. Suppose the corresponding Tate modules, modulo p, have triangle costs b_e and detecting set s as above. Include auxiliary inertial relations of the form $(U_y(\delta_v) - 1)o_v(U_y(\delta_v)) = 0$, where $\xi_e(\delta_v) = 1$. Assume also $e_j(\xi_e(\phi_j)) \equiv 1 \pmod p$ and $\prod_{k\neq e}(\xi_e(a) - \xi_k(a)) \mid D$. Then the Ribet element (140) exists with*

$$v_\pi(c_0) \le b_0 + \sum_e b_e + t'\left(\sum_v v_\pi(o_v(1)) + 2|s|v_\pi(D)\right) + (|J'| + 1)v_\pi(D). \tag{141}$$

Here the Tate-module assertions of triangle cost and detection are required for the modules whose members vanish on all δ_v and whose coboundaries are killed by $\xi_e(a) - 1$ (as in the source).

This is `ribet_first_order` in `New/Selmer/RibetFirstOrder.lean`. The block Fitting lemma `exists_mem_fittingIdeal_block` in the same file gives, for one Tate block, a relation in the Fitting ideal of the form $G_e \prod_{j\in I_e} X_j$ with constant-term bound $v_\pi(G_e(0)) \le b_e + (|I_e| + |s|)v_\pi(D)$. The proof successively enlarges presentations, uses the stated cost to clear one evaluation at a time, and then multiplies the resulting minors. When combining blocks, terms not supported in the interval of divisors of $\prod X_j$ can be discarded. More precisely, work modulo the ideal spanned by monomials not dividing $\prod X_j$, and at each step additionally keep track of previous types; the type ordering puts the unwanted earlier variables in the annihilator after extracting the earlier block factors. The direct-sum decomposition of the distinguished representations turns the determinant of a block-flattened presentation matrix into the product of its row determinant and its Tate determinants. The factors D and $o_v(1)$ are inserted to make the block operations integral, accounting for the additional losses in (141). Finally a group-algebra Cramer determinant identity turns these minors into an element of the trace ideal after multiplying by β. That identity is `exists_ribet_element_cramer` in `New/Selmer/RibetCramer.lean`; it uses integrality of characteristic polynomials of block matrices over the group algebra, supplied by the indicated Amitsur formula. These Fitting, block Cramer and multivariable support steps are specialized assertions of the formal-source proof, not consequences of the elementary determinant inequality alone.

In applying the triangle cost to a Tate component, choose L to be the prescribed local set above ℓ. The length lost by imposing vanishing there is bounded by the crystalline defect, since crystalline Tate cocycles vanish on that set. After imposing vanishing there, taking $z + z^c$ produces an even cocycle, and the triangle condition gives the required agreement of the conjugate pairs up to the factor D. Once all new inertia values vanish the cocycle is unramified outside ℓ. Its odd part $z^c - z$ now provides an element of the odd module, and the even part, by the triangle condition, is a coboundary up to a factor of controlled valuation. The assumption on coboundaries makes the kernel of this odd projection killed by D; in particular its length modulo D-torsion is bounded by the length of the odd module. The assertion is `hasTriangleCost_tateTwist` in `New/Selmer/TateTwistSteps.lean`; it treats $e = 0$ directly. Thus b_e is the local defect plus the length in Proposition 8.16, without any even-class length term.

8.11 Hecke operators realizing the first-order relations

The formal-source proof now passes from the group algebra to forms. For a fixed auxiliary pair, label tuple and twist, choose the base space at the corresponding level and consider its

enlargements at the Taylor–Wiles places. On each of these finite-dimensional spaces there is a finite set of eigensystems, and a closed integral Hecke algebra of scalar functions of the eigensystem preserving the lattice of Weyl-integral forms. The characteristic polynomial coefficients of the normalized Galois representations, as functions of the eigensystem, belong to this algebra: the attached representation comparison provides them as limits of integral Hecke coefficients on generic split Frobenius data. In particular $\mathrm{tr}(P(U_y(h))U_y(g))$ belongs to the algebra for every integral polynomial P, by expanding P and using the trace at each $h^i g$. This is `BaseLevel.charpoly_mem_heckeAlgebra` and `BaseLevel.traceElem` in `New/Patching/OnePrimeTie.lean`, with the closedness and local–global compatibility of the Hecke algebra taken as formal-source inputs.

A system is called weakly lifted for this step if every $U_y(g)$ has eigenvalue one; call this set $\mathfrak{W}_1$. It is potentially larger than $\mathfrak{X}$. A normalizer can be chosen as in (138) so that $\beta_y \neq 0$ off $\mathfrak{W}_1$, but the argument does not try to separate $\mathfrak{W}_1 \setminus \mathfrak{X}$ by this determinant. A further polynomial $P(\beta)$ over the ℓ-adic integers is chosen to vanish at all the nonzero determinant values on the old eigensystems and to have nonzero constant term (take the product of their linear factors, with integral scalar factors as necessary). Thus $P(\beta) \neq 0$ on $\mathfrak{W}_1$ and vanishes on all other old systems. It is not asserted that the coefficients of P or the divisions needed in clearing this cutoff are units.

Proposition 8.19 (Formal-source input: Hecke string)**.** *For a weakly lifted eigensystem of an auxiliary pair, its rescaled unramified Hecke polynomial at a generic split place of prime norm q_w has the full string of roots $1, q_w, \dots, q_w^{t'-1}$. More precisely, if its normalized Galois characteristic polynomial at a generic Frobenius is congruent coefficientwise, modulo the maximal ideal, to*

$$S(X) \prod_{i<t'} (X - \epsilon_\ell(\phi)^{-i}) \quad \textit{with } |S(1)|_\ell \geq 1,$$

then its rescaled Hecke polynomial factors as the corresponding full string times a residual factor dividing an integral polynomial $e(X)$ with $e(1) = 1$.

This is `isRoot_hecke_string` and `eventually_heckePoly_eq_string_mul` in `New/Selmer/HeckeString.lean`. The underlying representation-theoretic input is the string lemma for spherical unitary representations and the prescribed local type. The integral polynomial assertion follows from comparing monic string factors modulo the maximal ideal and from the bound $|S(1)|_\ell \geq 1$: for a monic B dividing the string, the Gauss norm of a product with B equals the product of the norms; hence comparison of AB and $A'B$ forces comparison of A, A'. Once the residual factor A is integral and $A(1)$ a unit, use $e = A/A(1)$. This is the elementary part of `exists_integral_multiple_eval_one`. For finitely many relevant eigensystems one takes the product of the resulting polynomials, still with $e(1) = 1$.

The required Frobenius conditions in Proposition 8.17 ensure that this proposition can be applied at every chosen auxiliary prime simultaneously. A simple reason such conditions are available is that at the distinguished row block a scalar inertial element with long prime-to-ℓ order allows one to move a Frobenius condition without changing its cyclotomic value. If an integral matrix A has rank r', the polynomial in Y given by $\det(1 - YA)$ has degree at most r' and constant term one. Among $r' + 1$ distinct residual scalars Y, at least one avoids its roots. The scalar-inertia condition of the row provides such scalars even after taking large ℓ-powers of the inertial element. By continuity the nonvanishing is a neighborhood condition. The formal statement is `exists_eventually_no_eigenvalue_one`. The attached representations and the generic Frobenius filter then yield the simultaneous string conditions for the eigensystems congenial to the target residual system. This is the condition imposed in the choice of the triangular primes in `denominatorLe_rowBudget`.

To realize the relations (139), at a Taylor–Wiles place let $\Gamma = \mathrm{GL}_{r'+t'}(k_w)$ and $H \subseteq \Gamma$ the stabilizer of a line. The character of H to $\mathbb{Z}/\ell^\nu$ given by the logarithm of its scalar action on the line is a diamond-operator character. Its transfer to Γ factors through the determinant. The relevant normalization, after multiplying by $(r'+t')!$, sends a chosen primitive determinant scalar to one modulo the logarithmic modulus. This follows from the transfer identity on scalar matrices and the count of lines; it is `exists_line_stabilizer_log` in `New/Patching/OnePrime.lean`. The proof then constructs a first-order rather than a semisimple deformation of the trace. At one place, if ρ denotes the residue-group action, define for a function f_0 on the logarithmic group

$$W(f_0) = \sum_{h \in H} f_0(\log h)\rho(h).$$

Changing h to $g^{-1}h$ or shifting the logarithmic parameter shows that multiplying by a diamond element shifts the argument of f_0. Taking a representative of the logarithm gives the infinitesimal version of this identity modulo ℓ^ν. The source performs the induced-representation transfer, using coset representatives for Γ/H, to construct a local operator F_j and the norm-sum operator $N_j = \sum_{\gamma \in \Gamma} \rho(\gamma)$, with the trace relations acting as follows: if $t_j(g) = \mathrm{tr}((U(\tau_j) - 1)U(g))$ and $r_j(g) = \mathrm{tr}((U(\phi_j) - 1)e_j(U(\phi_j))U(g))$, then, modulo ℓ^ν and up to the specified integer factors, $F_j t_j(g) = N_j r_j(g)$ and $N_j t_j(g) = 0$. There are additional identities controlling how F_j moves the old-space invariants. This is the formal-source input `BaseLevel.exists_tie` in `New/Patching/OnePrimeTie.lean`, proved using `one_prime_first_order` in `New/Patching/OnePrime.lean` and local line-stabilizer operators. The nonzero factorial factor is retained rather than divided out at the torsion level.

The multiple-prime coefficient extraction is motivated by the following simple polynomial observation. If operators F_j, N_j at distinct places commute, multiplying the first-order relation by products of operators at the other places transports the contribution from a monomial X^d to the monomial X^{d+e_j}. Terms with exponents larger than one are discarded. The product of the sums N_j is the full residue-group norm mapping to the old-level subspace, while the complementary product of the F_j probes its infinitesimal deformations. Thus evaluating the trace-ideal relations on this first-order coefficient system annihilates the trace ideal modulo ℓ^ν. On the distinguished systems, all coefficients of F in (140) properly dividing $\prod X_j$ vanish; only the top coefficient remains. The source proves that this coefficient gives an integral diagonal operator at old level, localized to the cutoff determined by $P(\beta)$, after controlling the movement from the old space. The auxiliary depth ν must exceed the valuation of an explicit factor involving the constant term of P, the level automorphism factors and the factorial. It can be made arbitrarily large since the index set of the triangle was bounded before ν was chosen.

Precisely the result needed is `BaseLevel.exists_diagOp_mem_ellIntegralForms` in `New/Patching/OnePrimeTie.lean`, a formal-source input. Given the string condition near the distinguished residual eigensystem, enough divisibility by ℓ^ν, and a Ribet element of top-coefficient valuation with a factor M_0 at most Γ_0, it supplies a scalar function s of the eigensystem with $v_\ell(M_0 s(y_0)) \leq \Gamma_0$. The diagonal operator with scalars s sends a vector supported on $\mathfrak{X}$ to an integral vector up to a specified level factor and any original denominator, provided that vector agrees with an integral old form on the eigensystems where $P(\beta) \neq 0$ in the sense of the diagonal-operator test in the source statement. Thus the integral output does not require that all weakly lifted systems themselves belong to $\mathfrak{X}$. A tested form minus its projection to Z is orthogonal to the unwanted weakly lifted components by the Hecke string assertion and (119), after inserting the chosen label projector. It follows that the diagonal operators satisfy the required integrality hypothesis on the projected tested forms.

8.12 The denominator bound from the local length budgets

We now combine the congruence terms numerically. Write b_1 for the rounded value of the thick weight and let

$$F_\ell^{\mathrm{num}} = (2b_1)! \left(\prod_{\max(b_1-30h_2^3,0)<i\le b_1+30h_2^3} i \right)^{h_2^3}.$$

This integer is independent of ℓ; the notation emphasizes its use in the local estimate. With $g = g_0 g_1$, define

$$\mathrm{loss}_\ell = 10^6 \frac{m}{k} \left(b_1 v_\ell(g) + v_\ell(F_\ell^{\mathrm{num}}) \right), \qquad D_\ell = h_2^{27}(c_\ell + v_\ell(X)) + 4f\,\mathrm{loss}_\ell .$$

Here X is from Proposition 8.16. These are `loss`, `lambdaLoss` in `New/Selmer/DeterminantBound.lean`, and `rowBudget` in `New/Selmer/RibetElement.lean`. The detailed local comparison of torsion cocycles and the crystalline relations is developed in the next part of the proof. Its crucial formal-source bound, `sum_blockLoss_le` in `New/Selmer/RibetElement.lean`, is

$$\sum_{i \text{ blocks}} \lambda(\text{loss of crystalline relations on block } i) \le 4r' f\,\mathrm{loss}_\ell, \tag{142}$$

for arbitrary torsion depths of the individual blocks. This controls the failure of extracted cocycles to satisfy the crystalline local conditions in the Selmer theorem. One cannot apply that theorem directly to all cocycles extracted from congruent matrix differences.

Proposition 8.20 (Formal-source input: arithmetic Ribet element)**.** *Use the apt columns and rows of an auxiliary pair and a twist satisfying* (134) *and* (135) *at a depth $M > \lfloor rD_\ell \rfloor$. Suppose their normalized matrices satisfy* (121) *on the distinguished systems, and for new twist inertia the normalized columns satisfy $(U(g) - \psi(g))(U(g) - 1) = 0$. At a triangular auxiliary level, with the trace ideal as above, the determinant from* (138) *has a Ribet element with top coefficient c_0 such that*

$$v_\ell(M_g^3 c_0) \le r' D_\ell.$$

This is `ribet_columns` in `New/Selmer/RibetElement.lean`. The required quadratic inertia relation is `normalisedRep_quadratic_on_inertia` in `New/Selmer/LabelledColumns.lean`; it follows from the shape of the parahoric type at the twist conductor. The cocycle length for the row block is at most its normalized Selmer bound $(10fr'+1)c_\ell$ plus its share in (142). The Tate costs in (141) contribute at most $v_\ell(X)$ per block plus their shares in (142). The scalar-separation and detecting terms contribute multiples of $h_2 h$ and $h_2^{22} h$; the old and new tame polynomial factors have valuations bounded by the number of places times c_ℓ and by the sum in (134). The aggregate of these noncrystalline terms, including the factor M_g^3, fits inside $r' h_2^{27}(c_\ell + v_\ell(X))$ by the admissible size inequalities. The source proves the explicit inequality in the last step of `ribet_columns`. Notice why the row modulus M is chosen larger than rD_ℓ: it is larger than the resulting bound on the row-cocycle length, as required in the first-order Ribet lemma.

We give the final linear algebra of combining individual integral operators, since it avoids multiplying the denominator cost by the number of distinct eigensystems. Let $\{p_i\}_{i=1}^q$ be a basis of Z consisting of nonzero vectors from the spanning family of lifted forms. Given integral test vectors u_a, v_b in the denominator condition, write $P_Z v_b = \sum_i P_{ib} p_i$. Suppose for each j a diagonal comparison operator acts on p_i by scalars a_{ji}, after a common injective comparison map, with $|M_0 \iota(a_{jj})|_\ell \ge e^{-\Gamma}$, and that the matrices

$$(\iota \langle u_a, \textstyle\sum_i a_{ji} P_{ib} p_i \rangle M_0)_{a,b}$$

are integral for every j. Choose coefficients t_j in the integral valuation ring of $\overline{\mathbb{Q}}_\ell$ so that for every i,

$$|M_0 \sum_j t_j \iota(a_{ji})|_\ell \geq |M_0 \iota(a_{ii})|_\ell \geq e^{-\Gamma}.$$

This is possible since for each i the finite linear polynomial has its Gauss norm attained at some tuple of units, and a finite set of nonzero linear polynomials over the infinite residue field can be avoided simultaneously after dividing each one by a coefficient of maximal absolute value. The linear combination of the integral pairing matrices is integral. Its determinant equals the determinant of the matrix $(\iota\langle u_a, P_Z v_b\rangle)$ multiplied by $\prod_i (M_0 \sum_j t_j \iota(a_{ji}))$, by writing all pairing matrices as a product of the basis-change matrices and the relevant diagonal matrices. Consequently the absolute value of the desired determinant is at most $e^{q\Gamma}$. This is the linear algebra in `norm_gram_det_le_of_operators` in `New/Selmer/DeterminantBound.lean`.

Apply the arithmetic Ribet element proposition and the first-order trace deformation over the triangle to a single basis eigensystem in the space of lifts. In the setting of the tested forms, the local operator integrality described above, together with the class-sum pairing and automorphism bounds absorbed into M_g^3, provides exactly the operators for this linear algebra with $\Gamma = r' D_\ell \log \ell$. This yields

$$\mathrm{DenominatorLe}(\iota, Z, D_\ell \log \ell), \tag{143}$$

that is (120) with $\beta = D_\ell \log \ell$. The precise formal-source theorem, including the spanning hypothesis that each generator is apt, lifts a fixed system having a row for all twists, and has the needed Selmer bound, is `denominatorLe_rowBudget` in `New/Selmer/DenominatorBoun d.lean`. The use of tested forms, not arbitrary forms, is why the weakly lifted eigensystems suffice. Equation (143) marks the end of the congruence argument.

8.13 Global numerical summation and proof of the fit theorem

All the terms in the local allowances are nonnegative, and the factorization of each positive integer A gives $\sum_{\ell \in T} v_\ell(A) \log \ell \leq \log A$ for every finite set of primes T. The other part of c_ℓ has bounded sum, since $\sum_{n \geq 2} 2^{-n} \log n < \infty$. Therefore

$$\begin{aligned}
&\sum_{\ell \in T} 9 c_\ell \log \ell \leq 9(\log M_g + \sum_{n \geq 2} 2^{-n} \log n), \\
&\sum_{\ell \in T} D_\ell \log \ell \leq h_2^{27} \left(\log M_g + \log X + \sum_{n \geq 2} 2^{-n} \log n \right) \\
&\qquad\qquad + 4 \cdot 10^6 \frac{fm}{k} \left(b_1 \log g + \log(2b_1)! + h_2^3 \sum_{\max(b_1 - 30h_2^3, 0) < i \leq b_1 + 30h_2^3} \log i \right).
\end{aligned}$$

These estimates reduce the infinite collection of local lengths to a handful of ordinary integer sizes. For example $\log(2b_1)! \leq 2b_1 \log(2b_1)$, and the short-interval contribution is at most $60h_2^6 \log(b_1 + 30h_2^3)$. Also $\log M_g$ is bounded by a constant times $f(rt)^2 \log g$ by the order formula for finite general linear groups, and $\log X$ is bounded by $4(N_E + t + 3)^{27} \log 2 + \log h_E$. The explicit admissibility inequalities in the source bound these contributions by

$$\sum_{\ell \in T} 9 c_\ell \log \ell \leq x^{999}, \qquad \sum_{\ell \in T} D_\ell \log \ell \leq B_0. \tag{144}$$

See the numerical calculation in `exists_isFitFamily` in `New/Selmer/FitTwistsExist.lean`, using the class-number estimate and `Data.le_of_hasLogBound`. These are formal-source

bounds with the full numerical hypotheses on the datum. To see their scale, along the admissible family with fixed $\varepsilon > 0$, b exceeds every fixed power of x; h_2, N_E, r and t are polynomial in x. Thus the first term of the second bound is smaller than B_0 by a power of b. In the remaining factorial term b_1 has order b, $\log g$ and $\log b_1$ have order $\log b$, and the leading contribution is $O(fmb \log b/k)$. Its ratio to $B_0 = fa(\log b)^3/(mk)$ is $O(m^2 b/(a(\log b)^2)) = O((\log b)^{-2})$ by $m^2 = O(x)$ and $a \geq xb$. This highlights the logarithmic slack without replacing the exact inequalities by asymptotic claims.

Proof of Theorem 8.2, using the stated formal-source inputs. For each ℓ and ι, make the finite list of rows in Proposition 8.3, including all rank-one pairs, before choosing any twists. Apply Corollary 8.15 to this list at depth $M = \lfloor rD_\ell \rfloor + 1$. The resulting m-fold twist θ_ι has conductor prime to ℓ and satisfies the local conditions for its first m powers. At each arrow j, the good lift is spanned by the apt systems with the chosen labels and this common Selmer bound. For the rank-one lines the corresponding assertion follows from their apt label $X - 1$ and the rank-one row. Consequently (143) applies simultaneously to the desired spaces, with the same allowance $\beta_{j,\ell} = D_\ell \log \ell$ independent even of j. The second inequality of (144) is its required finite-sum bound. The special-value inequality (136) and the first inequality of (144) give the required height bound on the inverse ratios: in the height definition take the constant archimedean factor 1 and the local log allowance $9c_\ell \log \ell$. All these choices are independent between the pairs (ℓ, ι) but uniform over the finite lists for a given pair, as required. Thus the family satisfies Definition 8.1. This reproduces the final assembly of `exists_isFitFamily`. □

The upper bound on projection denominator and the lower bound on theta-lift covolume now meet at exactly the same error budget. The global Selmer argument has been used to produce the local denominator estimate, not to establish a vanishing theorem for a class group, a modularity lifting theorem, or a nonvanishing assertion about automorphic L-functions at critical points. In particular the omission of any even-class-group length term is a consequence of the first-order ties at the auxiliary triangle, and not of assuming those classes vanish.

9 Local crystalline defects and the integral cup-product estimate

The congruence argument for the lower bound on the determinant needs considerably more than the vanishing of the ordinary Bloch–Kato pairing on two crystalline representations. The two classes to be paired arise from lattices which need not be dual to one another in characteristic zero; only a pairing of their reductions is given. Moreover the lattices occurring in the congruences have weights and ranks increasing with the main parameter x. An error term depending exponentially on a rank, or even a loss of order b on every line of a rank r lattice at every small prime, would not suffice. This section formulates the local integral problem, explains the Kummer-chart calculation used to solve it, and shows exactly how the resulting loss enters the global congruence budget. The integral period-ring assertions quoted here as *formal-source inputs* are statements of the compact Lean development, not assertions for which we claim an independent verification of the proof. In contrast, we give the length, valuation, and budget deductions explicitly.

9.1 Lattices, crystalline cocycles and imperfectly dual reductions

Write $p = \ell$ for the rational prime throughout the local discussion, $G_p = \mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$, and let $j : \Gamma \hookrightarrow G_p$ identify Γ with an open subgroup. Write F for the corresponding finite extension

of $\mathbb{Q}_p$ and $d_F = [F : \mathbb{Q}_p]$. The coefficient ring O denotes the integers of a sufficiently large finite extension of $\mathbb{Q}_p$, with fraction field K_O. For its uniformizer π_O, put $e_O = v_{\pi_O}(p)$; our valuation v_p on K_O has $v_p(p) = 1$. In particular the normalized length of a torsion O-module M is

$$\operatorname{len}_p M = e_O^{-1} \operatorname{length}_O M.$$

It is additive on short exact sequences and $\operatorname{len}_p(O/p^aO) = a$. Infinite lengths can be understood as $+\infty$. For submodules S, T of a common module, we use the notation

$$\operatorname{rl}(S,T) = \operatorname{length}_O(S/(S \cap T)), \qquad \operatorname{rl}_p(S,T) = e_O^{-1} \operatorname{rl}(S,T).$$

These correspond to `relLength` and `CoeffRing.normLength` in the formal source. The normalization is essential: passing to a more ramified coefficient field must not artificially enlarge the estimates. For a finite-length module over O, its length over $\mathbb{Z}_p$ is $[O/\pi_O : \mathbb{F}_p]$ times its length over O. Consequently

$$\operatorname{rl}_p(S,T) = \frac{\operatorname{length}_{\mathbb{Z}_p}(S/(S \cap T))}{[K_O : \mathbb{Q}_p]}. \tag{145}$$

The statements below are uniform in O once the necessary finite coefficient fields and roots of unity are included. This is the meaning of the eventual quantifier on coefficient rings in the source.

A Galois lattice T is finite free over O, with continuous O-linear Γ-action. Put $T^*(1) = \operatorname{Hom}_O(T, O(1))$, where the cyclotomic character acts on $O(1)$. The jump convention is that $O(1)$ has jump -1. At an embedding $\gamma : O \to \mathbb{C}_p$ let $j_T(\gamma, u)$ denote the multiplicity of jump $u \in \mathbb{Z}$; thus

$$j_{T^*(1)}(\gamma, -u-1) = j_T(\gamma, u). \tag{146}$$

The local development defines $j_T(\gamma, u)$ by the dimension over $\overline{\mathbb{Q}}_p$ of the space of γ-semilinear functionals $f : T \to \mathbb{C}_p$ satisfying $f(gv) = \chi_{\mathrm{cyc}}(g)^{-u} g f(v)$ for g sufficiently close to 1. In particular its weights are Hodge–Tate weights in the above sign convention, not the oppositely signed integers sometimes used for cyclotomic twists.

The O-module of crystalline cocycles $Z_f(T)$ consists of maps $c : \Gamma \to T$ such that there is $y \in B_{\mathrm{cris}} \otimes_{\mathbb{Z}_p} T$ with

$$1 \otimes c(g) = gy - y \qquad (g \in \Gamma).$$

The tensor product here is over $\mathbb{Z}_p$, and the action is diagonal. The definition is useful also for potentially crystalline T. This is the inverse image in $Z^1(\Gamma, T)$ of the usual kernel of the map to $H^1(\Gamma, B_{\mathrm{cris}} \otimes_{\mathbb{Q}_p} T[1/p])$. The cocycle identity follows by applying $gh - 1 = (g-1) + g(h-1)$ to y and using injectivity of the natural inclusion. The module includes all integral coboundaries. It is saturated with respect to multiplication by a nonzero element of O: if ac is a crystalline cocycle, divide its period primitive by a in the tensor product with B_{cris}, where K_O acts and a is invertible. In particular

$$Z_f(T)/p^a Z_f(T) \hookrightarrow Z^1(\Gamma, T/p^aT). \tag{147}$$

The formal source also proves continuity, including the eventual local constancy of reductions of such cocycles. The relevant definitions and results are `crysCocycles`, `jump` and `HasJumps` in `Lib/PadicHodge/Crystalline/Defs.lean`, `crysCocycles.of_smul` in `Lib/PadicHodge/Crystalline/Jumps.lean`, and `crysCocycles.eventually_eq_pow_smul` in `Lib/PadicHodge/BlochKato/CocycleContinuity.lean`.

Proposition 9.1 (Finite local cocycle rank)**.** Formal-source input. *Let T be potentially crystalline. Then $Z_f(T)$ is finite free over O. Rank over O is at most $(1+d_F)\,\mathrm{rk}_O\,T$. More precisely, in the situation where the jumps are indexed by embeddings of the completion of E in $\overline{\mathbb{Q}}_p$ and hence by a set Σ_v of complex embeddings of E determining a finite place $v \mid p$,*

$$\mathrm{rk}_O\, Z_f(T) = \mathrm{rk}_O\, T + \sum_{\sigma\in\Sigma_v} \#\{u \in J_T(\sigma) : u < 0\}.$$

The same formula with a local field of degree d_F and with summation over its embeddings applies after extending coefficients so that all the Hodge jumps can be read there.

This is the Bloch–Kato rank calculation on cocycles (as opposed to on classes), expressed by `crysCocycles.finrank_eq` in `Lib/PadicHodge/BlochKato/BlochKato.lean` and the embedding form `finrank_crysCocycles_eq_add_neg_jumps` in `New/Selmer/CrystallineRank.lean`. The rank formula on classes comes from the fundamental period-ring exact sequence and the de Rham filtration; adding the dimension of the coboundaries cancels the contribution of invariants. The formal-source statement on $\mathbb{Z}_p$-ranks uses d_F times the number of negative jumps in the multiset of a restriction of scalars. In the indicated embedding setting, decomposing that multiset and dividing by $[K_O : \mathbb{Q}_p]$ gives the asserted formula. Freeness follows from torsion-freeness and finite generation over a DVR. In particular the rank bound requires no extra hypothesis on inertial eigenvalues.

We describe the class of lattices against which the defect is measured. Let U be a family of finite subsets of $\mathbb{Z}$ indexed by coefficient embeddings. The notation $\mathcal{Q}(F_0, \mathcal{K}, U)$ will mean the class `Lift.InClass` of the formal source; here F_0 is the filter on Γ that specifies the required fixed crystalline periods, and $\mathcal{K}(g, Q)$ specifies a family of polynomials annihilating the operator on potentially crystalline periods associated with g. More explicitly a member Y of the class is admissible for periods fixed almost everywhere for the supremum of the near-identity filter and F_0, has $j_Y(\gamma, u) = 0$ for $u \notin U(\gamma)$, and, for g mapping to inertia, each polynomial Q with $\mathcal{K}(g, Q)$ annihilates the linear operator $\beta_{g,0}$ on $B_{\mathrm{cris}} \otimes_{\mathbb{Z}_p} Y$. This operator is the B_{cris}-linearization of g on the potentially crystalline periods (the source calls it `linearFrob`). For $F_0 = \top$, $\mathcal{K} = \bot$ we just require crystallinity and the specified weights. All members are potentially crystalline. The jump and polynomial conditions are preserved when a lattice is a subrepresentation of a direct sum of the given representations. The dual of such a class has jump sets $-U - 1$ and inertial polynomial condition at inverse elements. These facts are expressed in `InClass.of_jointly_injective`, `InClass.dual` and `InClass.tensor`, `Lib/PadicHodge/Crystalline/PairingDefs.lean`.

An imperfectly dual pairing of lattices Y, T is specified by a positive modulus p^a, an O-linear form $\psi : Y \otimes_O T \to O(1)$, a $\mathbb{Z}_p$-linear functional $\lambda : O \to \mathbb{Z}_p$, and a crystalline cocycle $q \in Z_f((Y \otimes T)^*(1))$ satisfying

$$g\psi - \psi = p^a q(g). \tag{148}$$

Thus the induced bilinear form $\bar{\psi}$ with values in $\mathbb{Z}/p^a(1)$, obtained by applying λ and reducing modulo p^a, is equivariant; the unreduced pairing need not be equivariant. The cup product of $c \in Z_f(Y)$ and $c' \in Z_f(T)$ under $\bar{\psi}$ is represented by

$$(g, h) \longmapsto \bar{\psi}(c(g), gc'(h)) \in \mathbb{Z}/p^a(1).$$

The vanishing required in the source means that, after the inclusion of the cyclotomic module $\mathbb{Z}/p^a(1)$ into $\overline{\mathbb{Q}}_p^{\times}$, the displayed cocycle is a continuous coboundary. One should not replace this assertion by an assertion of vanishing in $H^2(\Gamma, \mathbb{Z}/p^a(1))$ without an argument. The chosen codomain is the one used by the index form of local Tate duality below. Both the imperfect equivariance and the third class q are necessary: it would be illegitimate simply to apply the

pairing of $D_{\mathrm{cris}}(T)$ with its exact dual to a representation merely congruent modulo p^a to that dual.

Define the submodule $T^{\perp_{\mathcal{Q}}}$ of maps $c' : \Gamma \to T$ by imposing this cup-product vanishing for every $Y \in \mathcal{Q}$, every $c \in Z_f(Y)$, and every pairing of the above form, including every modulus. Define its crystalline defect

$$\Delta_p(T; \mathcal{Q}) = \mathrm{rl}_p(Z_f(T), T^{\perp_{\mathcal{Q}}}).$$

We use the symbol a in pairing modulus p^a only in this local definition: in estimates below it is the inherited global derivative-order parameter unless specifically renamed. The Lean definitions of pairings, their cup product, and the orthogonal submodule are, respectively, `Lift.Pairing`, `Pairing.CupBounds`, and `cupOrthogonal` in `Lib/PadicHodge/Herr/GaloisLattices.lean`.

Lemma 9.2 (Exponent-to-length conversion)**.** *If $p^s Z_f(T) \subseteq T^{\perp_{\mathcal{Q}}}$, then*

$$\Delta_p(T; \mathcal{Q}) \leq 2d_F\, s\, \mathrm{rk}_O\, T.$$

Proof. There is a surjection from $Z_f(T)/p^s Z_f(T)$ onto the defect quotient, because its denominator contains $p^s Z_f(T)$. The former has normalized length $s\,\mathrm{rk}_O\, Z_f(T)$. Proposition 9.1 yields $s(1 + d_F)\,\mathrm{rk}_O\, T \leq 2sd_F\,\mathrm{rk}_O\, T$, since $d_F \geq 1$. This is precisely the numerical content of `normLength_defect_le` in `New/Selmer/CrystallineRank.lean`. □

9.2 Why this defect bounds the local discrepancy of congruent cocycles

Take a finite list of global integral lattice representations $u_i : G \to \mathrm{GL}_{n'}(O)$ and a block representation $\rho : G \to \mathrm{GL}_{r'}(O)$. Fix a modulus p^M and write $R = O/(p^M)$. The global image of the reducibility ideal gives, on the block, cocycles $c : G \to R^{r'}$ satisfying not only the cocycle identity, but also all linear relations among the matrices $(u_i(g) - 1)_i$. Let $\mathcal{Z}$ denote the module of cocycles satisfying these relations. Call its local crystalline-property submodule at $v \mid p$ the intersection with the module of cocycles whose restriction to G_v obeys the relations of $Z_f(T_v)$, where $T_v = (O^{r'}, \rho|_{G_v})$. Here "the relations of a module" $Z \subseteq \prod_g O^{r'}$ are finitely supported O-linear relations vanishing on every element of Z; they are imposed modulo p^M. This is the definition in `obeyingCocycles` and `crysRels`, in `New/Selmer/BlocksAndCocycles.lean` and `New/Selmer/DeterminantBound.lean`. This local condition is stronger than continuity and the cocycle identity, and must not be assumed automatically from the construction of a congruence. The entire point of the local estimate is to bound how far $\mathcal{Z}$ is from imposing it.

Here is the algebraic reason these cocycles are amenable to crystalline estimates. Let P be the finite free module of the tuple of matrices, with the action by left multiplication by $(u_i(g))_i$. Put $D(g) = (u_i(g) - 1)_i$ and $A = \sum_g OD(g) \subseteq P$. Then $D(gh) = D(g) + u(g)D(h)$, so A is stable under G. Since O is Noetherian, A is finite. If c obeys the displayed relations then a linear map $q : A \to R^{r'}$ is defined by $q(D(g)) = c(g)$; it is equivariant by the two cocycle identities. By the Artin–Rees lemma, for some j large enough $A \cap p^{Mj}P \subseteq p^M A$. Consequently q extends to a well-defined equivariant map

$$q' : Y = A + p^{Mj}P \longrightarrow R^{r'}, \qquad q'(a + p^{Mj}z) = q(a).$$

Indeed if two such expressions have equal first component then their difference in A belongs to $p^M A$, on which q vanishes. The action on $p^{Mj}P$ is stable. The map satisfies $q'(D(g)) = c(g)$, and Y is a finite free lattice in P. Equivalently the graph of q' projects injectively into P. This is the constructive role of the graph subrepresentation in `obeyingCocycles.exists_graph_subrepresentation`. In particular after restriction to G_v, Y belongs to the class $\mathcal{Q}$ of

lattices inherited from the u_i, whenever those representations have the requisite admissibility, jump sets and annihilating polynomials. The map $g \mapsto D(g)$ is itself a crystalline cocycle whenever the ambient representations are potentially crystalline: over B_{cris} its primitive is the tuple of identity matrices, and multiplication by p^{Mj} puts this primitive in Y after extending scalars. Thus every obeying cocycle has the local form of a reduction of a linear image of a crystalline cocycle on Y. It need not be a reduction of a crystalline cocycle on T_v.

For the duality step, lift q' to an O-linear map $q_1 : Y \to T_v$ (possible by freeness). The reduction is equivariant, so $gq_1 - q_1 = p^M q_2(g)$ with $q_2(g)$ integral. Under the perfect pairing $T_v \otimes T_v^*(1) \to O(1)$ this is exactly (148) on $Y \otimes T_v^*(1)$. Its third class q_2 is crystalline: $p^M q_2$ is a coboundary, and saturation shows that q_2 lies in $Z_f((Y \otimes T_v^*(1))^*(1))$. Hence every reduction in $\mathcal{Z}|_{G_v}$ annihilates the reduction of the orthogonal module $(T_v^*(1))^{\perp_{\mathcal{Q}}}$ under the local Tate pairing. The same conclusion holds for reductions from $Z_f(T_v)$ itself, since $T_v \in \mathcal{Q}$ and the usual perfect pairing is permitted.

Proposition 9.3 (Local congruence loss)**.** Formal-source input (the local Tate-duality index estimate). *Under the preceding hypotheses,*

$$\mathrm{rl}_p\left(\mathcal{Z}, \{c : c|_{G_v} \text{ obeys all relations of } Z_f(T_v)\}\right) \le \Delta_p(T_v^*(1); \mathcal{Q}).$$

For clarity, the content of the index estimate is the following. A finite module $V = T_v/p^M T_v$ has the Cartier dual provided by the reduction of the $\mathbb{Z}_p$-dual Tate lattice. A $\mathbb{Z}_p$-linear functional on O can be chosen to make the pairing perfect: the $\mathbb{Z}_p$-dual of O is a rank-one O-module (the inverse different). The images A and A' of $Z_f(T_v)$ and $Z_f(T_v^*(1))$ have orders whose product is

$$p^{Md(2+d_F)}, \qquad d = \mathrm{rk}_{\mathbb{Z}_p} T_v. \tag{149}$$

Indeed, by (147) their orders are the powers of p with exponents M times their $\mathbb{Z}_p$-ranks. The dual jump relationship (146) pairs each nonnegative jump of T_v with a negative jump of its dual; their ranks therefore sum to $(2 + d_F)d$. The size bound from local Tate duality says that the number of extra classes of the restriction $\mathcal{Z}|_{G_v}$ modulo its intersection with A divides the index of the orthogonal part of $Z_f(T_v^*(1))$. It is applied at the cocycle level, not after omitting the coboundaries; the count (149) includes those terms. Every member of A satisfies the indicated crystalline relations, so passing from the condition of belonging to A to that of satisfying its relations can only decrease the loss. Pulling the resulting estimate back along restriction proves the inequality from the cited index estimate. The exact formal implementation is `IsPotCrys.relIndex_dvd_of_orthogonal` in `Lib/ClassField/LocalDuality/CrystallineDuality.lean` and `relLength_obeyingCocycles_le` in `New/Selmer/RowBalanced.lean`. A separate proof of the former would require developing the local Tate pairing and its index formula; here we have isolated where that input is used. Note especially that the right side is independent of M. This uniformity is what makes the first-order global congruence construction effective.

9.3 The Kummer chart, letter decomposition and height of a frame

We continue the local discussion, writing p for ℓ and reserving E without an additional subscript for the global CM field. In particular, the Eisenstein element in the chart will be denoted by $\mathfrak{e}$ (the Lean notation for it is E). Let $A_{\mathrm{inf}} = W(\mathcal{O}_{\mathbb{C}_p^\flat})$ and B_{dR}^+ be Fontaine's usual period rings, with the map $\vartheta : A_{\mathrm{inf}} \to \mathcal{O}_{\mathbb{C}_p}$ and Frobenius φ on A_{inf}. Write t_{dR} for the uniformizing parameter of B_{dR}^+, so $\mathrm{Fil}^n B_{\mathrm{dR}} = t_{\mathrm{dR}}^n B_{\mathrm{dR}}^+$ for $n \in \mathbb{Z}$. This avoids confusion with the rank difference $t = 5r$ of the global proof.

The local chart fixes an unramified coefficient ring $\Lambda = W(\mathbb{F}_{p^{f_0}})$, a ring $\mathfrak{S} \cong \Lambda[[z]]$ embedded in A_{inf}, and an element

$$\mathfrak{e} = z^{e_0} A + p \quad (A \in \mathfrak{S}^\times)$$

associated with a generator of $\ker \vartheta$. Here z is the Teichmüller lift of the element of the tilt arising from a tower of roots of a local uniformizer. For the completion-group case, $e_0 f_0 = d_F$. The chart also contains an open subgroup Γ' of index prime to p in Γ, preserving the indicated constants and the untilt of z, with the required transitivity on roots. Thus local cup product vanishing may be tested after restriction to Γ', since restriction followed by corestriction multiplies by the prime-to-p index, and the relevant classes are p-power torsion. The unramified chart has $e_0 = 1$, $\mathfrak{e} = z + p$. For the quantitative depth steps the relevant number is not the full discriminant of the global CM field but

$$\delta_0 = \begin{cases} 0, & e_0 = 1, \\ v_p(e_0) + 1, & e_0 > 1. \end{cases} \tag{150}$$

In particular δ_0 is a specified upper bound used in the derivative estimate of the chart and is not the actual different exponent in an arbitrary normalization. The existence of such a chart in the completion-group case, including the subgroup and the compatibility when the extension is unramified, is a *formal-source input*, `exists_kummer_chart` in `Lib/PadicHodge/WachKisin/KummerChartExists.lean`. See also the definition of `Lift.Chart` in `Lib/PadicHodge/WachKisin/HodgeFrame.lean`.

Let $q_0 = p^{f_0} - 1$, and choose a primitive q_0-th root of unity ζ in O (hence the initial enlargement of the coefficient ring). The source indexes the full set of units of $\mathbb{F}_{p^{f_0}}$ by the term *letters*; it is convenient to write $I = \mathbb{F}_{p^{f_0}}^\times$. A letter $i \in I$ indexes the summand of the Witt-vector ambient module on which multiplication by ζ on the lattice factor agrees with multiplication by the Teichmüller root $\omega(i)$ on the Witt-vector factor. The projector to this summand on $\mathcal{V}_T = W(\mathbb{C}_p^\flat) \otimes_{\mathbb{Z}_p} T$ is

$$e_i = \frac{1}{q_0} \sum_{j=0}^{q_0 - 1} \left(\omega(i)^{-1} \zeta \right)^j . \tag{151}$$

Because q_0 is prime to p this operator is integral. Orthogonality and idempotence follow by summation of a geometric progression over distinct roots of unity; $\omega(i) - \omega(j)$ is a unit for $i \neq j$. The sum of the projectors is the identity, and Frobenius sends the summand i to the summand i^p. Denote the permutation $i \mapsto i^p$ by σ_0 and put $i[-a] = \sigma_0^{-a} i$ for $a \geq 0$. The coefficient embeddings of O belonging to letter i are the $\gamma : O \to \mathbb{C}_p$ with $\gamma(\zeta) = \vartheta(\omega(i))$. Define

$$U_i = \{0\} \cup \bigcup_{\gamma(\zeta) = \vartheta(\omega(i))} U(\gamma), \qquad m_T(i, u) = \sum_{\gamma(\zeta) = \vartheta(\omega(i))} j_T(\gamma, u).$$

The zero in U_i permits crystalline extensions by a trivial lattice. The sum of $m_T(i, u)$ over all u and all i is $[K_O : \mathbb{Q}_p] \operatorname{rk}_O T$. In the arguments below it matters that the sum is over letters, whereas f_0 measures the size of the residue field extension: these are different integers. The projector formula, its Frobenius compatibility, and the definitions of the jumps attached to a letter are `letterProj`, `phi_pow_smul_mem_letterPart_iff`, and `jumpMultiplicity` in `Lib/PadicHodge/WachGamma/KisinModuleDefs.lean` and `Lib/PadicHodge/WachGamma/Letters.lean`; the enlarged weight sets are `KummerChart.letterJumps` in `New/Lift/ExponentTheorems.lean`.

We explain the choice of integral lattice used for this calculation. The source works with $W(\mathbb{C}_p^\flat)$ over A_{inf} and an auxiliary element $\tau = [\epsilon^{1/p}] - 1$ of A_{inf}, where ϵ is the compatible

tower of p-power roots of unity. Put $A_T = A_{\mathrm{inf}} \otimes_{\mathbb{Z}_p} T$ inside $\mathcal{V}_T$. An invariant form on T is an A_{inf}-linear functional $\mu : \mathcal{V}_T \to B_{\mathrm{dR}}$ satisfying $\mu(1 \otimes gv) = g\mu(1 \otimes v)$ for $g \in \Gamma'$; write $\mathcal{H}_T$ for the family of such forms. Let M_T be the source's submodule `kisinAinf`, consisting of elements in the τ-saturation of A_T such that

$$\mu(\varphi^a v) \in B_{\mathrm{dR}}^+ \quad (\mu \in \mathcal{H}_T,\ a \geq 1).$$

For a more intrinsic way of specifying the finite-height module $\mathfrak{N}_T$ in M_T, impose the additional value-field condition in the source: $\mu(v)$ must belong to the chart's value field for all invariant forms. This defines `kisinModule`; the value-fields and their Frobenius transforms are as in `Lib/PadicHodge/WachKisin/DeRhamValues.lean`. Let $M_T^{\mathrm{loc}} = M_T[1/\mathfrak{e}_0, 1/\mathfrak{e}_1, \ldots]$ inside $\mathcal{V}_T$, where $\mathfrak{e}_a = \varphi^a \mathfrak{e}$, and let $\mathfrak{N}_T^{\mathrm{loc}}$ be its points satisfying the same value-field condition. We do not identify the latter with an arbitrary scalar extension of an abstract Breuil–Kisin module; the proof uses the above integral model and its retraction. In particular it is important that the period-frame relations are preserved by certain $\mathfrak{S}$-linear retractions on its scalar coefficients.

Proposition 9.4 (Integral frame package)**.** Formal-source input. *Let T be admissible for the chart and let H_T bound the absolute values of its jumps. Then*

$$M_T = A_{\mathrm{inf}}\mathfrak{N}_T, \qquad \tau^{H_T} M_T \subset A_T, \quad \tau^{H_T} A_T \subset M_T, \quad \mathfrak{e}^{H_T}\varphi(\mathfrak{N}_T) \subset \mathfrak{N}_T.$$

Relations among value-field points remain valid after applying an $\mathfrak{S}$-linear map $A_{\mathrm{inf}} \to \mathfrak{S}$ to the coefficients. Moreover let $l, A, \delta \geq 0$ with $3e_0 \leq p^l$, $e_0 A \leq p^l \delta$. Then the value-field points of $\varphi^l(M_T)$, after multiplication by $p^{4\delta}$, are spanned over $\mathfrak{S}$ by value-field points invariant modulo $\mathrm{Fil}^A B_{\mathrm{dR}}$ under all Γ' when tested by all invariant forms. The invariant crystalline forms also test divisibilities in the frame.

These assertions are components of a single theorem, `Lift.KummerChart.JumpLattice.frame_theorem`, in `Lib/PadicHodge/WachKisin/FrameTheorem.lean`, with preparation in `PeriodMatrixFrame.lean`. The underlying results concern Breuil–Kisin lattices and comparison periods, but the last estimate is used with varying l, A, δ; it is not enough to cite finite height or crystalline comparison with unspecified constants. In the formal proof the period matrix is used to compare the natural A_{inf}-lattice with its Frobenius transports, and functionals arising from the invariant period frame bound orders on each Hodge-graded line. The package including the $p^{4\delta}$ estimate is a specific input whose concrete proof would be part of a longer expositional reconstruction. For the current section it will be important to retain this exact exponent. The other uses of the chart involve only the parameters e_0, δ_0 and the Hodge jump sets, rather than the matrix size in the period frame.

9.4 Exponent lattices and the cost of increasing depth

For integer-valued functions $c(a, u)$ on nonnegative integers a and allowed jumps $u \in U_{i[-a]}$, define the exponent lattice $M_T(i, c)$ to be the subset of M_T^{loc} in the letter i such that

$$\mu(\varphi^{-a} v) \in \mathrm{Fil}^{c(a,u)} B_{\mathrm{dR}} \tag{152}$$

for every a and every invariant form μ of type u at $i[-a]$. Here the type condition is $\mu(e_{i[-a]}(1 \otimes T)) \subseteq \mathrm{Fil}^{-u} B_{\mathrm{dR}}$. This is `withExponents` in `New/Lift/ExponentLattices.lean`. We call c the exponents of the point and call c' its depth if, for all $g \in \Gamma'$, $(g-1)v \in M_T(i, c')$. Multiplication by $\mathfrak{e}_a$ raises the exponent at a by one, and Frobenius transports indices $i \mapsto i^p$. In particular the exponents of level l are

$$\mathrm{st}_l(a, u) = \begin{cases} 0, & a < l, \\ -u, & a \geq l. \end{cases} \tag{153}$$

Larger exponents mean a smaller module. A useful operation on exponents is

$$\mathrm{raised}_s(a,u) = -u + s\mathbf{1}_{a=0}.$$

Thus raised_s increases only the exponent at the original Eisenstein element. Depth is distinguished from the exponents of a point itself; the former can be improved without requiring arbitrarily large orders on the nonzero constants to which the points project.

The basic numerical cost attached to a step (a,n) is

$$c_p(a,n) = v_p(n) + \delta_0 + 3a, \tag{154}$$

where we put $v_p(0) = 0$ in this formula only. One point of the construction is that $c_p(a,n)$ is independent of the rank. Its origin is the following assertion underlying the finite-difference operation.

Lemma 9.5 (One-step moving operator)**.** Formal-source input. *For $n \neq 0$, write $\mathfrak{e}_a = \varphi^a\mathfrak{e}$. There exist $g \in \Gamma'$, $\omega \in A_{\mathrm{inf}}^{\times}$ and $\alpha, X \in A_{\mathrm{inf}}$ with*

$$g\mathfrak{e}_a = \omega\mathfrak{e}_a, \quad \mathfrak{e}_a \mid gs - s \ (s \in \mathfrak{S}), \quad z\tau X = \omega^n - 1, \quad \mathfrak{e}_a \mid \alpha X - p^{c_p(a,n)}.$$

If $n = 0$, a sum of p^{δ_0+3a} elements of Γ' has invariant action on the constants modulo $\mathfrak{e}_a$. After the chart retraction these identities supply an additive endomorphism of the localized module raising depth from n to $n+1$, modulo invariant constants, with cost $p^{c_p(a,n)}$ on the quotient of an extension by its kernel.

The identities used here are `Lift.KummerChart.exists_depth_step`, `Lift.KummerChart.exists_depth_step_zero` in `New/Lift/ExponentLattices.lean`, and `exists_reaches_depth_succ` in `New/Lift/HodgeExpansion.lean`. To explain the cancellation for $n \neq 0$, suppose that the leading contribution of an invariant form at the divisor $\mathfrak{e}_a = 0$ is $\mathfrak{e}_a^n h$ modulo higher order, and that h is invariant to order zero. Taking the difference $g-1$ gives $(\omega^n-1)\mathfrak{e}_a^n h$ modulo the next order. The congruence for αX then implies that the combination

$$p^{c_p(a,n)}v - \frac{\alpha}{z\tau}(gv - v)$$

cancels this leading contribution. When $n = 0$, an averaging operator is used instead. The formal identity for the step includes precisely the hypotheses which allow the quotient by $z\tau$ and the application of the coefficient retraction: it is not legitimate to divide by $z\tau$ in an arbitrary integral submodule. The estimate for $c_p(a,n)$ is obtained by lifting a unit under ϑ, the bound for the differentiated Eisenstein equation and the valuation of $\omega^n - 1$. The latter contributes the factor $v_p(n)$ because lifting the exponent in a multiplicative congruence gives a factor dividing the valuation of n. The additional term $3a$ is the allowance for the Frobenius divisor. This explains why subsequent costs consist of sums of valuations of ordinary integers.

Consider an exact sequence of integral lattices $0 \to T \to \widetilde{T} \to C_0 \to 0$ in which C_0 is trivial, described by an O-linear map $\alpha : C_0 \to Z_f(T)$. The underlying O-module of $\widetilde{T}$ is $T \oplus C_0$, with action

$$g(v,c) = (gv + \alpha(c)(g), c).$$

It is crystalline over the chart because each $\alpha(c)$ is a coboundary after extending to B_{cris}. Denote the natural projection by π. In letter i a point over $p^\kappa c$ means $v \in \mathfrak{N}^{\mathrm{loc}}_{\widetilde{T}}$ satisfying $\pi(v) = p^\kappa e_i(1 \otimes c)$ and having prescribed exponents and depth. The relation $\pi(gv - v) = 0$ means that its Galois differences lie in a kernel lattice. Strictness of the kernel for the exponent conditions, namely every point in the kernel of π satisfying the exponent conditions on the extension comes from a point of $M_T(i,c)$ and similarly for localized value-field points, is a *formal-source*

input: `exists_inclKer_eq_of_projTriv_eq_zero` in `New/Lift/LongGap.lean`. Its proof uses extension of invariant forms from the sublattice to the extension while retaining their filtration degree, via `exists_extend_form` of the same file.

The following consequence of the moving operators is the useful form for bounds. Let S be a finite set of steps (a, n) containing those required to raise depth from ρ to ρ': for each a, u it contains every pair (a, n) with $\rho(a, u) \le n < \rho'(a, u)$. Starting with a point over $p^{\kappa}c$ of depth ρ, one obtains a point over $p^{\kappa+\sum_S c_p(a,n)}c$ of depth ρ'. Its exponents improve to $\rho'(a, u)$ wherever its old exponents were positive, and to at least $\min(0, \rho'(a, u))$ on the other lines. In the precise statement the integers may be raised further by paying more powers of p. This is *formal-source input* `pointsOver_nonempty_of_costs`, `exists_reaches_deep` and `mem_withExponents_of_deep` in `New/Lift/LongGap.lean` and `New/Lift/HodgeExpansion.lean`. The induction is on the largest step index to be increased and its depth. Compose the operators of Lemma 9.5; when operator D has cost p^e, the difference $Dv - p^e v$ belongs to the kernel of π, so the next operator preserves the accumulated scalar. Thus the cost of a composition is the sum of the integer exponents. The crucial difference from a dimension argument is that the same operator raises the specified depth on every line of the pertinent jump, so multiplicities do not multiply this cost. The exponent improvement statement uses that a de Rham value whose Galois differences lie in a deep filtration can be made invariant modulo that filtration by subtracting a constant from the local field. Positivity of the prior exponent forces that constant to be zero; without positivity one retains at least the zero-filtration threshold. These are the precise properties needed in the later cup-product calculation.

We record the elementary valuation estimate controlling these operators. For integers $n \ge 0$ and c, set $I(n, c) = [-n, n] \cup [c-n, c+n]$ (integer intervals). If $n \le pM$ and $|c|+n \le H$ with $H \ge 1$, then for $d \ge 0$,

$$\sum_{j \in I(n,c)} (v_p(j) + d) \le (4M + 2)\lfloor \log_p H \rfloor + (4n + 2)d. \tag{155}$$

Indeed at most $4n + 2$ integers occur and, in each interval of $2n + 1$ integers, at most $2\lfloor n/p \rfloor + 2$ can be multiples of p. More sharply, an interval of length $2n$ contains at most $\lfloor 2n/p \rfloor + 1 \le 2M + 1$ multiples if $n \le pM$; hence the two intervals together contain at most $4M + 2$. For any nonzero j of absolute value at most H, $v_p(j) \le \lfloor \log_p H \rfloor$, and zero contributes no valuation by convention. This proves the bound. It is `sum_twoWindows_le` in `New/Lift/CokernelKillers.lean`. For steps with $a = 0$, if p divides no nonzero integer within distance n of $0, \pm c$ then the valuation contribution vanishes; in that case all the steps in these intervals cost only δ_0 each. With the rough count $8n+4$ of their integers the total cost is at most $(8n + 4)\delta_0$. The formal-source derived assertion `pointsOver_nonempty_of_noMultipleNear` therefore improves a point at the original divisor at precisely this cost. In an unramified chart the cost is zero. This free improvement is the first of several places where the proof needs a gap in the congruence classes of large weights, not just a gap in their ordinary archimedean sizes.

9.5 Cokernel estimates: divisibility by p and by the Eisenstein element

For an extension side (T, C_0, α) as above let $J_l(i; s)$ be the ideal of A_{inf} defined as the sum of (s) and the ideal of those multipliers which send the letter-i part of $A_{\mathrm{inf}} \otimes C_0$ into the image under π of the extension's exponent lattice with exponents st_l. Thus, for a scalar $u \in A_{\mathrm{inf}}$, the condition $u \in J_l(i; s)$ means that one can choose $v \in A_{\mathrm{inf}}$ such that for every $c \in C_0$ an exponent-st_l point lifts $(u - sv)e_i(1 \otimes c)$. We say that $p^{\kappa}\mathfrak{e}^s$ kills the cokernel at i if

$$p^{\kappa}\mathfrak{e}^s \in J_0(i; 0).$$

These are `liftIdeal` and `KillsCokernel` in `New/Lift/CokernelKillers.lean`. This definition is a statement about actual multipliers, not just equality after inverting p. Coefficient retraction allows the corresponding lifts to be taken in the localized value-field module. In particular if $p^\kappa \mathfrak{e}^s$ kills, after dividing by $\mathfrak{e}^s$ one finds a point over $p^\kappa c$ having both exponents and depth raised$_{-s}$. This last assertion is *formal-source input* `pointsOver_nonempty_of_killsCokernel` in `New/Lift/LongGap.lean`.

Proposition 9.6 (Cokernel calculus)**.** Formal-source input. *For the ideals just defined, the following are valid.*

(i) *The level and letter are transported by Frobenius: in particular $p^c z^{pa} \in J_1(i^p; r_0)$ if and only if $p^c z^a \in J_0(i; r_0)$, for a fixed modulus r_0. If all allowed jumps are $\le n$, multiplication by $\mathfrak{e}^n$ sends $J_0(i; r_0)$ into $J_1(i; r_0)$.*

(ii) *If $p^c z^v \in J_1(i^p; r_0)$, $pa \le v$ and $c + v_p(v!) \le \kappa + v_p((pa)!)$, then $p^\kappa z^a \in J_0(i; r_0)$.*

(iii) *If $p^c z^a \mathfrak{e}^n \in J_1(i; r_0)$, then $p^{c+\delta_0 n + v_p(n!)} z^{a+n} \in J_1(i; r_0)$.*

(iv) *If all jumps on the side are $\le n$, for each $\kappa \ge 0$ there exists $a \ge 0$ with $(p-1)a \le e_0(n+\kappa)$ and $z^a \in J_0(i; p^\kappa)$ for all i.*

(v) *In the unramified chart and for $p \ge 5$, if all jumps are $\le (p-3)(\kappa + s)$ then $p^\kappa \mathfrak{e}^s$ kills the cokernel at every letter.*

All assertions occur in `New/Lift/CokernelKillers.lean`, respectively `p_pow_mul_z_pow_mem_liftIdeal_iff` and `mul_E_pow_mem_liftIdeal_one`, `mem_liftIdeal_zero_of_mem_one`, `mem_liftIdeal_of_mul_E_pow`, `exists_z_pow_mem_liftIdeal`, and `killsCokernel_of_isUnramified`.

We recall how the factorials arise. The finite-difference calculation from Lemma 9.5, at level zero or one, acts on $J_l(i; r_0)$ by the following rule. If $f \in \mathfrak{S} \cap J_1(i; r_0)$ and there is an identity

$$z\tau\mathfrak{e}(s + \beta f) = \alpha(gf - f), \qquad g \in \Gamma',$$

with the appropriate integral scalar coefficients, then $s \in J_1(i; r_0)$. It follows by taking the difference of integral lifts in the defining multiplier ideal, using a retraction to preserve the membership conditions. This rule is *formal-source input* `mem_liftIdeal_of_derivation`. Applied to $f = p^c z^v$, the rule lowers v by one up to $p^{v_p(v)}$ in the coefficient. Inducting from v down to pa gives the additive cost $v_p(v!) - v_p((pa)!)$; transporting the remaining power back by Frobenius proves (ii). Applied to $f = p^c z^a \mathfrak{e}^n$, the rule reduces n by one and increases the z-exponent, at step cost $\delta_0 + v_p(n)$. Their sum is $\delta_0 n + v_p(n!)$, giving (iii). These summations are independent of any rank. In the unramified case the polynomial identity $\mathfrak{e} = z + p$ permits one to combine the multipliers of (ii) with a binomial expansion of $\mathfrak{e}^s$. The resulting containment modulo $p^{\kappa+s+1}$ is upgraded to an actual containment by the elementary Nakayama argument: if $p^c \in J_0(i; 0) + (p^{c+1})$ then $p^c(1 - pq) \in J_0(i; 0)$ for some $q \in A_{\mathrm{inf}}$, and $1 - pq$ is invertible. Thus $p^c \in J_0(i; 0)$ and adding the higher modulus does not enlarge $J_0(i; 0)$. The quantitative minimum on a in (iv) and binomial comparison in (v) are the specific remaining cokernel estimates; they cannot be replaced by an unquantified assertion that some power of p annihilates the cokernel.

9.6 Cup-product vanishing from the three sides

Each pairing in (148) introduces three crystalline extension sides: the side Y carrying c, the side T carrying c', and the side $(Y \otimes T)^*(1)$ carrying the third class q. On the last side the

quotient lattice is O with extension determined by q, and the allowed jumps, including zero, are the numbers $-u-v-1$ with u and v running over the jumps on the first two sides. The integral equivariant pairing of this third extension with $Y \otimes T$ is

$$((\xi, s), u \otimes v) \longmapsto (p^M \xi + s\psi)(u \otimes v),$$

where p^M is the pairing modulus. Its equivariance follows directly from (148), since the extra term $p^M q(g)$ is exactly the change of ψ. On reduction it agrees with the imperfectly dual pairing. This is why controlling the first two extensions is not enough: the period computation must also supply a lift of a constant in the third extension.

Over the tilt consider $\mathcal{M} = W(\mathbb{C}_p^\flat) \otimes \mathbb{Z}/p^M(1)$ and let $\mathcal{M}_+$ be the image of $W(\mathfrak{m}_{\mathbb{C}_p^\flat})$. Call a pairing of Witt-vector components safe with cost κ if p^κ times its value lies in $\mathcal{M}_+$. The submodule $\mathcal{M}_+$ is stable under Frobenius and Γ. Different letters are orthogonal: for example if two vectors have labels i, j, then multiplying their value by $\omega(i)$ or $\omega(j)$ gives the same result, since the pairing is O-bilinear before applying the functional and reducing. For $i \neq j$ their difference is a unit. This is also true of the sum of their pairings, as required by `isSafe_sum`.

Here is the algebra of the safe-pair criterion. We use the elementary Artin–Schreier–Witt property of $\mathcal{M}$ with its small part: $\varphi - 1$ is surjective; if $(\varphi - 1)z \in \mathcal{M}_+$ then z is congruent modulo $\mathcal{M}_+$ to a unique element of the constant module $\mathbb{Z}/p^M(1)$; and the intersection of the constants with $\mathcal{M}_+$ is zero. This property, including the convergence for Galois differences needed to get continuous cochains, is a *formal-source input*, `TateTwist.smallPart_spec` in `Lib/PadicHodge/WachGamma/KisinModuleDefs.lean`.

Suppose x, w in the first extension lie over the same constant $p^{e_1}c_0$, and y in the second lies over $p^{e_2}c_0'$. Write X, W, Y for their respective first components in the Witt-vector tensor modules of the kernels. Write a_g, b_g for the crystalline cocycles on these kernels, scaled by p^{e_1} and p^{e_2}. Thus the first component of $(g-1)x$ equals $D_g = gX - X + a_g$, that of $(g-1)w$ equals $D_g^W = gW - W + a_g$, and that of $(g-1)y$ equals $T_g = gY - Y + b_g$. Put $C = \varphi Y - Y$. Denote the pairing multiplied by p^{e_3} by P. We assume the following four expressions are zero in $\mathcal{M}/\mathcal{M}_+$, for all $g, h \in \Gamma'$:

$$P(D_g, gT_h), \quad P(W - X, T_g), \quad P(D_g^W, gC), \quad P(\varphi X - W, g\varphi Y - \varphi Y + b_g). \tag{156}$$

These are precisely the four safe-pair hypotheses in `Pairing.ScaledCupBounds.of_isSafe`, `New/Lift/CupProductBound.lean`. The last second argument is the first component of $(g-1)\varphi y$. Since the scaled cocycle values are constants for Frobenius it is also φT_g. The goal is to see that $P(a_g, gb_h)$ is a coboundary.

First define $k_0(g) = P(a_g, gY) - P(X, T_g)$. The graded Leibniz rule for the group-cohomology differential, which can equally be checked by expansion in degree one, gives

$$\mathrm{d}k_0(g, h) = P(a_g, gb_h) - P(D_g, gT_h).$$

This equality can be read in the quotient because P is equivariant by construction. By the first safety condition the last term is zero. It remains to choose the cochain with values in constants, rather than just the large Witt-vector quotient. Choose Z with $(\varphi-1)Z = P(W, C)$ by the Artin–Schreier–Witt property and set $k(g) = k_0(g) + gZ - Z$. In the quotient, direct expansion gives

$$\begin{aligned}(\varphi - 1)k(g) &= P(a_g, gC) - P(\varphi X, gC - C) \\ &\quad - P(\varphi X - X, T_g) + P(gW, gC) - P(W, C) \\ &= P(D_g^W, gC) - P(W - X, T_g) - P(\varphi X - W, \varphi T_g).\end{aligned}$$

This is zero by the remaining three conditions. Hence each $k(g)$ is congruent to a unique constant $\kappa(g)$. The differential of k equals that of k_0, and $P(a_g, gb_h)$ is a constant. Injectivity of the constants into the quotient implies that $\mathrm{d}\kappa(g,h) = P(a_g, gb_h)$ as an identity of constants. Continuity of κ is part of the preceding formal-source small-part estimate: the lift formula involves finite congruence data, and sufficiently small changes under the Galois action contribute to the small part, so the unique constant representative is locally constant. This proves vanishing after restriction to the chart group Γ', and then in the cohomology group on Γ by corestriction, since its index is prime to p. We may further increase the total exponent $e_1 + e_2 + e_3$. This gives the claimed scaled cup product, and is the mechanism behind the formal theorem `Pairing.ScaledCupBounds.of_isSafe`. Notice that the third condition involving w is not redundant: it provides precisely the correction to make the cobounding cochain Frobenius-invariant modulo the small part.

The remaining challenge is to produce safe points with sufficiently small exponents e_1, e_2, e_3. The main safety test is a strict filtration inequality. Suppose v, w in the two kernels lie in exponent lattices of exponents $o(a,u), o'(a,v)$. Suppose that in the third extension one has a point over p^κ at the predecessor letter $i[-l]$, with exponents $o_2(a,u'')$, antitone as a function of the jump. The required numerical inequality is

$$o(a,u) + o'(a,v) + \mathbf{1}_{a\geq l} o_2(a-l, -u-v-1) > 0 \tag{157}$$

for each a and all the jumps u, v at $i[-a]$. Under these hypotheses multiplying either input by $z\tau$ makes its pairing safe with cost κ. This is *formal-source input* `SidePairing.isSafe_of_exponents` in `New/Lift/LongGap.lean`. It uses the equivariant third pairing above and the rule that multiplication of filtered values adds exponents. The target line $O(1)$ has jump -1; applying the frame comparison for that line produces the factor τ, and the positive valuation of z puts the result in $W(\mathfrak{m}_{\mathbb{C}_p^\flat})$. The strict inequality is indispensable, as safe is a condition of vanishing modulo the maximal ideal of the tilt, not merely integrality. A Galois difference of a point above a constant projects to zero and is divisible by the auxiliary $z\tau$ in the required localized kernel; this is how the depth inequalities for x, w, y imply (156).

One uses cuts on Frobenius cycles of letters. Let $I_0 \subset I$ be a subset meeting each cycle, $l(i)$ the distance to the first strictly preceding cut, and choose one leading side at each cut. The point at every other letter is transported from a preceding cut by Frobenius; at a cut one may choose directly constructed points on the leading side. The complementary choice gives two systems of points such that the Frobenius of one system is the other. More concretely the integer shifts are $d_P(i) = 0$ if i is a cut at which the side leads, and $d_P(i) = l(i)$ otherwise. For the complementary systems the identity

$$d_{\mathrm{lag}}(i) = d_{\mathrm{lead}}(i[-1]) + 1$$

uses minimality of the preceding cut. Consequently summing $\varphi^{d_P(i)}$ of the chosen points at $i[-d_P(i)]$ yields the points in the previous safe-pair criterion. Orthogonality of different letter summands reduces all four tests to the separate inequalities (157), with appropriate depth and minimum exponents of differences of points at a cut. The precise assertion with the additional depth inequality for the leading side is *formal-source input* `SidePairing.scaledCupBounds_of_points_at_cuts` in `New/Lift/LongGap.lean`; it concludes that an exponent $e_1 + e_2 + \kappa$ is sufficient. The optimization of the cut selection matters, because a Frobenius walk of length l introduces new filtration inequalities at all preceding divisors $\mathfrak{c}_a$.

9.7 The window estimates: the three ways to obtain safe points

For a sign $\epsilon \in \{1, -1\}$ put

$$I(\eta, \epsilon, b, W_0) = [-\eta, \eta] \ \cup \ \epsilon[b, b + W_0],$$

and write $\mathrm{NM}(p,n,b)$ for the condition that no nonzero integer within distance n of $0,b,-b$ is divisible by p. These two definitions are `IsInJumpWindow` and `NoMultipleNear` in `New/Lift/CokernelKillers.lean`. The estimates use opposite signs for the far windows of paired sides. The third extension has many possible jump sums, but for two opposite windows the sum of any pair of large jumps is near zero if both are far, and near a far window if precisely one is far. This is the basic reason that the window structure is stable enough for (157) to be useful.

Sparse thick letters. First consider opposite windows $I(\eta,\mp\epsilon_i,\beta_i,V)$ at letter i, where $\beta_i=b$ for thick letters and $\beta_i=G$ for thin ones, and $G+V\le\eta$. Let $W_0\ge V+4\eta$, and suppose that either all the signs are constant or the chart is unramified and $p\ge 7$. For an integer $J\ge 0$ assume that there cannot be three thick letters $i, i^{p^s}, i^{p^{s+t}}$ with $1\le s,t\le J$. The following full set of size hypotheses is sufficient:

$$8e_0\le p^{J+1},\qquad 2e_0b\le p^{J+1}X,$$
$$11(p^J+1)W_0\le b,\qquad 4(J+1+\delta_0+\lfloor\log_p(2b)\rfloor)W_0\le X,\qquad \eta\ge 1.$$

Under them every cup product of crystalline cocycles of the paired sides is killed, in the sense of (156), by $p^{10^4X(J+1)}$. This is *formal-source input* `SidePairing.scaledCupBounds_sparse_thick` in `New/Lift/SparseThickLetters.lean`. One must use these hypotheses on both sides, not just on the distinguished block.

The combinatorial operation in its proof is concrete. Define a cut at i if the next J letters $i^p,\dots,i^{p^J}$ are thin. On a Frobenius orbit every letter has a preceding cut at distance at most $2J+1$. Indeed, for $J=0$ all letters are cuts. For $J>0$, suppose every one of the preceding $2J+1$ positions fails to be a cut. Failure at distance $J+1$ yields a thick letter at distance a with $1\le a\le J$. Failure at distance $a+J+1\le 2J+1$ yields a thick one at distance $a+b$, $1\le b\le J$. Failure at distance a yields a third thick letter at forward distance c, $1\le c\le J$, from the first. This is a forbidden triple with gaps b,c. The formal implementation of this selection is the first part of `scaledCupBounds_sparse_thick`; the cut definition also gives the short thin run used in the bound on $J_0(i;s)$ of each extension. Along the run, multiplying by a suitable power of $\mathfrak{e}$ and transporting under Frobenius changes the required exponent of z by a factor p per step. Proposition 9.6(ii),(iii) gives the extra factorial and different costs at each stage. The number X in the hypotheses bounds those costs. The starting powers of z are obtained from Proposition 9.6(iv). At cuts the resulting points can improve their depth in the two windows, paying the bound (155) and summing over at most a bounded multiple of $J+1$ steps. With budget κ satisfying $12(J+1)X<\kappa$, the source obtains the points over $p^{25\kappa}$ on both sides and a third point over p^{κ}. Thus 51κ pays for the safe-pair criterion. Taking κ of order $(J+1)X$ and allowing ample slack gives the displayed exponent $10^4X(J+1)$. The general cut lemma specifying both the permissible sign pattern and the gap of the large weight is `SidePairing.scaledCupBounds_of_cut_windows`; it is a further formal-source input. This is one of the technically new results needed for the proof, not a standard property of p-adic representations cited without the quantitative bound.

Two sizes over an unramified chart. The second general regime assumes $p\ge 7$ and $e_0=1$, $\mathfrak{e}=z+p$. The windows are again $I(\eta,\pm\epsilon_i,\beta_i,V)$ with $\beta_i=b$ or G, and $W_0\ge V+4\eta$, $\eta\ge 1$. If

$$5W_0\le G\le b,\qquad b<Kp^2,\quad 2G<Kp,\quad W_0\le K,$$

then the exponent

$$10^4K(\lfloor\log_p(2b)\rfloor+1) \tag{158}$$

annihilates all imperfect cup products. This is *formal-source input* `SidePairing.scaledCupBounds_two_sizes` in `New/Lift/SparseThickLetters.lean`. It is used when p is large and a valuation budget is available, regardless of the sparsity of far-apart letters. Here the cut choice uses one Frobenius step and at i an additional order

$$s_i = \begin{cases} \max(0, \lfloor b/p \rfloor - 7K), & i^p \text{ thick}, \\ 0, & \text{otherwise}. \end{cases}$$

It is a cut if $s_i < \beta_i$. Some cuts require $\mathfrak{e}^{s_i}$ in the cokernel bound, rather than only p^κ; Proposition 9.6(v) is the point that makes these useful. When traversing the single large gap the necessary deep point is constructed by extracting a coefficient in an Eisenstein expansion, avoiding the large cost of paying for every intermediate jump. This explains the role of the long-gap lemma below. For the precise verification of the cup-product depth inequalities in this regime one uses the same `scaledCupBounds_of_cut_windows` with $J = 1$ and $n = K$.

The long-gap operation. The auxiliary statement used for these cuts is worth formulating. In an unramified chart suppose $l > 0$, $3e_0 \le p^l$, p^κ kills the cokernel at $i[-l]$, $\mathfrak{e}^s$ kills at i, all allowed jumps at $i[-j]$ are $\le G_1$ for $0 \le j < l$, and put

$$N = s + G_1 \sum_{j=0}^{l-1} p^j.$$

If $N < A$ and $e_0 A \le p^l \delta$, there is a point in the extension over $p^{\kappa+4\delta}c$ with level-l standard exponents modified at $a = 0$ by adding $\min(A, -u - N)$ on the line of jump u, and with depth modified at $a = 0$ by adding A independent of u. This is *formal-source input* `PairingSide.pointsOver_nonempty_of_long_gap` in `New/Lift/LongGap.lean`. A brief account of the mechanism suffices to locate the nontrivial ingredients: by starting with a point over the pure Eisenstein annihilator, one multiplies by a monic polynomial in $\mathfrak{e}$ of degree N to arrive in a Frobenius-transported module. In this step each Frobenius predecessor contributes the corresponding power of p to the degree; explicitly the polynomials for the unramified chart arise from $\varphi^j(z + p) = (\mathfrak{e} - p)^{p^j} + p$. Applying the $p^{4\delta}$ estimate of Proposition 9.4 expresses the resulting element as a linear combination of nearly invariant points. Expand the coefficients of that combination in powers of $\mathfrak{e}$ using $\mathfrak{S}/\mathfrak{e} \simeq \Lambda$. Comparing the coefficient of degree N gives a new point, now with exponents controlled by N rather than by the full far weight. The strict inequality $A > N$ ensures that the original leading coefficient survives the comparison. The precise invariant coefficient extraction is expressed by `JumpLattice.coeff_mem_constantsModE` in that file. Thus the exponent 4δ in the frame theorem and the geometric degree N are both visible in the cost of the long-gap operation.

Few low jumps. In the last regime it would be wasteful to multiply every class by the exponent needed for a small exceptional part. The module of crystalline cocycles of the distinguished side is used as the trivial quotient C_0, with its identity cocycle map. In an unramified chart suppose $p \ge 5$, $\eta \ge 1$, and all paired jumps lie at each letter in the opposite windows $I(\eta, \pm\epsilon_i, \beta_i, W_0)$. Let A_0 be a nonnegative integer such that for every letter

$$\beta_i + W_0 + \eta \le (p-3)A_0, \qquad 6(A_0 + \eta + W_0) < \beta_i, \qquad \mathrm{NM}(p, 6(A_0 + \eta + W_0), \beta_i). \quad (159)$$

Then the following estimate holds.

Theorem 9.7 (Length estimate for few low jumps)**.** Formal-source input. *Let T have the indicated second jump window and let $\mathcal{Q}$ be the crystalline lattices with the opposite first jump*

sets. Suppose that at every coefficient embedding the sum of the multiplicities of jumps between $-\eta$ *and* η *is at most* μ*. Then*

$$\Delta_p(T;\mathcal{Q}) \le 11 f_0 \left(\mu(A_0+1) + (W_0+1)\,\mathrm{rk}_O\, T\right).$$

Moreover $\Delta_p(T;\mathcal{Q}) = 0$ *if neither of the original jump sets meets* $[-A_0, 0)$*, or if at every embedding the second jump set is contained in* $[-\eta, 0]$*.*

This is `KummerChart.few_low_jumps` in `New/Lift/ExponentTheorems.lean`. Its bound involves the width of the far window and the number of low jumps separately, rather than the rank times A_0; this refinement is essential in summing the loss.

Here is the internal length argument behind it. For each letter, consider the constants which lift to points with exponent raised$_0$ and prescribed depth. To increase depth on one graded line of multiplicity $m_T(i,u)$ at the Frobenius divisor $\mathfrak{e}_a$, the index can be controlled by counting relations modulo p among its leading coefficients. For an unramified chart, if the depth conditions at $\mathfrak{e}$ can be advanced to the next Frobenius threshold without a nonzero multiple of p intervening, then the leading coefficients of $f_0 m_T(i,u)+1$ points satisfy a nontrivial integral relation with one coefficient equal to 1 modulo the deeper submodule. The leading-term calculation and normalization of a relation by its coefficient of least valuation are made precise by `exists_comb_of_jumpMultiplicity_le` in `New/Lift/LongGap.lean`. Indeed the leading terms on a Hodge line of multiplicity m may be represented in at most m copies of the finite unramified coefficient module of rank f_0 over $\mathbb{Z}_p$; the rank count supplies a relation. The theorem includes the important compatibility of the retraction with this leading-term comparison, so it is a statement about the actual integral exponent modules, not just dimensions of their rational Hodge pieces.

For finite free modules the passage from this relation property to a length bound is elementary. If $p^d L \subset L'$ and every $q+1$ elements have an integral combination, one coefficient a unit, belonging to L', then the elementary-divisor decomposition of L/L' shows that at most q cyclic elementary divisors can be nonzero, and all their exponents are at most d. The length is therefore at most dq over $\mathbb{Z}_p$. Applying this estimate to each depth step and using subadditivity of relative length bounds the cost of changing the exponents from $c(a,u)$ to $c'(a,u)$ by a sum of step costs weighted by the corresponding multiplicity. More concretely, let n_i and r_i bound f_0 times the low and the total jump counts at i. The prescribed module of deep constants at letter i has colength at most

$$(2A_0+1)n_i + 3(W_0+1)r_{i[-1]} + 6A_0 n_{i[-1]}. \tag{160}$$

The precise modules and sign-dependent depths are `deepConstants`, `relLength_deepConstants_le`, and `exists_relLength_deepConstants_le` in `New/Lift/FewLowJumps.lean`. If the sign of the second side at i is positive its depth increases by $2A_0+\eta$ at the single Eisenstein element. For the negative sign the prescribed depth is $\beta_i + W_0 + 2A_0$ at $a=0$ and $W_0 + 1 - u + 2A_0 \mathbf{1}_{|u|\le\eta}$ at $a=1$, with $-u$ elsewhere; this is the depth underlying (160). The verification of (160) from the individual steps is a formal-source input; the source further checks that these depths give safety through the cut lemma. Taking the intersection over all letters can only cost the sum of their colengths (inject the quotient of the intersection into the direct sum of quotients). Since $i \mapsto i[-1]$ is a permutation, summing (160) gives a bound by

$$11 \sum_i \left((A_0+1)n_i + (W_0+1)r_i\right).$$

For $n_i = f_0 \sum_{|u|\le\eta} m_T(i,u)$ and $r_i = f_0 \sum_u m_T(i,u)$ this is at most $11 f_0 [K_O:\mathbb{Q}_p](\mu(A_0+1) + (W_0+1)\,\mathrm{rk}_O\, T)$ over $\mathbb{Z}_p$. Dividing by $[K_O:\mathbb{Q}_p]$ as in (145) proves the displayed normalized

length bound from the integral depth statements. The zero-loss part of Theorem 9.7 does not follow by making this right side zero: when the low negative jumps are absent, pure Eisenstein annihilators and zero-cost depth steps give safe points at exponent zero. The second alternative is treated separately by the no-multiple condition and one Frobenius transport of the low-jump side. These are the two cases in `SidePairing.scaledCupBounds_of_killsCokernel`, `New/Lift/LongGap.lean`.

9.8 The tame type bound, uniform in the ramification degree

At the selected primes g_0, g_1 the local representations may only be potentially crystalline. Here the ramified inertial type has a special form. Let r_0, N_0' be bounds and suppose that for every inertial element s the polynomial class $\mathcal{K}(s,\cdot)$ contains X^M-1, where $M=\operatorname{lcm}_{1\le j\le r_0}(p^j-1)$, and contains a monic polynomial in $\mathbb{Z}_p[X]$ of degree at most N_0'. This is precisely `IsTameType` in `New/Lift/ExponentTheorems.lean`. The dual condition is formed by using polynomials at s^{-1}. Suppose now that the local field is $\mathbb{Q}_p$, the paired jump sets lie in opposite windows $I(\eta,\pm\epsilon,b,W_0)$, and

$$\kappa_0 = r_0^2(N_0'+1)^2, \qquad b+6(W_0+\eta)+\kappa_0+2<p.$$

Then multiplication by

$$p^{3\kappa_0+150(\eta+W_0+1)} \tag{161}$$

sends $Z_f(T)$ into the orthogonal part for all lattices in the opposite polynomial class. This is *formal-source input* `tame_types_exponent`, in the same file.

For the uniformity of (161), it is especially useful to recall its proof mechanism. All the roots of X^M-1 are prime-to-p roots of unity; one may use a tame chart of ramification index M, which can be extremely large. But $M\le\prod_{j\le r_0}p^j\le p^{r_0^2}$, and the monic small-degree annihilator allows one to avoid paying for all that ramification. A generator s of the tame quotient acts on the chart coordinate as $z\mapsto\xi z$ for a Teichmüller root of unity of order M. If Q is a monic annihilator on the kernel lattice, then $Q(X)(X-1)$ kills the extension of a trivial lattice by that kernel. Indeed $s-1$ maps the extension into the kernel on crystalline periods, and the polynomial Q kills that kernel. The corresponding operator on framed sections is the polynomial of the tame generator; its image is divisible by z in the integral frame. This last assertion is the specific formal-source input `TameChart.exists_z_smul_eq_sum_tameGen` in `New/Lift/SparseThickLetters.lean`. It is proved by applying the relation on the period matrix and using the divisibility test in the frame theorem. To lower a power z^j in the multiplier ideal one now uses this polynomial evaluated at ξ^j: unless its reduction vanishes, it is a unit and no power of p is needed. Where it vanishes the finite-difference step pays at most $\log_p M\le r_0^2$. The reductions of ξ^j for $0\le j<M$ are distinct, hence there are at most $\deg Q+1$ exceptional steps. Thus the cost at a side is bounded by $r_0^2(\deg Q+1)$, independently of M itself. The third extension uses a polynomial of degree bounded by the product of the two degrees (plus the contribution of the trivial extension). The chosen κ_0 is sufficient for all three sides. This exact bound for the cokernel is `TameChart.killsCokernel_of_isKilledBy`. Together with the no-multiple condition, guaranteed by the size hypothesis $b+6(W_0+\eta)+\kappa_0+2<p$, the depth and safe-pair calculation gives (161); the factor 150 is a generous allowance for the three sides. The rank-independent cost of the earlier depth operators and the degree rather than ramification bound for the tame operators are distinct features of the proof.

In the selected-prime application, the root condition itself gives the small monic annihilator. A root of order dividing p^j-1 lies in an unramified extension of degree dividing j, hence has monic minimal polynomial over $\mathbb{Z}_p$ of degree $\le j$. The product of at most N such polynomials is a monic polynomial of degree $\le Nr_0$ vanishing on the indicated roots. This

explains `isTameType_of_roots` in `New/Selmer/RowBalanced.lean`, which is applied to the polynomial class of the finitely many columns. It contrasts with a direct estimate based on the order of the whole finite group at g_0, g_1.

9.9 Case analysis for the primes and the explicit local loss

The form in which the three crystalline-window estimates are combined is the following. We give the constants, since an asymptotic assertion without them would not specify the interface to the global proof. Temporarily write b_* and G_* for the two window sizes, h_* for the half-gap parameter, and let $Q > 0$ be the prime cutoff. Put

$$\lambda_p(h_*, b_*) = v_p\left((2b_*)!\left(\prod_{\max(b_*-30h_*^3,0)<j\le b_*+30h_*^3} j\right)^{h_*^3}\right). \tag{162}$$

Assume the size conditions

$$99(Q^2 + (W_0+1)Q + Q + W_0 + m) \le h_*, \qquad 99h_*^3 \le G_*,$$
$$99(G_*h_*^3 + G_*) \le b_*, \quad 2\log(4b_*) < m, \quad \delta < \eta, \quad 99\big((\eta + W_0)Q + \eta + h_*\big) \le h_*^3,$$

with $e_0 \le h_*$. Let the opposing jump windows be $I(\eta, \mp\epsilon, b_* + \delta, W_0)$ at thick embeddings and $I(\eta, \mp\epsilon, G_* + \delta, W_0)$ otherwise. Suppose that either all negative jumps on both sides satisfy $2u \le -h_*$ and the multiplicity of jumps of the second lattice T in $[-\eta, \eta]$ is at most μ per embedding, or the second jump sets consist entirely of jumps in $[-\eta, 0]$. At letters above a given embedding the sign and the thick indicator must be compatible whenever the embedding is thick or $Q < p$. If $p > Q$ the chart must be unramified with $p \ge 7$; at $p \le Q$ one may instead use constant signs and the sparsity of thick letters (no three successive thick letters with the two positive gaps at most $\lfloor \log_p(h_*^3) \rfloor$). This is the set of hypotheses used by `crystalline_defect_le` in `New/Selmer/CrystallineBound.lean`, where T is crystalline.

Under these hypotheses the formal-source local estimate is

$$\Delta_p(T; \mathcal{Q}) \le d_F\, 10^5 \left(\mathrm{rk}_O\, T \left\lfloor \frac{2\lambda_p(h_*, b_*)}{Q} \right\rfloor m + \mu\lambda_p(h_*, b_*)\right). \tag{163}$$

The same statement uses an arbitrary upper bound y for the parenthesized expression including 10^5 in its source formulation. Here are the quantitative subcases in the proof. Write $v = \lambda_p(h_*, b_*)$ and $K = \lfloor 2v/Q \rfloor$. By the factorial term and Legendre's formula,

$$\lfloor 2b_*/p \rfloor \le v, \qquad 2b_* < p(v+1), \qquad KQ \le 2v < (K+1)Q.$$

Also $\lfloor \log_p(4b_*) \rfloor < m$, since $p \ge 2$ and $2\log(4b_*) < m$.

- For $p \le Q$ the resource inequality $2b_* < p(v+1)$ forces a large bound K; after enlarging the small window to absorb the thin weight, use the sparse-thick exponent with $J = \lfloor \log_p(h_*^3) \rfloor$. The source compares $2(b_* + \delta)$ with $h_*^2 K$ and takes the depth budget compatible with K. The rank bound of Lemma 9.2 then suffices, at cost bounded by the first term of (163).
- If $p > Q$ and $v \ge h_*^3$, use the two-sizes exponent (158). The floor inequality, $p > Q$, and the size assumptions ensure the requisite bounds on $b_* + \delta, G_* + \delta$ in terms of K. The logarithmic factor is at most m, so by Lemma 9.2 the normalized loss is at most $2 \cdot 10^4 d_F\, \mathrm{rk}_O\, TKm$, absorbed by the first term.

- If $p > Q$ and $v < h_*^3$, the product factor in (162) has no multiple of p in its interval: a single such integer would contribute h_*^3 to v. The factorial term also implies that p is large compared with G_*, since $2b_* < p(h_*^3+1)$. These facts yield the no-multiple condition for the two sizes in Theorem 9.7 at $A_0 = 2(v+1)$. That theorem either gives zero loss (in the pure-low second case or when the negative weights avoid $[-A_0, 0)$) or gives a contribution of order $d_F(\mu(v+1) + (W_0+1)\,\mathrm{rk}_O\, T)$. In the nonzero case the large negative gap entails $h_*/5 \le v$, and the numerical assumptions imply $W_0 + 1 \le K$. Thus this too is absorbed by (163).

The exact floor and scale comparisons, including the sparse-thick budget in the first subcase, are part of the formal-source input `crystalline_defect_le`. Notice the structural improvement: the second term has the number of low jumps, not the rank, multiplying the large factorial valuation. Without that refinement, the eventual bound at rank-one string blocks and at the main row block would not have the required common budget.

9.10 Insertion of the global weights and the sum over all blocks

Return to the admissible datum $\mathcal{D}$. For this application put $h_* = h_2 = r^8$, $Q = kt$, and set

$$b_1 = N_1 \left\lceil \frac{a}{xN_1} \right\rceil, \qquad b_2 = N_1 \left\lceil \frac{a}{x_1 N_1} \right\rceil,$$

so $w_j(\sigma) = (b_1 \text{ or } b_2) + 2jh_2 + h_2$ at the two kinds of places. The main row has rank r' with $1 \le r' \le r$, and its block string has length $t' \le t$. Its jumps at a complex embedding σ satisfy the alternative

$$h_2 \le 2|u| \le 2h_2^2 \quad \text{or} \quad |\epsilon_\sigma u + w_j(\sigma)| \le t,$$

for one $j \le m$; the number of jumps of absolute value at most $2h_2^2$ is at most $r'm/k$ per embedding. Here ϵ_σ is the sign of the CM embedding, $+1$ for the chosen CM type, -1 for its conjugate. The source also requires the sum of negative jump counts at $\sigma, \bar\sigma$ to be at most r'. These are the hypotheses `IsRowType` in `New/Selmer/DeterminantBound.lean`, used for the actual main row representations. The one-dimensional extra blocks are cyclotomic powers with jumps $c \in [0, t')$. After taking Tate duals the new large windows are opposite to the original ones, by (146), while the rank-one blocks have all jumps small and nonpositive. The local substitution in (163) is $b_* = b_1$, $G_* = b_2$, $W_0 = 4t$, $\eta = h_2^2 + 1$, and $\delta = 2jh_2 + h_2 - 2t$. This is why the input to the crystalline theorem is a count of low jumps rather than an absolute bound on all weights.

The local embeddings above p have sign and place classes related by the Galois action on roots of unity. If $g'g^{-1}$ acts on the prime-to-p part of the cyclotomic generator by the p^n power, the ratio of the infinite-place classes is $\epsilon_\sigma \epsilon_{\sigma'} p^n$ modulo the relevant prime-to-p divisor of N_1. This is `placeClass_ratio`. Squaring removes the sign, and the three-term avoidance in $\mathcal{S}$ gives `not_thick_of_short_arc`: for $p \le kt$, if two thick embeddings are separated by a Frobenius arc of length $n \le 27s$ then a third at interior distance $0 < i < n$ cannot be thick (where s is the cyclotomic tower exponent). Furthermore at $p \mid N_1 d$ two thick embeddings in one local Galois orbit have the same CM sign, by `cmSign_eq_of_thick`. These are specific arithmetic *formal-source inputs*, `New/Selmer/ThickSpacing.lean`, used to verify the small-prime hypotheses of (163). At large primes away from the conductor the local extension is unramified and the larger-prime hypotheses apply. These separate facts explain why it is not sufficient merely to take a large set of thick places, without the progression-freeness built into $\mathcal{D}$.

Define now the local cost used in the global proof:

$$\mathrm{loss}_p = 10^6 \frac{m}{k}(b_1 v_p(g) + \lambda_p(h_2, b_1)), \qquad g = g_0 g_1. \tag{164}$$

At the selected split level primes use (161), with the degree bound for the polynomials derived from the good-type eigenvalues at those primes; the source uses $r_0 = r$ and $N_0' = (r+t)r$. At other primes use (163). With the preceding scale and spacing verifications the result is the following exact statement.

Proposition 9.8 (One-place block estimate)**.** Formal-source input. *In the above setting, let $v \mid p$ and let T_i be the lattices of the row plus string blocks (one row, and one for each $c \in [0, t')$). Let $\mathcal{Q}_v$ be the class of local lattices inherited from the chosen column representations, with the same jump restrictions and crystalline condition, or the stated tame polynomial condition when $p \mid g$. For all sufficiently large coefficient rings, a block belonging to this class has*

$$\Delta_p(T_i^*(1); \mathcal{Q}_v) \le [E_v : \mathbb{Q}_p] \operatorname{rk}_O T_i \frac{\operatorname{loss}_p}{q_i}, \qquad q_i = \begin{cases} 1, & i \text{ the row,} \\ t, & i \text{ a cyclotomic string block.} \end{cases}$$

This is `block_defect_le_at_place` in `New/Selmer/OnePlaceDefect.lean`. The verification of the numerical parameters and the precise embeddings-to-letters dictionary are included in that declaration. It uses the very large gap between $p = g_0, g_1$ and the weights in the tame case. The row-type assumptions for actual automorphic Galois representations must be proved by the other part of the congruence argument; the local proposition does not give their existence. In particular, the gain for string blocks arises from their special single low weight, not from an assertion that all rank-one blocks have zero crystalline defect.

Let $\operatorname{bloss}_{i,p}(M_i)$ be the normalized relative length of the module of obeying cocycles on the i-th block modulo those also satisfying the crystalline relations simultaneously at all $v \mid p$. By Proposition 9.3 and subadditivity of relative length on intersections,

$$\begin{aligned} \sum_i \operatorname{bloss}_{i,p}(M_i) &\le \sum_{v \mid p} [E_v : \mathbb{Q}_p] \left(r' + \frac{t'}{t} \right) \operatorname{loss}_p \\ &\le 4 r' f \operatorname{loss}_p. \end{aligned}$$

The last step uses $r' \ge 1$, $t' \le t$, and $\sum_{v \mid p} [E_v : \mathbb{Q}_p] = [E : \mathbb{Q}] = 2f$. It is valid for arbitrary choices of the moduli M_i. This is exactly `sum_blockLoss_le` in `New/Selmer/RibetElement.lean`. It gives the $4f \operatorname{loss}_p$ term in the row budget per rank.

Finally the summation of this term over all rational primes with logarithmic weights is especially transparent. Write $J_* = (2b_1)! \left(\prod_{\max(b_1 - 30h_2^3, 0) < j \le b_1 + 30h_2^3} j \right)^{h_2^3}$. For a finite set of primes P_0,

$$\begin{aligned} \sum_{p \in P_0} 4f \operatorname{loss}_p \log p &\le 4 \cdot 10^6 f \frac{m}{k} (b_1 \log g + \log J_*) \\ &\le 4 \cdot 10^6 f \frac{m}{k} (b_1 \log g + 2b_1 \log(2b_1) + 60 h_2^6 \log(b_1 + 30h_2^3)). \end{aligned}$$

Here unique factorization gives the first inequality, and the number of factors in the consecutive product is at most $60h_2^3$. On the admissible asymptotic family $b_1 \sim b$, $\log g \le 22 \log b$, and h_2^6 is polynomial in x whereas b grows faster than every fixed power of x. Thus the cost is $O(f(m/k) b \log b)$. Its ratio to $B_0 = fa(\log b)^3/(mk)$ is

$$O\left(\frac{m^2}{x(\log b)^2} \right) = O((\log b)^{-2})$$

using $a \ge xb$, $m^2 \le 4x$. The source proves the requisite fixed-budget inequalities for all admissible data, including the other part of the row budget, in `exists_isFitFamily`,

`New/Selmer/FitTwistsExist.lean`. The displayed calculation identifies why the arithmetic parameters make even the high-weight local integral comparison affordable. At this point a loss of order $\log J_*$ is acceptable, but the large powers of r entering the rank and level cannot be permitted to multiply the factorial part indiscriminately. This is the significance of the rank-uniform depth method, the separate tame-polynomial estimate, and the low-jump colength bound in the proof.

10 The arithmetic inequality and conclusion

We assemble the three arithmetic results in their exact quantitative form. The proof of their consequence is independent of the internal choices of kernels.

10.1 Collecting the arithmetic inputs

With the data of Section 3, set $n = \dim_{\mathbb{C}} W$, $c = \dim_{\mathbb{C}} C$, and retain

$$H = mk\mathcal{U}, \qquad B = fa(\log b)^3, \qquad L = B \log b.$$

For clarity, we give the particular form of the three theorems that is needed in the calculation. The notation in the three statements below is intentionally the same. The preceding sections give their derivation and identify specialized inputs still requiring independent written proofs.

Theorem 10.1 (A: Short forms, recalled)**.** *There are a rational number $q_0 > 0$ and a set $S \subseteq C^{\mathrm{ram}}$ of integral forms spanning C^{ram} over $\mathbb{C}$ such that, putting $T_s = q_0 P_s$, one has*

$$\|\Theta_1 T_s(v)\|_\sigma^2 \le \exp(H - m\mathcal{R} + L) \quad (v \in S,\ s \in \mathcal{G},\ \sigma \in \mathrm{Aut}(\mathbb{C})).$$

For every finite set $F \subseteq \mathcal{G}$ and every finite family y in the images $T_s(C^{\mathrm{ram}}_{\mathrm{int}})$, $s \in F$, one has

$$\mathrm{IndexLe}(y, |F|cB).$$

Theorem 10.2 (B: Enough forms, recalled)**.** *The spaces C and W have positive dimension and there is a finite set $F \subseteq \mathcal{G}$ such that*

$$|F|c \le 5n \log(1 + b^2), \qquad \dim\left(\sum_{s \in F} \mathrm{im}(P_s)\right) \ge (1 - 4/b^2)n.$$

Theorem 10.3 (C: Integral forms are long, recalled)**.** *Let $y = (y_1, \dots, y_q)$ be q independent row forms with $q \ge (1 - 4/b^2)n$. If $\mathrm{IndexLe}(y, d_{\mathrm{ind}})$ and for every σ one has $\mathrm{covol}^2_\sigma(y) \le e^V$, then*

$$n(H - 3L) \le V + 2d_{\mathrm{ind}}.$$

Together with the crude bound

$$|H| + m|\mathcal{R}| \le b^2 x^{99}, \tag{165}$$

these are precisely the inputs in `star_inequality` of `New/Main.lean`. The source proves the latter as `Data.abs_mainHeight_add_remainder_le`. We give the elementary verification along the admissible asymptotic family.

Lemma 10.4 (Crude size comparison)**.** *Along the data in Proposition 3.2, including the chosen second cutoff, inequality (165) holds for all sufficiently large x.*

Proof. The factor $\lambda_{d,1}(\nu)$ is bounded by 2ν. Thus $0 \le A(z,\sigma) \le 2az$ for $0 \le z \le a$, since each summand is at most $2z$. From $w_j \le 2b$, $|\log(\Omega/g)| \ll_{N_0,d} \log b$ (by (16) and the bounds on g), and $\log(j!) \le j \log j$, we have $|J(w_0(\sigma))| \ll b \log b$. The nominal weights range between a/x_1 and a/x, so $|A_0| \ll a \log b$. From the definitions,

$$|H| \ll mkf(b \log b + ab), \qquad m|\mathcal{R}| \ll mf(a^2 + a \log b).$$

Here $a \le 2xb$, $k \ll x^5$, $f \ll_{N_0} x^{11} \log x$, $m \ll x^{1/2}$, and $\log b = x^{\varepsilon} \le x$. These upper bounds are $O_{N_0,d}(b^2 x^{20})$ for large x. Enlarging the threshold gives (165). The proof in the compact source instead uses explicit logarithmic inequalities to have the same bound on the full admissible record, a stronger quantification not needed for the eventual inequality here. □

10.2 From A and B to an almost-basis of short forms

Choose F as in Theorem 10.2 and use the set S from Theorem 10.1. Since q_0 is nonzero and S spans C^{ram}, the set

$$\{T_s(v) : s \in F,\ v \in S\}$$

spans the sum of the images of the P_s, $s \in F$. Choose independent vectors $y_1, \dots, y_q$ from this set which form a basis of its span. In particular,

$$(1 - 4/b^2)n \le q \le n.$$

By A and the bound on the number of translates we may take

$$d_{\mathrm{ind}} \le 5n \log(1 + b^2)B.$$

For each conjugated pairing Hadamard's inequality gives

$$\begin{aligned} \mathrm{covol}_\sigma^2(y) &\le \prod_{i=1}^{q} \|\Theta_1 y_i\|_\sigma^2 \\ &\le \exp\left(q(H - m\mathcal{R} + L)\right). \end{aligned}$$

Thus $V = q(H - m\mathcal{R} + L)$ is a valid upper bound. This is the conclusion of `Data.exists_family_covolumeSq_le` in `New/Endgame/LongForms.lean`. The step is only linear algebra, but realizes the crucial purpose of Theorem B: it converts estimates on vectors arising from a particular automorphic construction into an estimate on almost the entire determinant of the row space.

10.3 Comparison with the lower bound

Proposition 10.5 (The main arithmetic inequality)**.** *For all sufficiently large x along the admissible family of Proposition 3.2, the three arithmetic inputs imply*

$$m\mathcal{R} \le fa(\log b)^5. \tag{$\star$}$$

Proof. Write $h = H - m\mathcal{R} + L$ and apply C to the family just constructed. Then

$$\begin{aligned} n(H - 3L) &\le qh + 2d_{\mathrm{ind}} \\ &= nh + (n - q)(-h) + 2d_{\mathrm{ind}}. \end{aligned}$$

We cannot simply replace q by n in qh, because h may be negative. Instead, use (165) and $L > 0$ to get $-h \le b^2 x^{99}$, and the dimension estimate to get $n - q \le 4n/b^2$. Consequently

$$n(H - 3L) \le n(H - m\mathcal{R} + L) + 4nx^{99} + 10n \log(1 + b^2)B.$$

Divide by $n > 0$, cancel the common H, and rearrange:

$$m\mathcal{R} \le 4L + 4x^{99} + 10\log(1+b^2)B. \tag{166}$$

The scale inequalities imply that the right-hand side is at most $(\log b)^2 B$. For the eventual assertion this is immediate from $a \ge xb$, $B = fa(\log b)^3$, and $\log b = x^{\varepsilon}$, since $x^{99}/B \to 0$ and $\log(1+b^2) = O(\log b)$. The Lean version uses its explicit admissibility bounds to make this last comparison for every admissible datum, not just sufficiently large x along this family. Substituting the value of B gives ($\star$). □

It is worth looking closely at the size of the three errors in (166). The $4L$ is the sum of the error per vector in the upper bound and the three units of L in the lower bound. The $4x^{99}$ is the penalty for the missing dimensions, which are so few that even a very crude bound on h suffices. The last term is twice the index. All are smaller than the final budget by a factor of order $\log b$ or more along the asymptotic family. The leading term H has disappeared completely. Without the matching of its exact leading coefficient in A and C, this argument would not give the desired bound.

10.4 Extracting the analytic remainder

Here is the short algebra behind (23). Use the notation of Section 2 and write

$$S_M(z) = \sum_{\nu \le z} \lambda_{d,M}(\nu)(1/\nu - 1/z).$$

For a fixed integer ν, the number of $\sigma \in \mathcal{T}$ such that the place $s_\nu \sigma$ lies in $\mathcal{S}$ is $|\mathcal{S}|$ if $\nu \equiv 1 \pmod{N_0}$ and zero otherwise. Indeed, multiplication by s_ν carries the whole coset $\mathcal{T}$ onto the coset of class $-\nu$ when ν is invertible modulo N_1. The thick places have class -1 modulo N_0. Every $\nu \equiv 1 \pmod{N_0}$ is a unit modulo N_1 in this tower. In the remaining nonunit cases the source's place convention places the product back in $\mathcal{T}$, so it is not thick. It follows that replacing the weight a/x_1 by a/x inside the sum over the tops gives the identity

$$\frac{1}{a}\sum_{\sigma \in \mathcal{T}} A(a,\sigma) = |\mathcal{S}|(S_{N_0}(x) - S_{N_0}(x_1)) + |\mathcal{T}|S_1(x_1).$$

Here S_M denotes the sum before subtraction of its main terms. By the analytic normalization,

$$S_M(z) = R_M(z) + \frac{\log z}{\varphi(M)} + c_{d,M}.$$

The constants for $M = N_0$ cancel in the difference. For $M = 1$, we have $c_{d,1} = -1 - \kappa_K$. Moreover

$$\log a - \overline{w} = \log x_1 + \frac{|\mathcal{S}|}{f}(\log x - \log x_1).$$

Since $f = \varphi(N_0)|\mathcal{T}|$, we have exactly

$$\frac{|\mathcal{T}|A_0}{a} = |\mathcal{T}|(\log x_1 - 1 - \kappa_K) + \frac{|\mathcal{S}|}{\varphi(N_0)}(\log x - \log x_1).$$

Subtract these main terms from the previous sum identity. This proves (23), including its constants. In particular, no estimate involving the Chowla–Selberg constant is needed at this final stage.

10.5 Choosing the second cutoff and the power saving

Fix $0 < \varepsilon < 1/100$ and choose admissible data by Proposition 3.2 for every sufficiently large x. The quantity $X = |\mathcal{T}|/|\mathcal{S}|$ satisfies $0 < X \le \log b$. Write $T = 10^6 \log x$. For large x we have $X \le e^T$. Apply the window lemma of Section 2 to the two functions $u \mapsto R_{N_0}(e^u)$ and $u \mapsto R_1(e^u)$. It supplies $u \in [T, T^2]$ such that

$$R_{N_0}(e^u) - \frac{|\mathcal{T}|}{|\mathcal{S}|} R_1(e^u) \le e^{-T}.$$

Take $x_1 = e^u$ in the auxiliary data, as permitted by Proposition 3.2. This is the decisive use of the windowed argument: there is no need for a strong pointwise estimate at the initially specified, extremely large, second scale.

Solve (23) for the desired term and use ($\star$):

$$\begin{aligned} R_{N_0}(x) &= \frac{\mathcal{R}}{a|\mathcal{S}|} + R_{N_0}(x_1) - \frac{|\mathcal{T}|}{|\mathcal{S}|} R_1(x_1) \\ &\le \frac{f(\log b)^5}{m|\mathcal{S}|} + e^{-T} \\ &\le \frac{\varphi(N_0)(\log b)^6}{m} + x^{-10^6} \\ &\le \varphi(N_0) x^{-1/2+6\varepsilon} + x^{-10^6}. \end{aligned}$$

Absorbing the fixed constant using 2ε of additional slack gives, for all sufficiently large x,

$$R_{N_0}(x) \le x^{-1/2+8\varepsilon}. \tag{167}$$

This is the precise exponent in the compact development's `eventually_primeError_le_rpow`. Replacing ε by a smaller positive number proves a bound with any fixed exponent greater than $-1/2$.

Remark 10.6 (Where the critical line enters). Everything up to ($\star$) concerns norms, divisibilities, dimensions and heights. The saving $x^{-1/2}$ arises from dividing the error by $m \ge \sqrt{x}$. The large field degree, large weights and rank do not themselves account for it. One cannot hope to improve the analytic conclusion by choosing m arbitrarily larger: the arithmetic theorems and their errors depend crucially on the admissible relationship between the length, the weights, ranks and local losses. A proof of stronger short/long-form estimates on other ranges would be an additional mathematical result.

10.6 GRH

To finish, suppose a Dirichlet character χ modulo q had a zero ρ to the right of $1/2$, excluding the pole convention of Theorem 1.1. Choose $N_0 = 18480q$ and regard χ as a character of level N_0 by induction; the induced L-function differs by finite Euler factors with no zeros in this region. The class $N_0/2 - 1$ is a unit modulo N_0 and forces $d \equiv 7 \pmod 8$. For every prime d in this class, the above arithmetic argument gives (167). Choose ε positive and sufficiently small that $\beta = 1/2 + 8\varepsilon < \Re(\rho)$. The one-sided criterion of Proposition 2.9 forbids the zero ρ. This proves the deduction of Theorem 1.1 from the arithmetic inputs; a complete conventional proof requires independently establishing the remaining formal-source inputs listed in the final section. The standard functional equations imply the statements about the critical strip. This completes the deduction.

11 Dependency audit and remaining proof obligations

11.1 Status of the reconstruction

The principal claims of the manuscript are conditional only in the following explicit sense: an assertion taken from the supplied formal development without reconstructing its proof is designated a *formal-source input*. This is not a new mathematical axiom, and the authors of the formal source report that it is itself established in Lean; it is nonetheless an outstanding task for a proof that is to be checked solely by reading a conventional paper. Several inputs are stronger than what is usually stated in published versions of the relevant theory. The following table is a guide to the highest-priority proof obligations, not an assertion that the displayed packages exhaust every imported background declaration.

Package	Primary compact-source anchor(s)	What must be justified in a written proof
CM period and formal group (A)	`New/Curve/PeriodDefs.lean`; `FormalParameter.lean`; `ConjugateGenerator.lean`; `TightFrameExists.lean`; `FrameSections.lean`	Compatible period scale, integral logarithmic identity, uniform global generators and Frobenius pairing in the frame expansion.
Schur jets and inert-prime losses (A)	`New/Curve/JetDivisibility.lean`; `New/Columns/JetExponentBound.lean`; `New/Curve/LiveSum.lean`	Check exact integral coefficient and differential identities for the specific Weyl-coordinate model, including the marked/unmarked terms.
Iterated seesaw (A)	`New/Seesaw/SeesawTheorem.lean`; `UnfoldingStages.lean`; `GoodDatumExists.lean`	Construction of compatible lattices and kernels, complete unfolding and denominators for each twist, with a single normalization independent of the twist.
Hecke elementary divisors (A)	`New/TwistedHecke/`; `New/Columns/LocalStatement.lean`	Character projection through the integral operators; exceptional-prime multiplicity and root estimates for the actual automorphic lattices.
Galois-conjugate sizes (A)	`New/Columns/ChainValuation.lean`; `New/Columns/LocalStatement.lean`; `New/Endgame/ShortAndEnough.lean`	Archimedean bounds with common CM normalization; compatibility of the correction with ramified level N.
Finite-level transfer (B)	`New/Trace/GoodIdempotent.lean`; `New/Trace/BadShare.lean`; `New/Filling/FillingTheorem.lean`	Dimension bounds for the specified real types and principal-level finite group projector; verification of exact trace-transfer hypotheses.
Bessel systems (B)	`New/Filling/FillingData.lean`; `New/RealPlace/BandLimits.lean`; `New/RealPlace/OpenCellBound.lean`; `New/DiscreteSeries/`	Uniform lower normalization of peaks, truncation and Schur bounds as functions of the growing weights, discriminants and ranks.
Relative second moment (B)	`New/Filling/Moments.lean`; `CrudeEnergy.lean`; `GlobalSieve.lean`; `GlobalToLocal.lean`	Unfolding with precise measures, local-to-global transfer for coherent pairs, sharp main term with leading constant one and small residual terms.

Package	Primary compact-source anchor(s)	What must be justified in a written proof
Exchange and auxiliary rows (C)	`New/Exchange/`; `New/Exchange/CrudeRows.lean`	Complementary-rank identities of theta periods, local factors at ramified places, compatibility of parameters throughout the two grids.
Volume and one-lift ratios (C)	`New/Rows/RowVolume.lean`; `New/Rows/LiftVolumeHeight.lean`; `New/Rows/HeckeValueHeight.lean`	Check adjoint and Rallis pairing identities, triangular volume decomposition, normalization of auxiliary forms and exact rank-one special-value ratios.
False lifts and comparison projections (C)	`New/Rows/ComparisonForms.lean`; `New/Rows/StackedDeterminants.lean`; `New/Selmer/FitTwists.lean`	Identify precisely which nonlift eigenforms are annihilated by prescribed Hecke operators and why integral determinant bounds suffice.
Galois input and congruence (C)	`New/GaloisRep/`; `New/Selmer/`; `New/Patching/`; `Lib/GaloisRep/`	Match the exact normalized Hecke polynomials to the weak attached representations; triangular auxiliary primes, first-order congruence classes and dual Selmer conditions.
Twist selection (C)	`New/Selmer/FitTwistsExist.lean`; `New/Selmer/FitTwists.lean`; `New/Selmer/`	Simultaneous Selmer control at every local embedding and nonzero value ratios with specified log budget; verify finite-presentation and twisting hypotheses.
Integral crystalline loss (C)	`New/Lift/`; `Lib/PadicHodge/`; `Lib/PadicComparison/`	Integral Kummer-chart frame lemma, depths and cokernel ideals, rank-uniform duality defect and local crystalline conditions when residual types vary.

The standard results used in Sections 2 and 3 (Dirichlet L-functions, theta transformation, Pólya–Vinogradov, Siegel–Walfisz, class-group and cyclotomic facts) are independent of GRH. The zero-density argument here is a proof confined to the elementary estimates and classical inputs cited there. In particular one does not need to assume an unproved uniform error term for primes in arithmetic progressions in order to pass from $\mathcal{R}$ to $R_{N_0}(x)$.

11.2 Reference matching is a mathematical task

It is profitable to separate two different kinds of work. First, reconstruct the exact theorems in the previous table and their deductions, keeping all scaling factors and quantifiers. Second, compare the background they use with the published results at the appropriate level of generality. A standard theorem can be cited instead of proved only after its hypotheses, conventions and uniformity have been shown to imply the needed assertion. For instance, qualitative theta nonvanishing does not give the uniform nondegeneracy and norm estimates required by the volume argument; a standard crystalline comparison theorem over a fixed field does not automatically supply a bound for a varying integral lattice after reduction modulo a large power of a prime.

The formal repository's `REFERENCES.md` in the compact variant lists 196 citations compared to varying extents: 57 directly checked and precisely located, 30 locations identified at second

hand, 102 works found without a precise location, and 7 not compared. Several other files were outside this check. We do not treat this list as a verified bibliography of every assertion in the reconstructed `Lib` directory. The readable edition's `INDEX.md`, `CONTENTS.md`, and `GUIDE.md` are particularly useful for matching a large formal theorem to its intended mathematical counterpart. They likewise are documentation, not external peer review.

The three principal reductions are now narrow mathematical interfaces. In the local-to-global upper bound the outstanding proofs are concentrated in the normalized CM kernel and the Hecke/seesaw packages; in spanning they are concentrated in the uniform Bessel and sieve estimates; in the determinant lower bound they are concentrated in exchange/period identities and the congruence and crystalline frame machinery. This separation is intended to support independent checking and more efficient development of the complete research proof. The paper should be understood as a substantial intermediate draft on that route.

References

[1] L. Alpöge, with Claude, and the authors of the shortened development, *GRH for Dirichlet L-functions, in Lean 4*, supplied snapshot `grh-lean-main`, November 2026. Compact and readable editions (Apache 2.0); the root README links the announcement `https://alpo.ge/remark.pdf`. Main statements: `src/Statements.lean`. Main mathematical interfaces: `src/New/Main.lean`, `src/New/Endgame/ShortAndEnough.lean`, `src/New/Endgame/LongForms.lean` in the compact edition.

[2] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I: Classical Theory*, Cambridge Studies in Advanced Mathematics 97, Cambridge University Press, 2007. In particular Chs. 9–11 and Cor. 11.19, and Ch. 15 for Landau arguments. https://doi.org/10.1017/CBO9780511618314.

[3] M. Jutila, On mean values of Dirichlet polynomials with real characters, *Acta Arith.* **27** (1975), 191–198. https://doi.org/10.4064/aa-27-1-191-198.

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

[5] L. C. Washington, *Introduction to Cyclotomic Fields*, 2nd ed., Graduate Texts in Mathematics 83, Springer, 1997. https://doi.org/10.1007/978-1-4612-1934-7.