hexagon:2611.00809v2

A PRELIMINARY NOTE ON ALPÖGE'S PROPOSED PROOF OF THE GENERALIZED RIEMANN HYPOTHESIS

ROBIN ZHANG

Contents

0. Statement from the author 1
1. Mathematical overview 2
2. Detecting zeros from sums over prime powers 14
3. Recovering sums over prime powers from divisibility of periods 19
4. Bounding sums over prime powers with automorphic periods 23
5. Controlling denominators with Galois cohomology 48
6. Choosing the parameters to obtain GRH 57
References 60

0. Statement from the author

The main purpose of this preliminary report is to begin looking for some of the sources, provenance, and mathematical history behind the proposed proof of the generalized Riemann hypothesis announced by Levent Alpöge using a Claude model [Alp26] on November 20, 2026. The mathematics involved reflects many lifetimes of work, so I would like to recognize as many as possible of the people whose work may have contributed to it: the mathematicians working today, our teachers, and all of our predecessors. This is only a beginning, and I apologize for all the people, papers, and other contributions I may have missed.

The ideas used in modern mathematical proofs inevitably come from people, and I think that sharing them carries obligations to those people; see Harris [Har26a] for a great account of this notion of mathematics as a "gift economy". I also wonder what becomes of the labor and generosity behind the ideas, and who is expected to do the further work of understanding any proof announcement. While I share Michael Harris's [Har26b] objection to mathematicians acting as unpaid labor for AI companies, I am attempting to interpret this AI artifact because I want to recognize some of the mathematical origins of [Alp26] and the credit that is due to the people involved (without claiming to offer a complete history or to settle any questions of ownership).

Date: November 20, 2026.

I hope we will continue to remember and celebrate the people whose ideas have allowed the rest of us to see a little further.

The scale of the material involved in [Alp26] makes this an unusual exercise in exposition: the Lean code has roughly 17 million lines of code, while Alpöge [Alp26, p. 2, footnote 2] describes an AI-generated exposition of more than 250,000 pages that he chose not to release. Neither offers a practicable prospect of reading the argument from beginning to end in its present form. As such, I have also used AI to begin writing this report: I prompted GPT 6 Pro to digest the Lean code on Alpöge's website, infer a human-readable mathematical proof, to search the literature carefully for earlier sources and analogues of the ideas in the argument, and then to produce a written summary. These prompts were an attempt to obtain something that people could begin to read, discuss, and eventually improve. I then spent November 20, 2026 rewriting the AI output into the current form. Given the impossibility of any human checking 17 million lines of Lean code in a short timeframe, I cannot claim that this summary is a faithful translation of the Lean code in [Alp26], let alone that it actually proves the generalized Riemann hypothesis.[1] I also apologize to the reader for the condensed and imperfect exposition here, as most of the text from Section 2 onwards was largely written by AI.

Part of the pleasure of mathematics, for me, is that an idea can pass from one person to another and take on a life that neither could have anticipated. There is an art in making that possible, through the patient work of explanation and the equally patient work of learning, rooted in the earliest oral traditions of human culture. I hope that mathematicians will come to understand both proposed proofs of the generalized Riemann hypothesis that were announced on November 20, 2026,[2] discover what is new in them, and explain what further questions they make possible. I would be very glad if this small report helped someone write such an account. Such understanding should be available to anyone who wishes to learn, regardless of access to powerful AI systems.

1. Mathematical overview

1.1. **Zeros of L-functions.**

The connection between prime numbers and L-functions began with a question about arithmetic progressions in the

[1] In the absence of a full paper, this would require, at minimum: checking that the Lean definitions agree with standard mathematical definitions, and checking the Lean verification using independent implementations of the Lean kernel.

[2] Alessandro Masullo [Mas26] has also announced a proposed proof of GRH on Ravi Vakil's Math 216 course website on November 20, 2026. Vakil's announcement reported that no AI was used in Masullo's work, and the paper seems to have been the result of nearly twelve years of work with a premature release prompted by the announcement of [Alp26] (not unlike the series of events that led to the release of [Sta26]). As a (monumental!) human effort that acknowledges previous ideas, it does not have the same need for this kind of AI-assisted historiographical exercise.

early 19th century. Dirichlet [Dir37] wanted to show that every arithmetic progression

$$a_1, a_1 + \delta, a_1 + 2\delta, \dots$$

with relatively prime first term a_1 and common difference δ contained infinitely many primes. Certain periodic multiplicative functions $\chi : \mathbb{Z} \longrightarrow \mathbb{C}$, now called Dirichlet characters, allowed him to distinguish the arithmetic progressions, and his L-functions for $\mathrm{Re}(s) > 1$,

$$L(s,\chi) := \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s},$$

made it possible to study the primes in each. Riemann [Rie59, p. 674] then connected the distribution of primes with the zeros of the zeta function, and in 1859 used the assertion that its nontrivial zeros lay on the critical line $\mathrm{Re}(s) = 1/2$ as a working hypothesis. The extension of the Riemann hypothesis to Dirichlet L-functions already appeared in the work of Piltz [Pil84, pp. 6, 25] in 1884, and is the generalized Riemann hypothesis considered in [Alp26].

Conjecture 1.1 (Generalized Riemann hypothesis for Dirichlet L-functions)**.** For every primitive Dirichlet character χ, every zero of $L(s,\chi)$ in the strip $0 < \mathrm{Re}(s) < 1$ lies on the line $\mathrm{Re}(s) = \frac{1}{2}$.

Throughout the rest of the note, $\mathbb{N}$ denotes the positive integers, $\mathbb{Z}_{\geq 0}$ denotes the nonnegative integers,

$$\Lambda(n) := \begin{cases} \log p & \text{if } n = p^k \text{ for a prime } p \text{ and } k \in \mathbb{N}, \\ 0 & \text{otherwise,} \end{cases}$$

denotes the von Mangoldt function, and $\gamma = 0.577\dots$ denotes Euler's constant. We will also assume from now on that χ is a primitive Dirichlet character, since removing Euler factors to obtain an imprimitive character introduces no new zeros of the corresponding L-function in the region $\mathrm{Re}(s) > 0$.

1.2. **Why sums over prime powers enter the argument.** To exclude zeros of $L(s,\chi)$ to the right of the critical line, [Alp26] studies its logarithmic derivative. For $\mathrm{Re}(s) > 1$, the identity $-L'(s,\chi)/L(s,\chi) = \sum_{n\geq 1} \Lambda(n)\chi(n)n^{-s}$ connects the zeros of $L(s,\chi)$ with prime powers: zeros give poles after meromorphic continuation of the left side, while $\Lambda(n)$ vanishes unless n is a prime power. To make this relation quantitative, [Alp26] uses the weighted sums, for $x \geq 1$,

$$\begin{aligned} S_\chi(x) &:= \sum_{n\leq x} \Lambda(n)\chi(n)\left(\frac{x}{n} - 1\right) \\ &= \sum_{p} \sum_{\substack{k\geq 1 \\ p^k \leq x}} (\log p)\chi(p^k)\left(\frac{x}{p^k} - 1\right). \end{aligned}$$

The weight $x/n - 1$ makes the Mellin transform of S_χ particularly simple:

$$\int_1^\infty S_\chi(x) x^{-s-1}\,dx = -\frac{L'(s,\chi)}{s(s-1)L(s,\chi)} \qquad \text{for } \mathrm{Re}(s) > 1.$$

The integral converts information about the growth of $S_\chi(x)$ into information about the poles of the expression on the right. A zero of $L(s,\chi)$ in the critical strip produces such a pole, so bounds for $S_\chi(x)$ can constrain the location of the zeros.

Before applying this observation, however, the contribution at $s = 1$ must be removed. For a nonprincipal primitive character, [Alp26] cancels the pole introduced by the factor $1/(s-1)$ by setting

$$G_\chi(x) := S_\chi(x) + xL'(1,\chi)/L(1,\chi).$$

The principal primitive character, denoted by $\mathbb{1}$, has conductor 1, and $L(s,\mathbb{1}) = \zeta(s)$. The pole of $\zeta(s)$ at 1 requires the different correction

$$G_{\mathbb{1}}(x) := \sum_{n \le x} \Lambda(n)\left(\frac{x}{n} - 1\right) - x\log x + (1+\gamma)x.$$

The Mellin transforms of the corrected functions G_χ are regular at $s = 1$, while retaining the poles contributed by zeros in the critical strip.

[Alp26] proves upper bounds for a particular linear combination of the real parts of G_χ, with primitive characters inducing powers of one character. For a primitive Dirichlet character ψ of order r, let $\mathcal{X}_\psi$ be the set of primitive characters inducing $\mathbb{1}, \psi, \dots, \psi^{r-1}$. The combination is

$$\mathcal{G}_\psi(x) := G_{\mathbb{1}}(x) + \frac{1}{3r+1} \sum_{\substack{\chi \in \mathcal{X}_\psi \\ \chi \neq \mathbb{1}}} \mathrm{Re}(G_\chi(x)). \tag{1.1}$$

The occurrence of all powers of ψ will come from character orthogonality. For the argument about zeros, the important feature is the positivity of the coefficients. Since $\mathcal{X}_\psi$ is stable under complex conjugation, the Mellin transform of $\mathcal{G}_\psi$ can be expressed using the logarithmic derivatives of these L-functions with positive coefficients. The contributions from a zero shared by several of the L-functions therefore cannot cancel.

The quantitative assertion needed from [Alp26] is the following upper bound.

Proposition 1.2. *Let ψ be a primitive Dirichlet character. For every $\eta > 0$ and every $A \ge 2$, there is $Y \ge A$ such that*

$$\mathcal{G}_\psi(x) \le x^{1/2+\eta} \qquad \textit{for every } x \in [Y, Y^{3/2}]. \tag{1.2}$$

Notice that Proposition 1.2 bounds $\mathcal{G}_\psi$ only from above; the function itself may take either sign. Moreover, the intervals $[Y, Y^{3/2}]$ need not cover every sufficiently large x. Section 2 explains why these upper bounds on arbitrarily large intervals nevertheless suffice to exclude nontrivial zeros off the critical line.

1.3. **How automorphic periods bound sums over prime powers.** To prove Proposition 1.2, [Alp26] relates the sums over prime powers to periods of automorphic forms. It normalizes the periods so that the values under comparison lie in a number field. At finite primes, [Alp26] proves lower bounds for their valuations: these bounds specify powers of prime ideals that divide the period values. The product formula suggests why this might help: if a nonzero algebraic integer is highly divisible at finite primes, its absolute values at the archimedean places cannot all be small. The sums over prime powers occur in the lower bounds for the valuations of period values at finite primes. An upper bound for the archimedean sizes of the periods can therefore constrain these sums. The difficulty is to make the comparison precise enough to obtain the bound for $\mathcal{G}_\psi$ in Proposition 1.2.

The periods and their relation to prime powers. The automorphic forms in [Alp26] are on unitary groups over a totally real field E^+, associated to Hermitian spaces over an auxiliary abelian CM field E. Here E^+ is the maximal totally real subfield of E, and E contains the abelian extension of $\mathbb{Q}$ cut out by ψ. For each of these unitary groups, its adelic points act on the spaces of automorphic forms by right translation. The automorphic representations used here are irreducible representations occurring in these spaces. Their local components at a real place of E^+ are representations of the real unitary group at that place. In [Alp26], the choice of these real local components is made by prescribing weights; the unitary groups and the corresponding parameters are described in Section 4.1.

The initial pairings in [Alp26] are Bessel periods, among the pairings studied by Gan, Gross, and Prasad [GGP12, § 12 and § 23]. Each pairing uses automorphic forms on two unitary groups, one embedded in the other. The first integration extracts a Fourier coefficient of the automorphic form on the larger group by integrating against the complex conjugate of a prescribed character of a unipotent subgroup. The second integration pairs the restriction of that coefficient to the smaller group with the complex conjugate of the automorphic form on that group. Both integrations take place over the corresponding adelic quotients by E^+-rational points.

After choosing models of the relevant spaces of automorphic forms over a coefficient number field K and normalizing the periods, [Alp26] expresses them as K-bilinear pairings. It is therefore possible to compare the valuations of a normalized period at finite primes of K with its absolute values at the archimedean places of K. The lower bounds for these valuations involve sums over powers of rational primes subject to conditions on Frobenius elements in abelian extensions of $\mathbb{Q}$. By character orthogonality, those conditions can be written as sums of Dirichlet character values. This is where the functions G_χ reappear: after an explicit main term is subtracted, the resulting expression is a combination of the corrected sums G_χ, as in Proposition 3.1. To recover $\mathcal{G}_\psi(x)$ from this expression, [Alp26] averages over suitable subgroups to retain the primitive characters inducing powers of ψ, and chooses the weights to

make their arguments close to $\mathfrak{x}$. The character identity and these choices are described in Section 3.

Comparing the periods produced by theta lifts. To increase the divisibility of the normalized period values, [Alp26] repeatedly applies theta lifting. Integration against a theta kernel, an automorphic function on the product of two unitary groups, carries an automorphic form from one group to the other. Let $\mathfrak{m}$ be the number of successive lifts in the proposed proof. The norms used here are Petersson norms: their squares are integrals of absolute squares over adelic quotients. Each lift increases the divisibility of the period values, but also changes the norms and dimensions of the spaces of automorphic forms. Taking more lifts is useful only if the gain in divisibility survives the errors in estimating those norms and dimensions. This is one reason for the choices of weights of the real local components of the automorphic representations.

But before comparing the periods, there is a question of which spaces they pair: the groups and weights vary during the lifts. Kudla [Kud84] proved identities that expressed the same pairing by integrating a theta kernel in two different orders. These seesaw identities relate the pairings before and after theta lifting. In [Alp26], the identities and the identifications of the selected spaces of test automorphic forms allow all the pairings under comparison to use one fixed finite-dimensional K-space of test automorphic forms. Write W for this space, and let α index the choices of weights in the other argument of the pairing.

To record the divisibility of period values, [Alp26] works with integral lattices. Let N_α be the modified lattice in the space of initial automorphic forms that records the divisibility obtained after the successive lifts, and let C be the Hermitian dual of the integral lattice chosen in W. These are lattices over $\mathcal{O}_K$, the ring of integers of K, with norms at its archimedean places. Thus $C \otimes K = W^\vee$. Writing $P_\alpha(f, g)$ for the normalized bilinear period, [Alp26] obtains maps

$$\mathfrak{b}_\alpha : N_\alpha \otimes K \longrightarrow C \otimes K = W^\vee,$$

defined by $\mathfrak{b}_\alpha(f)(g) = P_\alpha(f, g)$ for $g \in W$. In other words, $\mathfrak{b}_\alpha(f)$ records the periods obtained by pairing f with every test automorphic form in W. Using the lower bounds for the valuations of period values, [Alp26] chooses the lattices so that $\mathfrak{b}_\alpha$ takes N_α into C. Thus the maps $\mathfrak{b}_\alpha : N_\alpha \to C$ are integral.

To bound the norms of $\mathfrak{b}_\alpha$, [Alp26] uses formulas developed by Rallis [Ral87] that express Petersson norms of theta lifts through special values of L-functions and local integrals. The normalizations also involve periods of CM abelian varieties, namely integrals of algebraic differential forms over homology classes on those varieties. These formulas explain the particular main term that was subtracted from the sums over prime powers. When [Alp26] compares the valuations of period values at finite primes with their

archimedean norms, the formulas for CM periods, Petersson norms, dimensions of spaces of automorphic forms, and weighted counts of Hermitian lattices combine to produce exactly the main term in Proposition 3.1. The resulting cancellation is written out in Section 4.5. The upper bound for $\mathcal{G}_\psi$ then depends on bounding the errors left in this comparison.

Why many independent period functionals are needed. Here the period maps are used to bound the covolume of C from above. In the language of Arakelov geometry, this is a lower bound for the mean slope $\mu(\mathsf{C})$: for a fixed rank and field of definition, a smaller covolume means a larger mean slope. The precise normalization is given in Section 4.3. But the norms of the maps $\mathfrak{b}_\alpha$ do not by themselves determine the covolume of C. If, for example, every $\mathfrak{b}_\alpha$ took values in the same line, bounds for their norms would give no information about the other directions, which also contribute to the covolume of C. This is why [Alp26] seeks enough independent period functionals to span nearly all of $W^\vee$.

The first step is to bound the covolumes of the lattices of integral automorphic forms from which N_α is obtained. By interpreting these automorphic forms as sections of bundles on compactified moduli spaces of polarized abelian varieties with prescribed endomorphisms and level structure, [Alp26] uses arithmetic geometry to bound the covolumes of their integral lattices and to construct sections with prescribed vanishing or nonvanishing. Section 4.3 describes the bounds for covolumes and the construction of these sections.

There is a difficulty in passing from N_α to C: a lower bound for the mean slope of N_α does not give the same lower bound for a quotient of N_α, such as the image of $\mathfrak{b}_\alpha$. To handle this difficulty, [Alp26] uses the Harder–Narasimhan filtration: a filtration by sublattices whose successive quotients have decreasing mean slopes. The bounds in [Alp26] control the average slope of these lattices from below and their largest slopes from above, while allowing some successive quotients to have smaller slopes. For a real threshold $\mathfrak{t}$, [Alp26] retains the initial sublattice in this filtration whose slopes are at least $\mathfrak{t}$ plus an allowance for the averaged logarithmic norms of $\mathfrak{b}_\alpha$. This allowance accounts for how much $\mathfrak{b}_\alpha$ can enlarge norms. Since $\mathfrak{b}_\alpha$ preserves the integral lattices, its image then lies in the part of the Harder–Narasimhan filtration of C whose slopes are at least $\mathfrak{t}$.

Notice that this changes the spanning problem. The maps restricted to these sublattices still need to span a large subspace of $\mathsf{C} \otimes \mathsf{K}$. The bounds for the slopes of N_α control how much of the dimension of $\bigoplus_\alpha(\mathsf{N}_\alpha \otimes \mathsf{K})$ is removed at each threshold. Using bounds for inner products of period functionals, [Alp26] shows that when this loss is small, the remaining images still span nearly all of $\mathsf{C}\otimes\mathsf{K}$; the linear algebra needed for this step is recorded in Lemma 4.2. Since those images lie in the part of C with slopes at least $\mathfrak{t}$, only a small proportion of its slopes can be less than $\mathfrak{t}$. Together with a lower bound for the smallest slope of C, the bounds for this proportion can be

integrated as t varies to obtain the lower bound for $\mu(C)$ in Proposition 4.3, in Section 4.4.

Choosing the weights at the complex places. The choice of weights has another purpose: the periods are being used to estimate $\mathcal{G}_\psi$ at a prescribed value of χ. The weights belong to the real local components of the automorphic representations on the chosen unitary groups. Each real place of E^+ corresponds to a complex place of E, namely a pair of complex-conjugate embeddings $E \to \mathbb{C}$. When the sums recording divisibility of period values are expanded in Dirichlet characters, the arguments of the functions G_χ depend on these weights. At a selected set of complex places, [Alp26] chooses the weights so that these arguments are close to χ. At the other places, the arguments are much larger. At those larger arguments, [Alp26] bounds the absolute values of G_χ using regions in which the corresponding L-functions are nonzero.

The lower bounds for valuations of period values at finite primes require enough selected complex places, but also depend on where those places lie. Since $E/\mathbb{Q}$ is abelian, every unramified rational prime determines a Frobenius element acting on the complex places of E. To keep the losses in the bounds for valuations small, [Alp26] chooses the selected places so that the short segments of Frobenius orbits entering these bounds contain at most two selected places. To reconcile these requirements, [Alp26] adapts Behrend's [Beh46] construction of large sets of integers containing no nonconstant three-term arithmetic progression. The absence of certain relations among the integers in Behrend's construction yields this restriction on complex places, as described in Section 4.6.

Why denominators require a separate argument. There is one more difficulty with the integral lattices. To isolate the automorphic forms used in the periods, [Alp26] projects onto selected sums of simultaneous eigenspaces for the commuting Hecke operators on the spaces of automorphic forms over K. Such a projection need not preserve the lattice over $\mathcal{O}_K$. The resulting denominators can cancel part of the divisibility gained through theta lifting. An upper bound for these denominators is therefore part of the comparison of periods.

The obstruction is related to congruences between automorphic forms: distinct systems of Hecke eigenvalues over K may become congruent modulo a prime. To study these congruences, [Alp26] uses representations of the absolute Galois group of E whose Frobenius eigenvalues at unramified finite places are determined by the Hecke eigenvalues, as in Section 4.2. The congruences can then be related to extension classes of Galois representations. Among the contributions to the denominators are groups of classes that become trivial over prescribed completions of E, but need not be trivial over E itself. In Section 5, [Alp26] bounds the orders of these groups to control their contribution to the denominators.

The dependence on the automorphic weights matters here. As the weights vary, the cohomology groups involve different Tate twists, indexed by positive

integers j; a Tate twist multiplies the Galois action by a power of the cyclotomic character, which describes the action on roots of unity. Only a coarser upper bound for the orders of these groups is available at certain exceptional indices, but [Alp26] bounds how many such indices can occur in a given range. This is where the freedom to choose the intervals in Proposition 1.2 becomes useful. The interval in [Alp26] is chosen so that every exceptional index that contributes is small compared with x. The coarser bounds then keep the contribution of these groups to the denominators within the error allowed in the comparison of periods. Lemma 5.4 makes this choice precise; the resulting bound for the orders of the groups holds throughout the interval, as stated in Proposition 5.2.

How the exponent $1/2$ *appears.* The number m of theta lifts now explains the exponent $1/2$ in Proposition 1.2. After division by the coefficient of $\mathcal{G}_\psi(x)$ in the inequality obtained from periods, the two principal errors have sizes x/m and m, up to factors that are arbitrarily small powers of x. Increasing the number of lifts improves the first term but worsens the second. Taking m close to $\sqrt{x}$ makes both terms close to $\sqrt{x}$.

Of course, choosing m in this way does not settle the conditions on the number fields, automorphic weights, and selected complex places. Proposition 6.3 records the simultaneous choice obtained in [Alp26]. With these choices, the errors fit within $x^{1/2+\eta}$ for any prescribed $\eta > 0$. In Section 6, [Alp26] combines the errors from comparing period values, norms, and covolumes to obtain the upper bound for $\mathcal{G}_\psi$ in Proposition 1.2; Lemma 6.1 records the calculation.

It remains to apply the connection with zeros from Section 1.2. A zero to the right of the critical line would produce a pole in the Mellin transform of $\mathcal{G}_\psi$. The positive coefficients in Equation 1.1 ensure that a zero shared by several L-functions still produces a pole. The upper bounds on arbitrarily large intervals exclude such a pole, and the functional equation for Dirichlet L-functions excludes the remaining nontrivial zeros off the critical line, as in Propositions 2.1 and 2.3. Figure 1 summarizes these implications. Figures 2–4 indicate the results used to bound covolumes of lattices of automorphic forms, construct the period functionals b_α, and bound the orders of the Galois cohomology groups, respectively.

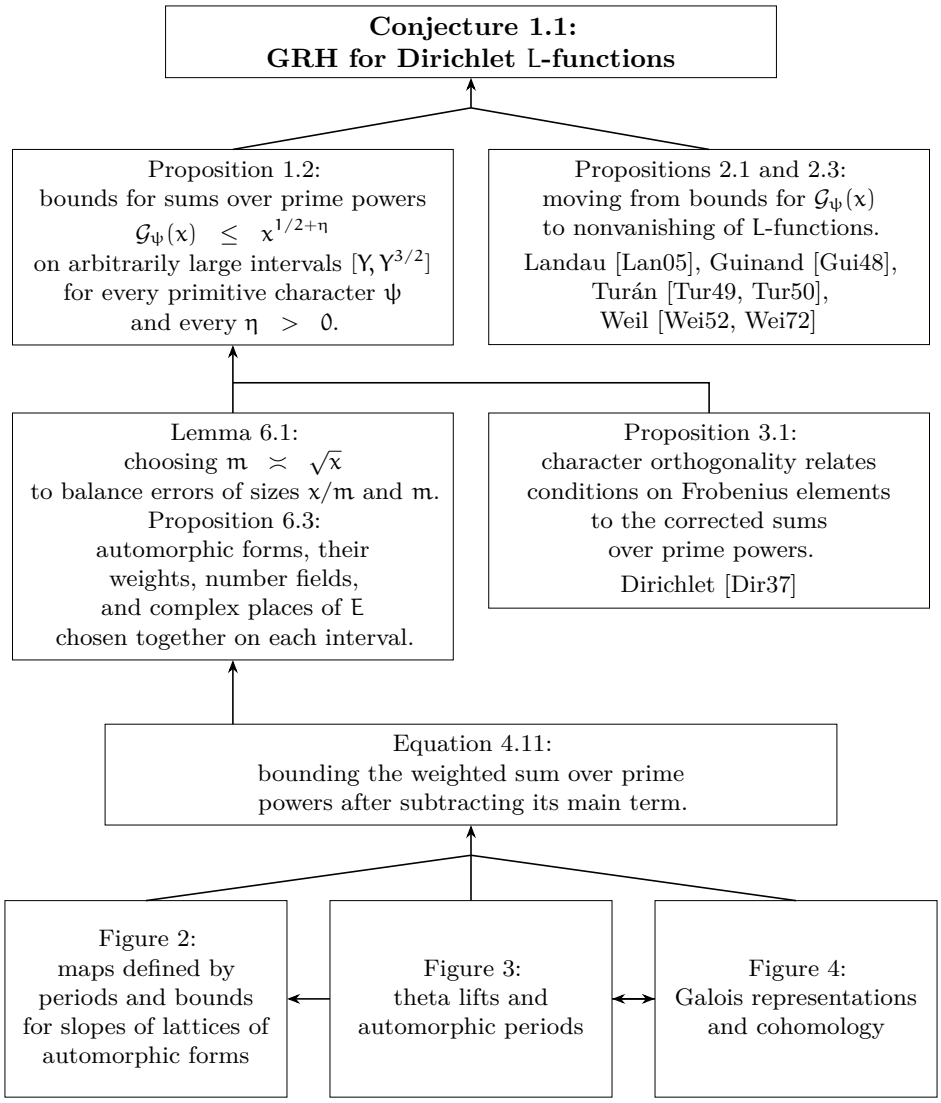


FIGURE 1. The main ingredients in [Alp26], with selected historical references.

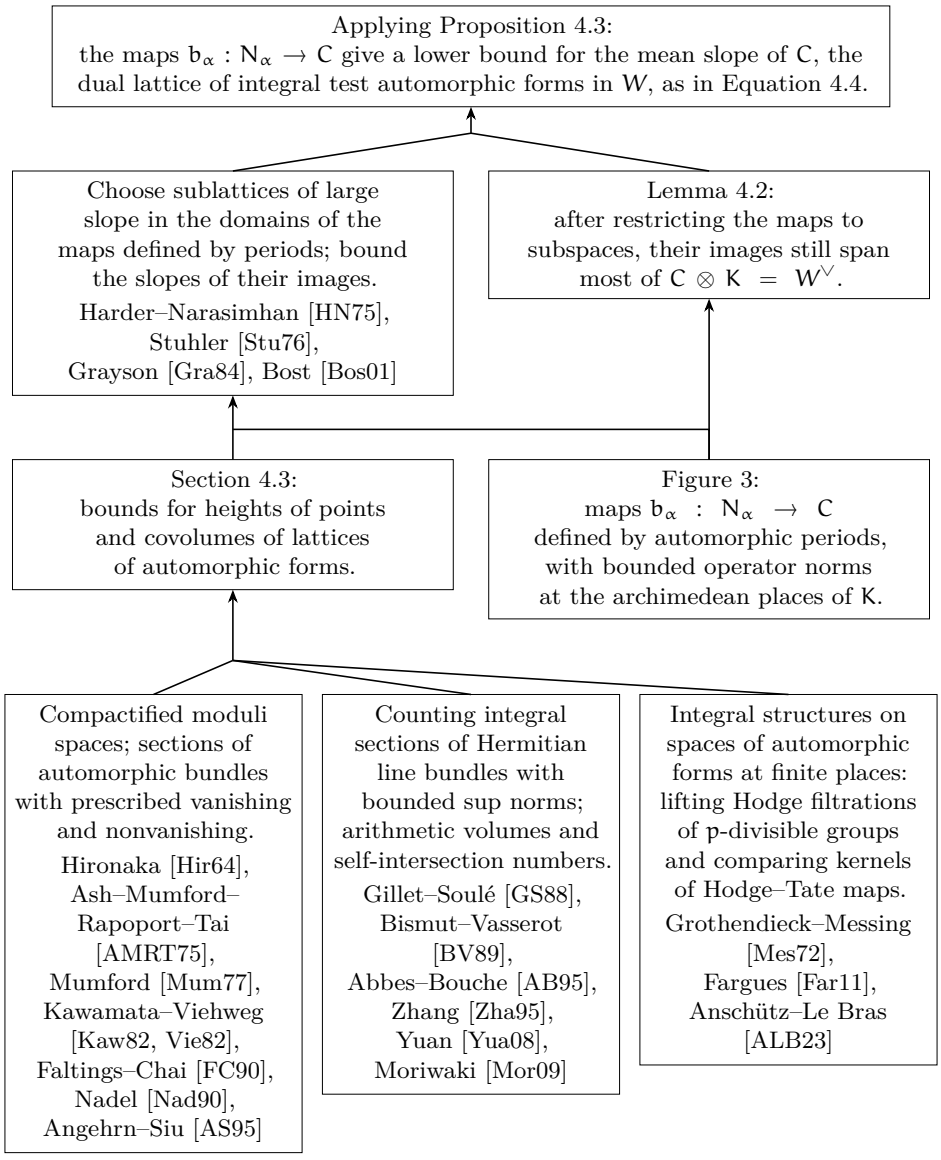


FIGURE 2. The maps defined by automorphic periods and the bounds for covolumes of lattices of automorphic forms in [Alp26], with selected historical references.

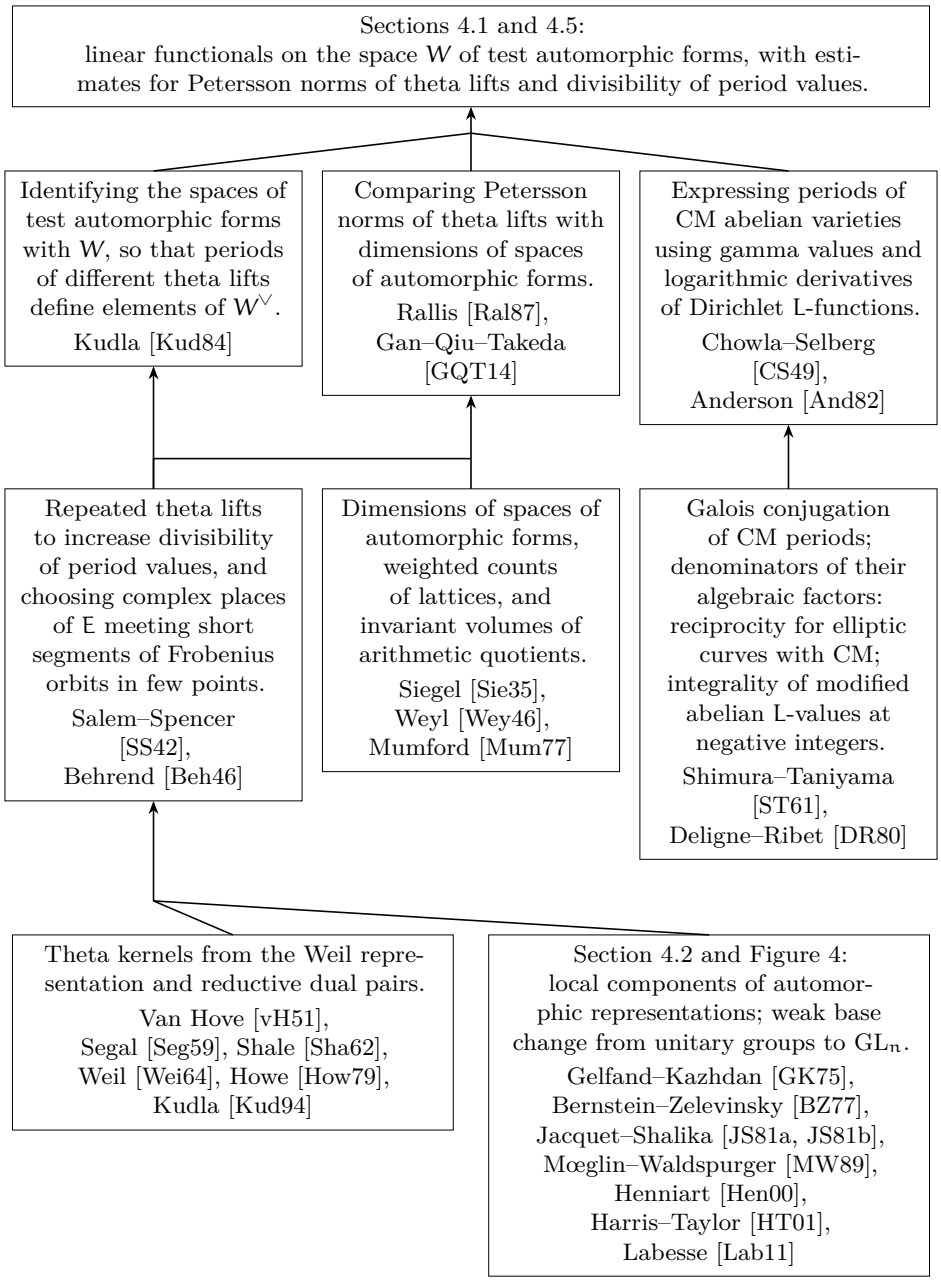


FIGURE 3. Theta lifts and CM periods in [Alp26], with selected historical references.

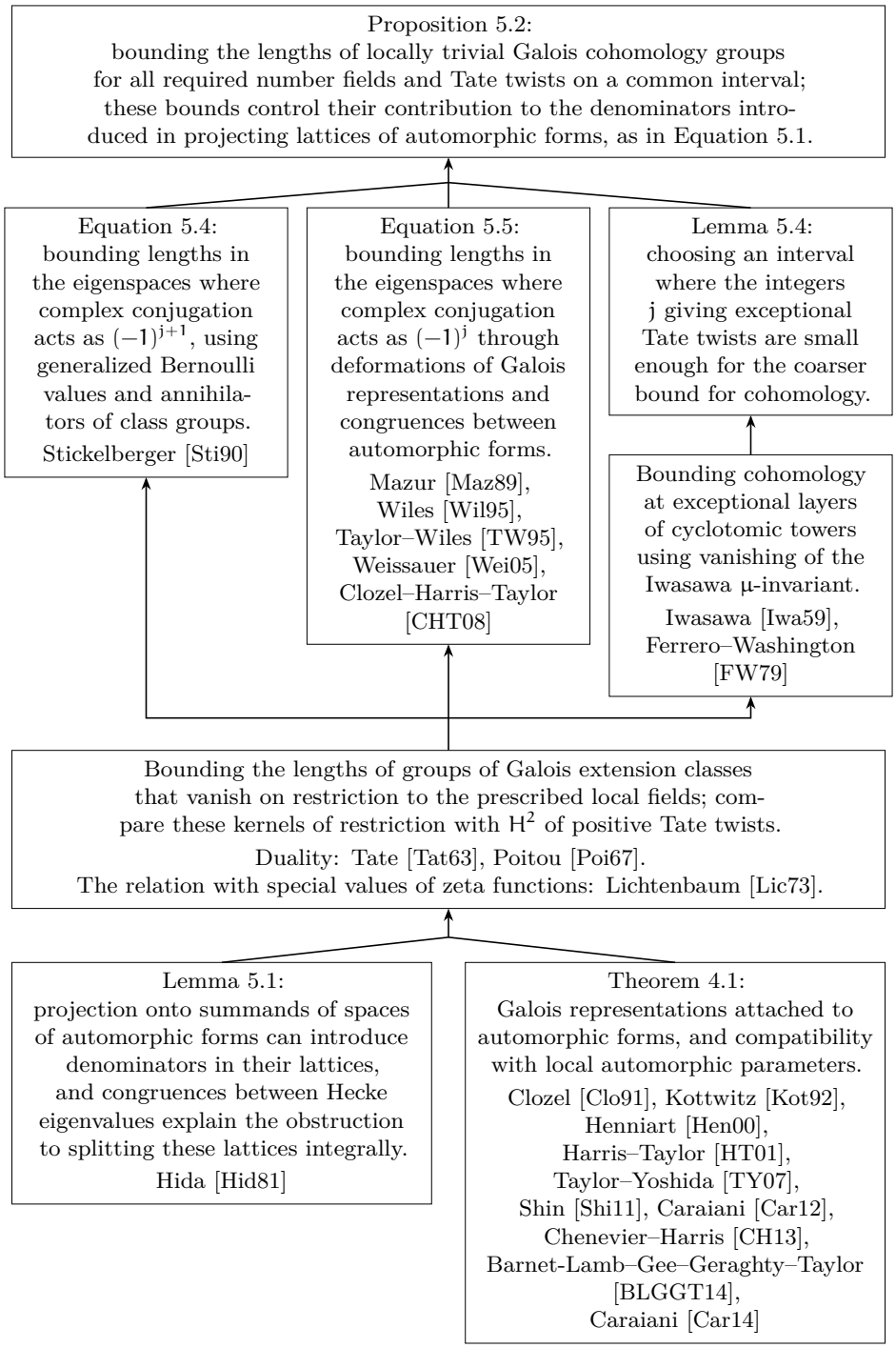


FIGURE 4. Galois cohomology and the denominators introduced in projecting lattices of automorphic forms in [Alp26], with selected historical references.

2. Detecting zeros from sums over prime powers

The estimate for $\mathcal{G}_\psi$ sought in [Alp26] has two limitations: it holds only on certain intervals $[Y, Y^{3/2}]$, with Y arbitrarily large, and even there it gives only an upper bound. To see why this suffices for GRH, consider the effect of a zero of an L-function on the Mellin transform of $\mathcal{G}_\psi$. The positive coefficients of the corrected sums G_χ prevent contributions from a shared zero from canceling. By differentiating the Mellin transform many times, [Alp26] detects a pole using the upper bound for $\mathcal{G}_\psi$ on a single interval. The freedom to choose these intervals is useful in Section 6, where [Alp26] chooses auxiliary number fields and weights subject to several arithmetic conditions.

The linear combination $\mathcal{G}_\psi$ in Equation 1.1 is one instance of a more general situation. Let $\mathcal{X}$ be a finite family of primitive Dirichlet characters, possibly including $\mathbb{1}$. For positive real numbers a_χ, consider the combination

$$M_{\mathcal{X}}(x) := \sum_{\chi \in \mathcal{X}} a_\chi \mathrm{Re}(G_\chi(x)),$$

where the functions G_χ are defined in Section 1.2. Since the coefficients are positive, the multiplicities of a zero in different L-functions will add in the Mellin transform of $M_{\mathcal{X}}$. Proposition 2.1 states the upper bound for $M_{\mathcal{X}}$ that will suffice.

Proposition 2.1. *Let $0 < \vartheta < 1$, $\delta > 0$, and $C \in \mathbb{R}$. Suppose that there are arbitrarily large $Y \geq 2$ such that for every $x \in [Y, Y^{1+\delta}]$,*

$$M_{\mathcal{X}}(x) \leq C x^{\vartheta}.$$

Then every $L(s, \chi)$ with $\chi \in \mathcal{X}$ is nonzero in $\mathrm{Re}(s) > \vartheta$, away from the pole of $\zeta(s)$ at $s = 1$.

There is a useful classical precedent for recovering information about singularities from a condition on signs. Landau [Lan05] proved that a Dirichlet series with nonnegative coefficients has a singularity at its finite real abscissa of convergence. In the present situation, however, the upper bound is available only on certain intervals, so Landau's theorem does not apply directly. To concentrate an integral on one interval where the bound holds, [Alp26] differentiates the Mellin transform of $M_{\mathcal{X}}$ many times. Between these intervals, a rough bound for $M_{\mathcal{X}}$ will suffice.

2.1. How the Mellin transform detects zeros.

The connection between primes and zeros begins with the logarithmic derivative of an L-function: its Dirichlet series involves prime powers, whereas its poles record zeros with their multiplicities. This is the starting point of the explicit formulas to which [Alp26, p. 3] refers. Guinand [Gui48] related weighted sums over prime powers to weighted sums over zeros, under the Riemann hypothesis. Weil [Wei52, Wei72] developed explicit formulas for L-functions over number fields, keeping track of the contributions from every place. For the weighted

sums S_χ used in [Alp26], the identity follows directly by taking a Mellin transform, without assuming the Riemann hypothesis. The factor $x/n - 1$ contributes $1/(s(s-1))$, and the corrections in G_χ remove the singularity at 1.

Indeed, the sum defining S_χ in Section 1.2 can be integrated term by term for $\mathrm{Re}(s) > 2$. Since

$$\int_n^\infty \Big(\frac{x}{n} - 1\Big) x^{-s-1}\,dx = \frac{n^{-s}}{s(s-1)},$$

the transform of the sum over prime powers is

$$\int_1^\infty S_\chi(x) x^{-s-1}\,dx = -\frac{1}{s(s-1)}\frac{L'}{L}(s,\chi).$$

Write $F_\chi(s)$ for the Mellin transform of G_χ. For a nonprincipal character, the term $xL'(1,\chi)/L(1,\chi)$ added in the definition of G_χ therefore gives

$$F_\chi(s) = -\frac{1}{s(s-1)}\frac{L'}{L}(s,\chi) + \frac{1}{s-1}\frac{L'}{L}(1,\chi). \tag{2.1}$$

For $\chi = \mathbb{1}$, the same calculation gives

$$F_{\mathbb{1}}(s) = -\frac{1}{s(s-1)}\frac{\zeta'}{\zeta}(s) - \frac{1}{(s-1)^2} + \frac{1+\gamma}{s-1}. \tag{2.2}$$

It is useful to check the cancellation at $s = 1$ explicitly. For a nonprincipal character, L is holomorphic and nonzero at 1, so the two terms in Equation 2.1 have opposite residues there. In Equation 2.2, the cancellation follows from the expansion $\zeta'/\zeta(s) = -(s-1)^{-1} + \gamma + O(s-1)$. Thus both transforms are holomorphic at 1.

Logarithmic coordinates make the intervals in Proposition 2.1 easier to handle. Put $g(u) := M_{\mathcal{X}}(e^u)$ for $u \geq 0$. The interval $[Y, Y^{1+\delta}]$ becomes $[\log Y, (1+\delta)\log Y]$. For $\mathrm{Re}(s) > 2$, the Mellin transform can therefore be written as the Laplace transform

$$F(s) := \int_0^\infty g(u) e^{-su}\,du = \frac{1}{2}\sum_{\chi\in\mathcal{X}} a_\chi (F_\chi(s) + F_{\overline{\chi}}(s)).$$

The pairing of χ with $\overline{\chi}$ accounts for the real part in the definition of $M_{\mathcal{X}}$. By Equations 2.1 and 2.2, F extends meromorphically to $\mathrm{Re}(s) > 0$. The residues now show why the coefficients were chosen to be positive. Let $m_\chi(\rho)$ denote the multiplicity of a zero ρ of $L(s,\chi)$, with multiplicity 0 if ρ is not a zero. If $\mathrm{Re}(\rho) > 0$ and $L(s,\chi)$ or $L(s,\overline{\chi})$ vanishes at ρ for some $\chi \in \mathcal{X}$, then

$$\begin{aligned} \operatorname*{res}_{s=\rho} F(s) &= \frac{\nu_\rho}{\rho(1-\rho)}, \\ \nu_\rho &:= \frac{1}{2}\sum_{\chi\in\mathcal{X}} a_\chi (m_\chi(\rho) + m_{\overline{\chi}}(\rho)). \end{aligned} \tag{2.3}$$

Notice that ν_ρ is positive even when several characters have a zero at the same point. Thus no such zero disappears from F. Hadamard [Had96, §10]

proved nonvanishing on $\mathrm{Re}(s) = 1$, and the Euler product excludes zeros for $\mathrm{Re}(s) > 1$. The poles of F in $\mathrm{Re}(s) > 0$ therefore form a discrete subset of $0 < \mathrm{Re}(s) < 1$. In particular, the residue at a real pole is positive; this sign is used in the proof of Proposition 2.1.

One further estimate is needed away from the intervals in Proposition 2.1. Since $|\Lambda(n)\chi(n)| \le \log n$, the corrected sums $G_\chi(x)$ are $O(x \log^2(2x))$. Consequently $|g(u)| \ll e^{2u}$ for $u \ge 0$. The proof of Proposition 2.1 uses this rough bound only outside the intervals where the stronger upper bound for $M_{\mathcal{X}}$ is available.

2.2. **Using the estimate on a single interval.** Each zero of $L(s,\chi)$ in the critical strip, for $\chi \in \mathcal{X}$, gives a pole of F. The remaining difficulty is to detect one of these poles from a bound for $M_{\mathcal{X}}$, despite possible cancellation between contributions from different poles. Turán [Tur49, Tur50] addressed precisely this difficulty in studying errors in prime counting, using lower bounds for power sums. The discussion in [Alp26, pp. 3–4] explicitly mentions this method. In [Alp26], repeated differentiation of F serves a related purpose. The point of differentiation is chosen so that one pole is nearer than all the others; its contribution then dominates the derivatives of high order.

There is also a reason to use derivatives of high order in the integral expression for F. Differentiating k times introduces a factor u^k. After division by $k!$, the resulting kernel is a constant multiple of a gamma density. When k is proportional to $\log Y$, most of its mass lies inside the interval where the upper bound for g holds. The following elementary bound makes the contribution from the complement small enough.

Lemma 2.2. *Fix $\delta > 0$. There is a $\sigma > 2$ such that, for every sufficiently large U, there is an integer $k \ge U$ with $k \asymp U$ and*

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

The constants depend only on δ and σ.

Proof. Set $\tau := 1 + \delta/2$. For a sufficiently large $q > 0$, choose $k := \lfloor q\tau U \rfloor$. The measure $q^{k+1}u^k e^{-qu}\, du/k!$ has total mass 1 and exponential moments $(q/(q-t))^{k+1}$ for $t < q$. The exponential Markov inequality bounds its two tails using the function $I(\nu) := \nu - 1 - \log \nu$. The ratios of the two endpoints to the mean tend to $1/\tau < 1$ and $(1+\delta)/\tau > 1$, respectively. Since I is positive at both values, there is a constant $c > 0$, depending only on δ, such that the mass outside $[U, (1+\delta)U]$ is at most $2e^{-ck}$ for sufficiently large U.

Choose q large enough that $q\tau > 1$ and $\log((q+2)/q) < c$, and set $\sigma := q + 2$. The integral in the statement is then at most $2q^{-k-1}e^{-ck} \ll \sigma^{-k}$. This choice ensures that $k \ge U$ and $k \asymp U$ for sufficiently large U. □

The absence of a lower bound for g requires a little care. An upper bound for g alone does not bound the absolute value of its transform. The proof instead combines $|F^{(k)}(\sigma + it)|$ with $(-1)^k F^{(k)}(\sigma)$, for real t. After division

by $k!$, the sum is bounded using only the positive part of g. A pole with real part greater than ϑ would force this combination of derivatives to grow faster than the assumed upper bound for g on those intervals permits.

Proof of Proposition 2.1. Increase C if necessary so that $C \geq 0$. In logarithmic coordinates, the hypothesis on $M_{\mathcal{X}}$ provides arbitrarily large U such that, for every $u \in [U, (1+\delta)U]$,

$$g(u) \leq Ce^{\vartheta u}.$$

Choose σ as in Lemma 2.2, and set $R := \sigma - \vartheta$. The factor $u^k/k!$ from Lemma 2.2 appears after differentiating k times and dividing by $k!$. Absolute convergence justifies differentiation under the integral for F. For real t, write

$$A_k(t) := \frac{(-1)^k}{k!} F^{(k)}(\sigma + it) = \int_0^\infty \frac{u^k}{k!} g(u) e^{-\sigma u} e^{-itu}\, du.$$

Write $g(u)_+ := \max(g(u), 0)$. Since g is real-valued,

$$|A_k(t)| + A_k(0) \leq 2\int_0^\infty \frac{u^k}{k!} g(u)_+ e^{-\sigma u}\, du.$$

On $[U, (1+\delta)U]$, the assumed upper bound for g and positivity of the integrand imply

$$\int_U^{(1+\delta)U} \frac{u^k}{k!} Ce^{-Ru}\, du \leq C\int_0^\infty \frac{u^k}{k!} e^{-Ru}\, du = CR^{-k-1}.$$

On the complement, the rough bound $|g(u)| \ll e^{2u}$ is sufficient by Lemma 2.2. Consequently, along an unbounded sequence of integers k,

$$|A_k(t)| + A_k(0) \ll R^{-k}, \tag{2.4}$$

uniformly for real t. The implied constant is independent of the interval and of t.

It remains to compare Equation 2.4 with the contribution of a pole of F. Suppose that F has a pole with real part greater than ϑ. The argument separates into two cases, according to the poles nearest to the real point σ.

First, suppose that the unique nearest pole is real, say β, and that $d := \sigma - \beta < R$. Choose r with $d < r < R$ such that the disk of radius r centered at σ contains no other pole. By Equation 2.3, the residue c at β is positive. After subtracting $c/(s-\beta)$, the remainder is holomorphic on the disk of radius r, so that Cauchy's estimates imply

$$A_k(0) = \frac{c}{d^{k+1}} + O(r^{-k}).$$

Since $d < R$, the resulting lower bound for $A_k(0)$ contradicts Equation 2.4 at $t = 0$.

In every other case, there is a real t and a unique pole ρ nearest to $z := \sigma + it$ such that, on writing $d := |z - \rho|$,

$$d < R,$$
$$d < \operatorname{dist}(\sigma, \{\text{poles of } F\}).$$

To see this, suppose first that a nonreal pole is nearest to σ at distance less than R. Move the center a small distance vertically towards this pole, thereby decreasing its distance to the center. If no pole has distance less than R from σ, instead move the center to the imaginary part of a pole with real part greater than ϑ. In either case, the displayed inequalities are strict and persist under a further sufficiently small perturbation. The perturbation can be chosen so that the nearest pole is unique. Indeed, only finitely many poles are relevant in the compact region under consideration, and equality of the distances to two different poles excludes at most one point of the vertical line. Poles with the same imaginary part cannot be equidistant from points on $\operatorname{Re}(s) = \sigma$, since their real parts are less than σ.

Choose $r > d$ smaller than R, the distance from σ to every pole, and the distance from z to every pole other than ρ. Let $c \neq 0$ be the residue at ρ. Subtracting its principal part makes Cauchy's estimates applicable at z; at σ, the function F is already holomorphic on the disk of radius r. Consequently,

$$A_k(t) = \frac{c}{(z-\rho)^{k+1}} + O(r^{-k}),$$
$$A_k(0) = O(r^{-k}).$$

Hence

$$|A_k(t)| + A_k(0) \geq |c| d^{-k-1} - O(r^{-k}),$$

which again contradicts Equation 2.4. Thus F has no pole in $\operatorname{Re}(s) > \vartheta$. By Equation 2.3, $L(s, \chi)$ has no zero in $\operatorname{Re}(s) > \vartheta$ for any $\chi \in \mathcal{X}$. □

2.3. **Excluding zeros on both sides of the critical line.** To apply Proposition 2.1 to $\mathcal{G}_\psi$, recall that Equation 1.1 expresses it as a positive linear combination of the real parts of the corrected sums G_χ, including G_ψ. Thus the upper bound for $\mathcal{G}_\psi$ sought in [Alp26] rules out a zero of $L(s, \psi)$ to the right of the critical line. The functional equation then treats the other half of the strip. For completeness, the following deduction also includes the passage from primitive characters to all Dirichlet characters.

Proposition 2.3. *If Equation 1.2 holds for every primitive character ψ and every $\eta > 0$, then GRH holds for all Dirichlet L-functions.*

Proof. Suppose that $L(s, \psi) = 0$ with $\operatorname{Re}(s) > 1/2$. Choose ϑ with $1/2 < \vartheta < \min(\operatorname{Re}(s), 1)$, and set $\eta := \vartheta - 1/2$. By Proposition 2.1, with $\delta = 1/2$, the estimate excludes this zero. Thus every primitive Dirichlet L-function is nonzero to the right of the critical line.

To treat the other half of the critical strip, let χ be primitive of conductor q, and let $\kappa_\chi \in \{0,1\}$ satisfy $\chi(-1) = (-1)^{\kappa_\chi}$. The completed function

$$\Lambda(s,\chi) := \left(\frac{q}{\pi}\right)^{(s+\kappa_\chi)/2}\Gamma\left(\frac{s+\kappa_\chi}{2}\right)L(s,\chi)$$

satisfies the functional equation $\Lambda(s,\chi) = \varepsilon_\chi\Lambda(1-s,\overline{\chi})$, where $|\varepsilon_\chi| = 1$. Since the gamma factors have neither poles nor zeros in $0 < \mathrm{Re}(s) < 1$, a zero to the left of the critical line would give a zero to its right for $\overline{\chi}$. Thus every zero in the critical strip has real part $1/2$.

Finally, let χ be any Dirichlet character modulo N, induced by a primitive character χ^* of conductor q. Then

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

For $\mathrm{Re}(s) > 0$, each added factor is nonzero because $|\chi^*(p)p^{-s}| < 1$. The conclusion follows for all Dirichlet characters, including the Riemann hypothesis in conductor 1. □

3. Recovering sums over prime powers from divisibility of periods

The criterion proved in Section 2 turns upper bounds for $\mathcal{G}_\psi$ into nonvanishing of Dirichlet L-functions. To obtain the sums over prime powers that occur in $\mathcal{G}_\psi$, [Alp26] measures the divisibility of automorphic period values at finite primes. Character orthogonality provides the connection: just as Dirichlet [Dir37] separated residue classes using characters, the argument here expresses a condition on Frobenius as a sum of character values. Proposition 3.1 makes this relation explicit, including the main term subtracted to obtain the functions G_χ. The choices of weights and averages over characters then recover the combination $\mathcal{G}_\psi$.

3.1. **Expressing the Frobenius condition with characters.** A quadratic character distinguishes the coset in an abelian Galois group whose Frobenius elements contribute to the divisibility of period values in [Alp26]. The identity can be stated using only data from a finite group, before choosing the abelian extension of $\mathbb{Q}$. Let Γ be a finite abelian group, written additively, and let $\kappa : \Gamma \longrightarrow \{\pm 1\}$ be a nontrivial quadratic character. Set $\Phi := \ker\kappa$ and $f := |\Phi|$, so that $|\Gamma| = 2f$, and write $\widehat{\Gamma}$ for the character group of Γ.

There is a small complication at ramified primes: a Frobenius element is defined only modulo the inertia subgroup. Averaging the character values over inertia removes the choice of a representative and produces the coefficients of the associated primitive Dirichlet L-function. To make this explicit, for each prime p, take a subgroup $I_p \subseteq \Gamma$ and a class $F_p \in \Gamma/I_p$. To evaluate a character without choosing a preferred lift of F_p, for each $j \geq 1$ and $\chi \in \widehat{\Gamma}$,

consider the average

$$a_\chi(p^j) := \frac{1}{|I_p|}\sum_{h\in I_p}\chi(jF_p+h).$$

If χ is trivial on I_p, the average equals $\chi(F_p)^j$; otherwise it is zero. In particular, it is independent of the chosen lift. Extend a_χ multiplicatively, with $a_\chi(1) := 1$. When Γ, I_p, and F_p come from an abelian extension of $\mathbb{Q}$, the numbers $a_\chi(n)$ are the coefficients of the primitive Dirichlet L-function associated to χ, including at ramified primes.

Conversely, the Frobenius condition can be recovered by summing over characters. For $g \in \Gamma$, the orthogonality relation is

$$\frac{1}{f}\sum_{\chi\in\widehat{\Gamma}}\overline{\chi(g)}a_\chi(p^j) = \frac{2}{|I_p|}\mathbf{1}_{\{g-jF_p\in I_p\}}. \tag{3.1}$$

Here $\mathbf{1}$ is 1 if the displayed condition holds, and 0 otherwise. Indeed, the inner sum over characters is $|\Gamma|$ precisely when $jF_p + h = g$, and is zero otherwise. Although individual characters may take complex values, their sum on the left is therefore real and nonnegative. This is the identity that converts a count at a prime into a sum of character values.

In the lower bounds for valuations of period values in [Alp26], for a positive parameter a, a prime power p^j contributes only while a/p^j exceeds a prescribed weight. This explains the positive part in the sum below. The weights are indexed by Φ, whereas the Frobenius terms run over $\Gamma - \Phi$. To match the indices, choose $c \in \Gamma - \Phi$ and $t \in \Phi$, and use the bijection $\rho_t(g) := c + g + t$ from $\Gamma - \Phi$ to Φ. For positive weights w_ρ, indexed by $\rho \in \Phi$, the relevant sum is

$$\mathcal{A}(a,t,w) := \frac{1}{f}\sum_{g\notin\Phi}\sum_{\chi\in\widehat{\Gamma}}\overline{\chi(g)}\sum_{n\ge1}\Lambda(n)a_\chi(n)\Big(\frac{a}{n}-w_{\rho_t(g)}\Big)_+, \tag{3.2}$$

where $u_+ := \max(u,0)$. Each sum over n is finite. Equation 3.1 expresses the same quantity as

$$\mathcal{A}(a,t,w) = \sum_p \log p\sum_{j\ge1}\frac{2}{|I_p|}\sum_{\substack{g\notin\Phi\\ g-jF_p\in I_p}}\left(\frac{a}{p^j}-w_{\rho_t(g)}\right)_+.$$

In particular, $\mathcal{A}(a,t,w) \ge 0$. The second expression keeps the contributions to divisibility of period values visible at each prime, while Equation 3.2 expresses the same sum in terms of the characters of Dirichlet L-functions.

3.2. **Removing the main contribution.** Expressing $\mathcal{A}(a,t,w)$ in terms of G_χ also introduces the correction terms from their definitions. Their sum is easy to compute once the logarithmic derivatives at 1 are treated as constants. For each nontrivial χ, let ℓ_χ be a complex number; in the

application it is $L'(1,\chi)/L(1,\chi)$ for the corresponding primitive character. Thus the corrected sums are

$$G_\chi(y) := \sum_{n\geq 1} \Lambda(n)a_\chi(n)(y/n-1)_+ + y\ell_\chi,$$
$$G_{\mathbb{1}}(y) := \sum_{n\geq 1} \Lambda(n)(y/n-1)_+ - y\log y + (1+\gamma)y.$$

For the arithmetic data, these expressions extend the definitions in Section 1.2 to every $y > 0$. Assume that ℓ_κ is real, as it is in the application. Since $\mathcal{A}(a,t,w)$ sums over the nontrivial coset of $\ker\kappa$, only the corrections for the trivial character and κ survive.

The surviving corrections form the real main term

$$\mathcal{M}(a,w) := a\left(\log a - 1 - \gamma + \ell_\kappa - \frac{1}{f}\sum_{\rho\in\Phi}\log w_\rho\right). \tag{3.3}$$

After removing this term, the remaining expression involves the corrected sums G_χ, which in the arithmetic application agree with those in Section 1.2. The following identity makes this correspondence explicit.

Proposition 3.1. *For $\mathcal{A}$ and $\mathcal{M}$ as in Equations 3.2 and 3.3,*

$$\mathcal{A}(a,t,w) - \mathcal{M}(a,w) = \frac{1}{f}\sum_{g\notin\Phi}\sum_{\chi\in\widehat{\Gamma}}\overline{\chi(g)}w_{\rho_t(g)}G_\chi\left(\frac{a}{w_{\rho_t(g)}}\right). \tag{3.4}$$

Proof. Since $w > 0$, $(a/n - w)_+ = w(a/(wn)-1)_+$. Thus each sum over prime powers in Equation 3.2 can be expressed in terms of G_χ. For $\chi \neq \mathbb{1}$, the correction is $-a\ell_\chi$; for the trivial character, it is

$$a\log a - a\log w_{\rho_t(g)} - (1+\gamma)a.$$

To see which corrections survive, sum over the nontrivial coset:

$$\frac{1}{f}\sum_{g\notin\Phi}\overline{\chi(g)} = \begin{cases} 1 & \text{if } \chi = \mathbb{1},\\ -1 & \text{if } \chi = \kappa,\\ 0 & \text{otherwise.}\end{cases}$$

Only $\chi = \kappa$ therefore contributes a nontrivial correction, namely $a\ell_\kappa$. Moreover, since ρ_t is a bijection, the terms involving the logarithms of the weights sum to $-(a/f)\sum_{\rho\in\Phi}\log w_\rho$. These are exactly the terms subtracted in Equation 3.3. □

Only logarithmic derivatives at 1 occur in this identity.

Remark 3.2. The correction uses logarithmic derivatives only at 1, where the nonprincipal Dirichlet L-functions are already known to be nonzero. Thus Equation 3.4 does not depend on nonvanishing inside the critical strip.

The problem is therefore to bound the excess $\mathcal{A}(a,t,w) - \mathcal{M}(a,w)$. Before using this identity to estimate $\mathcal{G}_\psi$ in Equation 1.1, [Alp26] selects the characters and the arguments at which G_χ occurs.

3.3. **Retaining the required characters and arguments.** The identity in Equation 3.4 sums over all characters of Γ, while $\mathcal{G}_\psi$ involves only the primitive characters inducing powers of ψ. Moreover, the arguments of the functions G_χ in that identity depend on the weights. To specify the arguments, [Alp26] uses weights of two different sizes. At selected complex places of an abelian CM field E, the argument a/w of G_χ is close to x; at the other complex places it is much larger, so estimates for L-functions control the contribution of $G_\chi(a/w)$. Averaging over a subgroup of the Galois group retains the required characters. To describe the weights and the average explicitly, first identify the complex places with elements of a finite group.

Here $\Gamma = (\mathbb{Z}/2\mathbb{Z})^2 \times H$ for a finite abelian group H, and Φ is the kernel of the first sign character. In the application, Γ is identified with $\mathrm{Gal}(E/\mathbb{Q})$. After fixing one complex embedding of E, the elements of Γ label all embeddings $E \to \mathbb{C}$. The complex places of E are pairs of conjugate embeddings, so their labels lie in the quotient of Γ by complex conjugation. In the arithmetic application, κ is odd, so Φ contains one label from each conjugate pair. Thus these representatives form a CM type of E, meaning a choice of one embedding from each complex-conjugate pair. They label the complex places where the weights are prescribed. At unramified finite primes, Frobenius acts on the complex places. In the chosen representatives, this action includes complex conjugation whenever the transported embedding lies outside the CM type.

For each translate $u \in H$, [Alp26] chooses weights $w_{u,j,\rho}$, with $1 \le j \le m$ and $\rho \in \Phi$, for the real local components of the automorphic representations in the m successive theta lifts. These integers specify the blocks of Harish–Chandra parameters added by the lifts, as in Section 4.1. The sets T and T_2 label two selected sets of complex places of E. Each $t \in T$ has a positive parameter a_t. The construction based on the method of Behrend [Beh46] concerns T_2. For a small $\varepsilon > 0$, set $b := \exp(x^\varepsilon)$ and $b' := xb/\exp(x^{\varepsilon/3})$. The parameters satisfy $a_t/b = x + O(1)$. At the places labeled by T_2, the weights lie between b and $b + \eta_w + 1$, where η_w is a tolerance for the weights, much smaller than b. Thus a_t/w is close to x at these places. At every remaining complex place, including those labeled by T, the parameter bounds ensure $w \le 2b'$ for sufficiently large x. Since $a_t/b = x + O(1)$, it follows that $a_t/w \gg \exp(x^{\varepsilon/3})$. This much larger argument is the reason for the second choice of scale: estimates for L-functions in regions without zeros control the terms $G_\chi(a_t/w)$ at these places. The choice of parameters accounts for the dependence of those estimates on the characters and conductors. The terms whose arguments are close to x are the ones retained.

It remains to select the characters. For a subgroup $H_0 \subseteq H$, orthogonality gives

$$\sum_{u \in H_0} \chi(u) = \begin{cases} |H_0| & \text{if } \chi|_{H_0} = \mathbb{1}, \\ 0 & \text{otherwise,} \end{cases}$$

so that averaging over H_0 retains precisely the characters trivial on H_0. In [Alp26], the subgroup H_0 and the selected complex places are chosen to retain the primitive characters inducing powers of ψ.

The other characters retained by the average still contribute, and the arguments close to χ have yet to be replaced by χ itself. In Section 6, [Alp26] bounds both errors as the fields, conductors, and weights vary. Equation 3.4 identifies the quantity to which the estimates for periods will be applied: the excess of $\mathcal{A}(a, t, w)$ over $\mathcal{M}(a, w)$. In Section 4, [Alp26] relates $\mathcal{A}(a, t, w)$ to divisibility of automorphic period values. Comparing these valuations with the Petersson norms of theta lifts and the dimensions of spaces of automorphic forms then bounds the excess.

4. Bounding sums over prime powers with automorphic periods

After the character calculation in Section 3, the remaining problem is to bound the arithmetic sum in Equation 3.2 after subtracting its main term. At first, this may seem rather far from automorphic forms. The connection in [Alp26] comes from period pairings realized over a coefficient number field K. The valuations of their values at finite places of K contain the summands of $\mathcal{A}(a, t, w)$, while estimates at the archimedean places limit how divisible the period values can be, as in the product formula. The comparison is useful only if its constants also recover $\mathcal{M}(a, w)$: the desired bound concerns $\mathcal{A}(a, t, w) - \mathcal{M}(a, w)$.

There are two difficulties in making this comparison precise. To obtain enough divisibility of period values relative to the errors in their archimedean estimates, [Alp26] repeats the theta correspondence. It also varies the spaces of automorphic forms being paired with the fixed test space W. Pairing with elements of W gives period maps into the dual lattice C introduced in Section 1.3. If their images were confined to a small subspace of $C \otimes K$, they would say little about its covolume. The linear algebra in Section 4.4 is what allows [Alp26] to use enough of these images to bound the mean slope of C. Figures 2 and 3 show how the construction of the periods and the estimates for lattices meet in this comparison.

4.1. **Increasing divisibility of period values by repeated theta lifts.**

Before describing the periods, it helps to recall where the theta correspondence came from. Its underlying representations grew out of questions in quantum mechanics: a linear change of position and momentum preserving the commutation relations should, when possible, be realized by a unitary operator. Van Hove [vH51] had already studied this problem for canonical transformations. In the real case, Segal [Seg59] recognized the resulting projective representation, and Shale [Sha62] developed the theory in his study of the linear symmetries of free boson fields. In his later account, Segal [Seg63, pp. 31–33] explained the relation to quantum mechanics and to van Hove's work.

The passage to number theory required an extension to local fields and adeles. In making this extension, Weil [Wei64, pp. 143–144] sought to understand the modularity of theta series and the appearance of symplectic groups in Siegel's theory of quadratic forms. He explicitly credited Segal [Seg63] for the principle of his existence proof and for the starting point in locally compact abelian groups. Howe [How79] then organized the representation theory around reductive dual pairs: two reductive groups acting as mutual centralizers in a symplectic group. Restricting the Weil representation to such a pair made it possible to pass between representations of its two members.

The unitary groups and the theta integral. Retain the auxiliary CM field E from Section 3, and write E^+ for its maximal totally real subfield. For a nondegenerate Hermitian space (V, h_V) over E, the group $U(V)$ is defined over E^+: its E^+-points are the E-linear automorphisms g satisfying $h_V(gv, gw) = h_V(v, w)$ for every $v, w \in V$. At a real place τ of E^+, a form of signature (p_τ, q_τ) gives the real group $U(p_\tau, q_\tau)$. Thus a totally definite Hermitian space gives a compact unitary group at every real place.

To write a theta kernel, first take a Hermitian space V and a skew-Hermitian space V' over E. The trace of the tensor product of their forms makes $V \otimes_E V'$ a symplectic space over E^+, on which $U(V) \times U(V')$ acts as a reductive dual pair. Multiplying a Hermitian form by an element $\delta \in E^\times$ with $\bar{\delta} = -\delta$ makes it skew-Hermitian without changing its unitary group. This explains how the usual convention with opposite signs applies when the groups are described using Hermitian spaces. Gan [Gan23, § 1.7 and § 2.2] gives an account of these conventions.

Write $\mathbb{A}_F$ for the ring of adeles of a number field F. The Weil representation also depends on a nontrivial additive character ψ_0 of $E^+\backslash\mathbb{A}_{E^+}$ and two unitary Hecke characters η_V and $\eta_{V'}$ of $E^\times\backslash\mathbb{A}_E^\times$, with

$$\eta_V|_{\mathbb{A}_{E^+}^\times} = \omega_{E/E^+}^{\dim_E V},$$
$$\eta_{V'}|_{\mathbb{A}_{E^+}^\times} = \omega_{E/E^+}^{\dim_E V'}.$$

Here ω_{E/E^+} is the quadratic character associated to E/E^+. Kudla [Kud94] constructed the splittings of the metaplectic cover of the symplectic group that these characters specify. Pulling back the Weil representation along the splitting gives a representation ω of $U(V)(\mathbb{A}_{E^+}) \times U(V')(\mathbb{A}_{E^+})$.

Choose a polarization $V \otimes_E V' = \mathbb{X} \oplus \mathbb{Y}$ into maximal isotropic subspaces. In the corresponding Schrödinger model, ω acts on $\mathcal{S}(\mathbb{X}(\mathbb{A}_{E^+}))$. For a Schwartz function φ in this space, the theta kernel is

$$\theta_\varphi(g, h) = \sum_{x \in \mathbb{X}(E^+)} (\omega(g, h)\varphi)(x).$$

For an algebraic group J over E^+, write $[J] := J(E^+)\backslash J(\mathbb{A}_{E^+})$. For a cusp form f on $U(V)$, the convention linear in f gives the theta integral

$$\Theta_\varphi f(h) = c_V \int_{[U(V)]} \theta_\varphi(g,h) f(g)\, dg.$$

The factor $c_V > 0$ records the normalization of the Petersson pairing. This is the convention used for the successive linear maps in [Alp26]. In the convention of Gan, Qiu, and Takeda [GQT14, §11.2], the integral contains $\bar{f}$; replacing f by its conjugate passes between the two conventions. For cuspidal f the integral converges absolutely and defines an automorphic form on $U(V')$.

Which groups occur in the successive lifts. The construction in [Alp26] uses two sequences of Hermitian spaces X_j and Y_j, which it is useful to keep separate. Write $k := \dim_E X_0$, and retain m for the number of lifts in each sequence. Both k and m are even. Put $r := k + m - 1$, and choose an odd integer $h \geq r + 5$ to govern the lengths of the blocks added by theta lifting. For $0 \leq j \leq m$, the dimensions are

$$\dim_E X_j = k + j(h+1),$$
$$\dim_E Y_j = r + jh.$$

At every real place, X_j is positive definite and Y_j has signature

$$\left(r + jh - \left\lfloor \frac{m-j}{2} \right\rfloor, \left\lfloor \frac{m-j}{2} \right\rfloor\right).$$

Thus Y_0 has signature $(k + m/2 - 1, m/2)$, whereas Y_m is totally definite. The isometric embeddings are $X_j \hookrightarrow Y_j$ for $j < m$ and $Y_m \hookrightarrow X_m$. Notice that the direction reverses at the last stage: Y_m has codimension one in X_m.

The Harish–Chandra parameters of the representations at real places make the increase in dimension quite concrete. Label these places by the embeddings τ in the chosen CM type of E. The passage from X_j to X_{j+1} retains the old parameters and adds the block

$$-e_{j,\tau} + \frac{h}{2}, \quad -e_{j,\tau} + \frac{h}{2} - 1, \quad \ldots, \quad -e_{j,\tau} - \frac{h}{2},$$

where $e_{j,\tau}$ is a prescribed integer. The parameters on Y_j acquire blocks of length h, together with the prescribed character twists. The initial representations, Hecke characters, levels, and integral lattices are chosen along with these parameters. In particular, the initial source spaces satisfy additional conditions at an auxiliary rational prime. These define the source subspaces used for the period maps; the conditions are written explicitly in Section 4.4.

Notice why the theta lift is repeated. Each additional lift increases the divisibility of the resulting period values, so the contribution to their valuations grows linearly with m. The analogy proposed in [Alp26] is with the method of Stepanov [Ste69]. To bound the number of points on a curve

over a finite field, Stepanov constructed a nonzero auxiliary polynomial that vanished to high order at those points. Its degree then limited the total prescribed vanishing. Schmidt [Sch73] extended this approach to general absolutely irreducible plane curves. Here m plays the part of the order of vanishing. But taking more lifts also enlarges the spaces of automorphic forms. The choice of m therefore depends on the errors in estimating their norms and dimensions, as well as on the gain in divisibility of the period values.

From Bessel periods to linear functionals. The embeddings between X_j and Y_j allow their automorphic forms to be paired. The initial pairing is a Bessel period for $U(X_0) \subset U(Y_0)$. The orthogonal complement of X_0 in Y_0 consists of $(m-2)/2$ hyperbolic planes and a nonisotropic line. Choose a maximal isotropic flag in this complement, and let N be the unipotent radical of its stabilizer in $U(Y_0)$. Choose a $U(X_0)$-invariant Bessel character ν of $N(\mathbb{A}_{E^+})$, trivial on $N(E^+)$, from the flag and the nonisotropic line as in Gan, Gross, and Prasad [GGP12, § 12 and § 23]. For cusp forms f on $U(X_0)$ and g on $U(Y_0)$, the period is

$$\mathcal{B}_\nu(f,g) = c_{X_0} \int_{[U(X_0)]} \int_{[N]} \overline{f(h)} g(nh) \overline{\nu(n)}\, dn\, dh.$$

Here h is regarded as an element of $U(Y_0)$ through the chosen embedding, and $[N]$ has volume one. The inner integral extracts a Fourier coefficient of g; the outer integral pairs its restriction to $U(X_0)$ with f. Thus the period is linear in g and conjugate-linear in f.

The point of the two sequences of lifts is that these pairings can be followed from one stage to the next. Kudla [Kud84] proved seesaw identities by integrating a theta kernel in two orders. In [Alp26], the resulting identities relate the successive periods to the final restriction pairing for $U(Y_m) \subset U(X_m)$, where the unipotent integral has disappeared because the codimension is one. At the intermediate stages, the difference $\dim_E Y_j - \dim_E X_j = m-1-j$ alternates in parity. The periods are Bessel periods for even j and Fourier–Jacobi periods for odd j; the latter include an additional theta function in the integrand, as in Gan, Gross, and Prasad [GGP12, §23]. The bilinear period used to compare lattices is obtained by conjugating the first argument of that sesquilinear pairing and then imposing the prescribed normalizations.

For the comparison of lattices, the direction of the lifts on the test automorphic forms matters as well. The successive maps from the spaces on Y_j down to the space on Y_0 allow the periods to be regarded as functionals on a fixed space of test automorphic forms. Its rational realization is the space W of automorphic forms on $U(Y_0)$ introduced in Section 1.3, with the chosen level and representations at infinity. Let α again index the choices of source weights. The tuples of auxiliary Hecke characters used in these lifts vary independently of the weights indexed by α. After the chosen identifications and normalizations, [Alp26] obtains K-bilinear periods $P_\alpha(f,g)$ with $g \in W$.

Let M_α be the chosen Hermitian lattice in the initial source subspace on $U(X_0)$. The lattice N_α lies in this same space: its integral structure records the conditions imposed on the successive lifts, together with the required rescaling by a fractional ideal. Retain C for the Hermitian dual of the chosen lattice in W. The periods then define K-linear maps

$$b_\alpha : N_\alpha \otimes K \longrightarrow C \otimes K = W^\vee,$$

with $b_\alpha(f)(g) = P_\alpha(f, g)$ for $g \in W$. An element $b_\alpha(f)$ records the periods obtained by pairing f with every test automorphic form in W. The lower bounds for the valuations of period values at finite primes ensure that $b_\alpha(N_\alpha) \subseteq C$. Thus b_α is an integral map of lattices. Changing the lattices introduces denominators, however, and [Alp26] bounds these using the Galois cohomology in Section 5.

To compare the sizes of b_α with the divisibility of their values, [Alp26] estimates Petersson norms of the theta lifts at the archimedean places of K. Rallis [Ral87] developed formulas expressing such norms through special values of L-functions and local integrals. Gan, Qiu, and Takeda [GQT14] treated the unitary case in their work on the regularized Siegel–Weil formula.

Here it is useful to write the identity used in [Alp26] on a subspace with fixed Hecke eigenvalues. Fix a complex embedding of K. Let B be a simultaneous Hecke eigensystem occurring in $W \otimes_K \mathbb{C}$, and write $W[B]$ for its eigenspace. At an unramified finite prime $\mathfrak{p}$ of E, let $B_\mathfrak{p}$ be its multiset of r Satake parameters, in the arithmetic normalization of [Alp26]. For a tuple $\mu = (\mu_1, \dots, \mu_m)$ of auxiliary Hecke characters, let S_μ be the finite set of primes omitted from the Euler products. With $q_\mathfrak{p}$ the norm of $\mathfrak{p}$ and $\mu_j(\mathfrak{p})$ its unramified character value, put

$$L^{S_\mu}(s, B \otimes \mu_j^{-1}) = \prod_{\mathfrak{p} \notin S_\mu} \prod_{\beta \in B_\mathfrak{p}} (1 - \beta \mu_j(\mathfrak{p})^{-1} q_\mathfrak{p}^{-s})^{-1}.$$

The parameters satisfy $|\beta| \le q_\mathfrak{p}^{(r-1)/2}$. Since $h \ge r+5$, the point $s = (h+1)/2$ lies in the region of absolute convergence.

Write Θ_μ for the composite map from the selected space of automorphic forms on $U(Y_m)$ down to $W \otimes_K \mathbb{C}$, and use the Petersson inner products to define its adjoint. With the prescribed normalizations, [Alp26] proves

$$\Theta_\mu \Theta_\mu^* \big|_{W[B]} = \kappa_0 \prod_{j=1}^{m} L^{S_\mu}\left(\frac{h+1}{2}, B \otimes \mu_j^{-1}\right) \mathrm{Id}_{W[B]}.$$

The positive factor κ_0 is independent of the tuple μ and the Hecke eigensystem B. Its value retains the archimedean factors and the chosen normalizations of measures and levels. Thus the adjoint of Θ_μ multiplies squared norms on $W[B]$ by the displayed scalar. This is the information about the periods that enters the comparison of lattices.

The constants in the identity are part of the reason that this estimate can be used for the corrected sums over prime powers. Together with the dimensions of the spaces of automorphic forms and the normalization of

the periods of abelian varieties with complex multiplication, they recover the main term $\mathcal{M}(a, w)$ that was subtracted in Section 3. The equality responsible for this agreement is written out in Section 4.5.

4.2. **From Hecke eigenvalues to Galois representations.** To study congruences between the automorphic representations, [Alp26] associates Galois representations to their Hecke eigenvalues. The passage involves two steps, and keeping them separate helps explain the Tate twists that occur later. First, base change transfers the representations of the definite unitary groups to general linear groups over E. The resulting representations need not be cuspidal. Instead, [Alp26] keeps track of pairs (Π_a, d_a), where Π_a is a unitary conjugate self-dual cuspidal representation of $\mathrm{GL}_{n_a}(\mathbb{A}_E)$ and $d_a \geq 1$. Here conjugate self-duality means $\Pi_a^c \simeq \Pi_a^\vee$, where c is complex conjugation in $\mathrm{Gal}(E/E^+)$ and the superscript $\vee$ denotes the contragredient. For a unitary group in n variables, their dimensions satisfy $\sum_a n_a d_a = n$. At a finite place w of E, write rec_{E_w} for the local Langlands parameter and $|\cdot|$ for the normalized local absolute value. At an unramified place, the parameter of the base change is

$$\bigoplus_a \bigoplus_{j=0}^{d_a-1} \mathrm{rec}_{E_w}(\Pi_{a,w}) \otimes |\cdot|^{(d_a-1)/2-j}.$$

Thus even a single cuspidal representation can contribute several different powers of the norm character. Mœglin and Waldspurger [MW89] classified the discrete spectrum of the general linear groups in these terms. The weak base change from unitary groups used here is of the kind established by Labesse [Lab11, Corollary 5.3].

The second step attaches a Galois representation to a cuspidal constituent, after the appropriate norm twist. Here are the conventions, since a change of normalization would change both the Frobenius eigenvalues and the Tate twists in Section 5. Fix a rational prime ℓ. Write $G_E := \mathrm{Gal}(\overline{E}/E)$, fix an isomorphism $\iota : \overline{\mathbb{Q}}_\ell \simeq \mathbb{C}$, and normalize local reciprocity to send a uniformizer to geometric Frobenius. For Hodge–Tate weights, use the convention in which the cyclotomic character ϵ_ℓ has weight -1; a Tate twist $V(j)$ means $V \otimes \epsilon_\ell^j$. Suppose (Π, ω) is regular algebraic polarized cuspidal: Π is cuspidal on $\mathrm{GL}_n(\mathbb{A}_E)$, ω is a Hecke character of E^+ with the same sign at every real place, and

$$\Pi^c \simeq \Pi^\vee \otimes (\omega \circ N_{E/E^+} \circ \det).$$

For each embedding $\sigma : E \hookrightarrow \mathbb{C}$, write $a_{\sigma,1} \geq \cdots \geq a_{\sigma,n}$ for the highest weights of the algebraic representation whose dual has the infinitesimal character of Π_∞. Here Π_∞ denotes the tensor product of the archimedean local components of Π. For a local Galois representation, WD denotes its associated Weil–Deligne representation; the superscript F-ss means Frobenius semisimplification, which retains the monodromy operator. Write G_{E_w} for the absolute Galois group of E_w and HT_τ for the multiset of Hodge–Tate weights at τ.

The construction and compatibility theorems can then be stated as follows, combining recent progress in the Langlands program. We use conventions that agree with Barnet-Lamb, Gee, Geraghty, and Taylor [BLGGT14, §1].

Theorem 4.1. *There is a continuous semisimple representation* $r_{\ell,\iota}(\Pi) : G_E \longrightarrow \mathrm{GL}_n(\overline{\mathbb{Q}}_\ell)$ *such that, at every finite place* w*,*

$$\iota\,\mathrm{WD}\big(r_{\ell,\iota}(\Pi)|_{G_{E_w}}\big)^{\text{F-ss}} \simeq \mathrm{rec}_{E_w}\Big(\Pi_w \otimes |\det|^{(1-n)/2}\Big). \tag{4.1}$$

At $w \mid \ell$*, the representation is de Rham, with*

$$\mathrm{HT}_\tau(r_{\ell,\iota}(\Pi)) = \{a_{\iota\tau,j} + n - j : 1 \le j \le n\}$$

for each $\tau : E \hookrightarrow \overline{\mathbb{Q}}_\ell$ *inducing* w*. If* Π_w *is unramified, the representation is crystalline at* w*.*

The existence of $r_{\ell,\iota}(\Pi)$ grew out of the cohomology of Shimura varieties: Clozel [Clo91], Kottwitz [Kot92], and Harris and Taylor [HT01] obtained the construction under additional hypotheses at a finite place; Shin [Shi11] then removed that restriction at a finite place, subject to an extra condition on the weights in even dimension; and then Chenevier and Harris [CH13] used deformation to treat the remaining weights and also established the crystalline assertion in Theorem 4.1. These improvements matter here because the dimensions and weights in [Alp26] vary with the parameters.

The identification with local Langlands parameters required further work, which we also briefly summarize. Harris and Taylor [HT01] (and later, Henniart [Hen00] and Scholze [**?**] with different proofs) proved the local correspondence for GL_n. For the Galois representations in Theorem 4.1, compatibility away from ℓ was established under successively more general hypotheses by Taylor and Yoshida [TY07] and Caraiani [Car12]. At places above ℓ, Barnet-Lamb, Gee, Geraghty, and Taylor [BLGGT14] proved compatibility up to semisimplification, and up to Frobenius semisimplification for the weights covered by Shin's construction. Caraiani [Car14, Theorem 1.1] identified the monodromy in the remaining cases.

At a place $w \nmid \ell$ where Π_w is unramified, the content of Equation 4.1 is especially concrete. If $\alpha_{w,1}, \dots, \alpha_{w,n}$ are the Satake parameters of Π_w and q_w is the cardinality of the residue field, then

$$\det(X - \iota r_{\ell,\iota}(\Pi)(\mathrm{Frob}_w)) = \prod_{j=1}^{n}\Big(X - q_w^{(n-1)/2}\alpha_{w,j}\Big).$$

The same normalization explains how [Alp26] chooses the twists of the cuspidal constituents. Suppose Π_0 is conjugate self-dual and the parameters z of its infinitesimal character are distinct at every complex embedding. Choose a real number a_0 for which every $a_0 - z$ is an integer. Applying Theorem 4.1 to $\Pi_0 \otimes |\det|^{(n-1)/2-a_0}$ gives Hodge–Tate weights $a_0 - z$ and unramified Frobenius eigenvalues $q_w^{a_0}\alpha_{w,j}$. The parameter a_0 is chosen to match the integral weights in the construction. This accounts for the shifts when the Galois representations of the different constituents are assembled.

There are two local consequences that [Alp26] needs in addition to the unramified characteristic polynomials. At the selected ramified places, the prescribed parameters have zero monodromy and controlled tame inertia; the compatibility compares the characteristic polynomials of inertia as well. At places above ℓ where the automorphic representation is unramified, crystallinity and the specified Hodge–Tate weights allow the argument to compare integral lattices. For a ramified parameter with zero monodromy, the corresponding comparison uses potential crystallinity. The distinctions are necessary: knowing the Frobenius eigenvalues at almost all places would not supply the local conditions on extension classes used in Section 5.

Recognizing the cuspidal constituents from their local parameters also uses older results. Jacquet and Shalika [JS81a, JS81b] proved the required uniqueness from almost all local components, as well as the bounds $q_w^{-1/2} < |\alpha| < q_w^{1/2}$ for unramified Satake parameters of unitary cuspidal representations. When [Alp26] compares a representation at conjugate places, admissibility from Bernstein and Zelevinsky [BZ77, Theorem 2.5] allows the use of Gelfand and Kazhdan [GK75]: the contragredient is obtained by composing with inverse transpose.

Finally, the integral step needs more than a semisimple representation over $\overline{\mathbb{Q}}_\ell$. After enlarging the coefficient field, [Alp26] chooses Galois-stable lattices and compares their reductions modulo a coefficient prime. The congruences between Hecke eigenvalues then constrain how the residual constituents can be joined by extensions. These extension classes are the objects used to bound the denominators of the projections onto Hecke eigenspaces in Section 5. Notice that the construction here starts with automorphic forms and attaches Galois representations to them; the automorphy lifting arguments in that section start with Galois representations and go in the opposite direction.

4.3. **Obtaining bounds for lattices of automorphic forms.** Estimates for the Petersson norms of individual automorphic forms at the archimedean places of K do not determine the covolume of their integral lattice. To bound these covolumes as the levels and weights vary, [Alp26] uses the geometric realization of automorphic forms as sections of bundles. The argument calls for integral sections with controlled archimedean norms, and for sections that do not vanish at specified points. Figure 2 records the results from algebraic and arithmetic geometry used here.

Sections on compactified moduli spaces. The bundles in [Alp26] are defined on compactified moduli spaces of polarized abelian varieties with prescribed endomorphisms and level structure. Ash, Mumford, Rapoport, and Tai [AMRT75] developed toroidal compactifications of arithmetic quotients over $\mathbb{C}$. The arithmetic questions here require models over rings of integers as well. For Siegel moduli spaces, Faltings and Chai [FC90, Chapters IV–V] developed this theory through degenerations of abelian varieties to semi-abelian varieties. In [Alp26], the compactifications and their boundary charts allow cusp forms to be treated as sections vanishing along the boundary. The line

bundles whose degrees enter the estimates are determinant lines formed from the Hodge filtration of the abelian varieties. The argument uses resolution and principalisation in characteristic zero, going back to Hironaka [Hir64], to work on smooth models and to make the base ideals of selected spaces of sections into Cartier divisors on a common modification. The comparison of their degrees with invariant volumes uses Mumford's [Mum77] study of metrics at the boundary. These degrees and volumes also enter the dimension formulas used in Section 4.5.

On the compactified moduli spaces, there are two questions about sections of the automorphic bundles to answer.

- First, how large are the spaces of sections, and when can local relations among sections be made global?
- Second, can a section be chosen nonzero at a prescribed point?

For the first question, [Alp26] uses the theorems of Kawamata [Kaw82] and Viehweg [Vie82] on vanishing of higher coherent cohomology to remove the cohomological obstructions. For the second, it uses multiplier ideals and vanishing in the sense of Nadel [Nad90], together with effective generation of sections of the kind proved by Angehrn and Siu [AS95], which bounds the positivity needed to obtain a section that does not vanish at the prescribed point. Together, these results are used in the bounds for dimensions of spaces of sections, orders of vanishing, and covolumes of their integral lattices in [Alp26].

Small sections and arithmetic intersection numbers. The integral sections carry norms at all the archimedean places of their field of definition. To compare their covolume with the geometry of the moduli space, [Alp26] uses arithmetic intersection numbers. It is useful to state the relevant facts for a regular, integral, projective scheme X flat over $\mathcal{O}_K$. Write $d := \dim X$, so that $\dim X_K = d - 1$, and let $\overline{L}$ be a line bundle with smooth Hermitian metrics invariant under complex conjugation. The number of integral sections of a tensor power with bounded norms is measured by

$$\widehat{h}^0(k\overline{L}) := \log \#\{s \in H^0(X, L^{\otimes k}) : ||s||_{\sup} \leq 1\},$$

where the supremum is over all archimedean places. Its leading growth is the arithmetic volume

$$\widehat{\mathrm{vol}}(\overline{L}) := \limsup_{k\to\infty} \frac{d!}{k^d} \widehat{h}^0(k\overline{L}).$$

Positive arithmetic volume means that $\overline{L}$ is arithmetically big; in particular, some positive power has a nonzero integral section with sup norm at most 1.

The positivity assumptions in this discussion have different roles. Nefness of L on the geometric fibres and semipositivity of its curvature concern the algebraic bundle and the variation of its metric. Arithmetic nefness requires, in addition, nonnegative degrees on horizontal arithmetic curves. Equivalently, the corresponding heights of algebraic points must be nonnegative. For an

algebraic point P defined over a finite extension F/K, properness extends P over $\mathcal{O}_F$, and its height is

$$h_{\overline{L}}(P) := \frac{\widehat{\deg}(P^*\overline{L})}{[F:\mathbb{Q}]}.$$

The degree here combines the orders of vanishing at finite primes with the logarithms of the archimedean norms. This is the usual normalization of heights.

For positive Hermitian line bundles, Gillet and Soulé [GS88] related the asymptotic covolumes of lattices of sections to arithmetic intersection numbers, using the estimates for analytic torsion of Bismut and Vasserot [BV89]. Abbes and Bouche [AB95] subsequently gave another proof. Zhang [Zha95, Theorem 1.4] extended this arithmetic Hilbert–Samuel formula to the case where L_K is ample, L is nef on geometric fibres, and its metric is smooth and semipositive. To state the formula, write $\chi_{\sup}(H^0(X, L^{\otimes k}))$ for the logarithm of the volume of the unit ball for the sup norm divided by the covolume of the lattice $H^0(X, L^{\otimes k})$. Both volumes use the same Lebesgue measure on $H^0(X, L^{\otimes k}) \otimes_{\mathbb{Z}} \mathbb{R}$. With $\widehat{c}_1$ denoting the first arithmetic Chern class and $\widehat{\deg}(\widehat{c}_1(\overline{L})^d)$ its top arithmetic self-intersection, the formula is

$$\chi_{\sup}(H^0(X, L^{\otimes k})) = \frac{\widehat{\deg}(\widehat{c}_1(\overline{L})^d)}{d!}k^d + o(k^d).$$

Zhang [Zha95, Corollary 1.9] applied Minkowski's theorem to deduce that positive arithmetic self-intersection gives a nonzero strictly small integral section of some positive tensor power.

The argument in [Alp26] also needs this conclusion when L_K is only nef. The further work of Yuan and Moriwaki matters at precisely this point. Yuan [Yua08, Theorem 2.2] proved an arithmetic analogue of the criterion for bigness of Siu [Siu93, §1]. For arithmetically ample Hermitian line bundles $\overline{A}$ and $\overline{B}$, the inequality

$$\widehat{\deg}\Big(\widehat{c}_1(\overline{A})^d\Big) > d\,\widehat{\deg}\Big(\widehat{c}_1(\overline{A})^{d-1}\widehat{c}_1(\overline{B})\Big)$$

ensures that $\overline{A} \otimes \overline{B}^{-1}$ is arithmetically big. The mixed intersection measures the cost of subtracting $\overline{B}$ from $\overline{A}$. Moriwaki [Mor09] adapted Yuan's estimates for changes in the number of small sections to prove continuity of arithmetic volume. His results give

$$\widehat{\mathrm{vol}}(\overline{L}) \geq \widehat{\deg}\Big(\widehat{c}_1(\overline{L})^d\Big)$$

when L is nef on every geometric fibre and its metric is smooth and semipositive. Under the stronger assumption that $\overline{L}$ is arithmetically nef, Moriwaki [Mor09] proved the arithmetic Hilbert–Samuel formula for the number of small sections:

$$\widehat{h}^0(k\overline{L}) = \frac{\widehat{\deg}(\widehat{c}_1(\overline{L})^d)}{d!}k^d + o(k^d).$$

The distinction is needed below: changing a metric by a constant preserves its curvature, but can change the signs of the heights on horizontal curves.

From small sections to heights of points. Here is how the existence of a small section gives the height comparison used in [Alp26]. Suppose L is nef on every geometric fibre, its metric is smooth and semipositive, and $\delta := \deg_{L_K}(X_K) > 0$. Normalize the arithmetic self-intersection by putting

$$h_{\overline{L}}(X) := \frac{\widehat{\deg}(\widehat{c}_1(\overline{L})^d)}{d[K:\mathbb{Q}]\delta}.$$

For a real number c, let $\overline{L}(-c)$ denote the same line bundle with norms multiplied by e^c. Arithmetic intersection theory gives

$$h_{\overline{L}(-c)}(X) = h_{\overline{L}}(X) - c,$$
$$h_{\overline{L}(-c)}(P) = h_{\overline{L}}(P) - c.$$

If $c < h_{\overline{L}}(X)$, the rescaled bundle still has positive arithmetic self-intersection. Moriwaki's theorem therefore supplies a nonzero integral section s of some $L^{\otimes k}$ with $||s||_{\sup,\overline{L}} \leq e^{-kc}$. For P outside the zero divisor of s, the finite valuations of $s(P)$ are nonnegative, while its archimedean norms satisfy this bound. Consequently $h_{\overline{L}}(P) \geq c$.

Zhang [Zha95, proof of Theorem 5.2, p. 210] used this rescaling and evaluation of small sections in his inequalities for successive minima. The version in [Alp26] uses Moriwaki's result to obtain the section under the stated nefness hypotheses. The two comparisons needed here are particularly simple.

- If $h_{\overline{L}}(P) \geq a$ for every algebraic point P, then $h_{\overline{L}}(X) \geq a$.
- If a Zariski dense set of algebraic points satisfies $h_{\overline{L}}(P) \leq b$, then $h_{\overline{L}}(X) \leq b$.

For the first comparison, $\overline{L}(-a)$ is arithmetically nef. Applying Moriwaki's [Mor09] arithmetic Hilbert–Samuel formula to this bundle shows that its top arithmetic self-intersection is nonnegative: the quantities $\widehat{h}^0(k\overline{L}(-a))$ are nonnegative because the zero section is included in the count. Since $h_{\overline{L}(-a)}(X) = h_{\overline{L}}(X) - a$, the desired bound follows. For the second, if $h_{\overline{L}}(X) > b$, choose c strictly between these two numbers. The section just constructed forces every point of height at most b into its zero divisor, contradicting Zariski density. Thus lower bounds valid at every point and upper bounds on a dense set give corresponding bounds for the arithmetic self-intersection.

For the application to Hodge bundles, there is a difficulty at the boundary: the natural metrics need not extend smoothly to the compactification. Let X now be a compactified component used in [Alp26], again of absolute dimension d, and let $\mathcal{L}$ be the determinant line obtained from its Hodge quotients. For a local nonvanishing section e of $\mathcal{L}$, write $\phi := -\log||e||$ for the metric potential. The modification in [Alp26] replaces ϕ by $\phi + U/2$.

For a point z in the interior of a complex fibre, let A_z be the polarized abelian variety with $\mathcal{O}_E$-action represented by z, and let $\Lambda_z := H_1(A_z, \mathbb{Z})$ be its integral period lattice. The function U is defined by

$$U(z) := \max\left\{0, \sup_I(-\log N_z(I))\right\},$$

where I runs over the nonzero saturated $\mathcal{O}_E$-sublattices of Λ_z isotropic for the alternating form supplied by the polarization. To define $N_z(I)$, choose an $\mathcal{O}_E$-linearly independent family in I whose span has finite index. Project its period vectors to the $\mathcal{O}_E$-eigenspaces of $\mathrm{Lie}(A_z)$ corresponding to the CM type chosen in this construction, with the Hermitian metrics given by the polarization. Then $N_z(I)$ is the product of the resulting Gram determinants divided by the index of the span in I. Thus this is a change of metric potential; the local frame norm is multiplied by $e^{-U/2}$.

The corrected potential is locally bounded near the boundary. To smooth it, [Alp26] convolves the local potentials with mollifiers and glues the results using a regularized maximum. The original Hodge potential is regularized separately: the local potentials are truncated from below before convolution, with the truncation level tending to $-\infty$ as the smoothing parameter tends to 0. Fix a sufficiently large cutoff B for rational primes. Write $\mathcal{L}_B$ for $\mathcal{L}$ tensored with the line bundles of the prescribed vertical divisors above primes at most B; these divisors are supported in the finite fibres. Denote the two resulting smooth metrized line bundles by $\overline{\mathcal{L}}^+_{B,t'}$ and $\overline{\mathcal{L}}_{B,t}$, respectively, where $t', t > 0$ are the smoothing parameters. To obtain the required positivity, [Alp26] also uses an auxiliary line bundle $\mathcal{A}$ on X, with a smooth metric of positive curvature, and, for integers $1 \le p \le q$, considers

$$\overline{\mathcal{M}}_{p,q,t'} = (\overline{\mathcal{L}}^+_{B,t'})^{\otimes q} \otimes \overline{\mathcal{A}}^{\otimes p}.$$

For fixed p and q and sufficiently small t', these bundles are nef on the geometric fibres, ample on the generic fibre, and have semipositive curvature and metrics invariant under complex conjugation.

The point estimates in [Alp26] give $h_{\overline{\mathcal{M}}_{p,q,t'}}(P)/q \ge a - (p/q)K_{\mathcal{A}}$ for every algebraic point P. Here a depends on B but not on p, q, t, or t', and $K_{\mathcal{A}}$ is a constant with $h_{\overline{\mathcal{A}}}(P) \ge -K_{\mathcal{A}}$. The estimates include points on the boundary as well as points in the interior, so the lower height comparison applies to all of X and gives the same lower bound for $h_{\overline{\mathcal{M}}_{p,q,t'}}(X)/q$.

It remains to remove the correction to the metric. A Stokes formula compares the arithmetic intersections for the corrected and original potentials. In [Alp26], the limits are taken first as $t' \to 0$, then as $t \to 0$, and finally with $p = 1$ and $q \to \infty$ to remove the auxiliary bundle. Write $\langle U \rangle$ for the average of U with respect to the $(d-1)$st power of the curvature form of the original Hodge metric on $\mathcal{L}$. More precisely, sum the integrals over the interiors of the complex fibres and divide by the same sum with U replaced

by 1. The resulting inequality, in the height normalization above, is

$$\lim_{t\to 0} h_{\overline{\mathcal{L}}_{B,t}}(X) \geq a - \frac{1}{2}\langle U\rangle.$$

The boundary estimates in [Alp26] justify these limits and the integration of U. The bound for $\langle U\rangle$ then turns the lower bounds for heights of points into the lower bound for arithmetic self-intersection used in the covolume estimates.

From heights to lines in Hermitian lattices. The upper comparison for heights of varieties uses a Zariski dense set of points of bounded height. On a Grassmannian, such points are quite explicit. For $0 < j < r$, let $\mathrm{Gr}(j, r)$ parametrize the j-dimensional subspaces of K^r, equipped with its standard Hermitian metrics. On the standard affine chart, represent a subspace by the rows of $[I_j\ Z]$, where Z is a $j \times (r-j)$ matrix. For the Plücker line bundle with its Fubini–Study metric, the squared norm of the resulting Plücker vector is $\det(I_j + ZZ^*)$, where Z^* is the conjugate transpose. Choose the entries of Z to be roots of unity. The Plücker coordinates are integral and include a coordinate equal to 1, so the finite places contribute nothing to their height. At every complex embedding, Hadamard's inequality gives

$$\frac{1}{2}\log\det(I_j + ZZ^*) \leq \frac{j}{2}\log(r-j+1).$$

These points are Zariski dense in the Grassmannian. The upper comparison therefore bounds its normalized arithmetic self-intersection by the same expression. The construction in [Alp26] also uses products of Grassmannians; the heights add for the product Plücker line bundle.

To see how this controls a lattice, let M be a Hermitian $\mathcal{O}_K$-lattice of positive rank r, with K totally imaginary. The note uses twice the usual normalized Arakelov slope:

$$\mu(M) := -\frac{2}{[K:\mathbb{Q}]r}\log\mathrm{covol}(M) - \log 2 + \frac{\log|\Delta_K|}{[K:\mathbb{Q}]}.$$

Here Δ_K is the discriminant of K; the correction makes $\mu(\mathcal{O}_K^r) = 0$ for the standard metrics. Thus a lower bound for $\mu(M)$ gives an upper bound for the covolume, with the rank and the field fixed. The maximum slope $\mu_{\max}(M)$ is the largest mean slope of a nonzero saturated sublattice; the minimum slope $\mu_{\min}(M)$ is the smallest mean slope of a nonzero torsion-free quotient with its quotient metrics.

The projective bundle $\mathbb{P}(M)$ has generic fibre $\mathbb{P}(M \otimes K)$, parametrizing lines in $M \otimes K$. With $\mathcal{O}(1)$ the dual of the tautological line and the induced Fubini–Study metric, the calculation in [Alp26] gives

$$h_{\overline{\mathcal{O}}(1)}(\mathbb{P}(M)) \leq -\frac{1}{2}\mu(M) + \frac{1}{2}\log r.$$

The lower height comparison then supplies an algebraic point whose height is at most this upper bound plus any prescribed positive error. If the point

is defined over K'/K and corresponds to the saturated line $L \subseteq M \otimes_{\mathcal{O}_K} \mathcal{O}_{K'}$, then $h_{\overline{\mathcal{O}}(1)}([L]) = -\mu(L)/2$. A point of small height therefore gives a line of large degree. Applying this to a sublattice attaining the maximum slope shows that, for every $\epsilon > 0$, some finite extension K'/K contains such a line with

$$\mu(L) \geq \mu_{\max}(M) - \log r - \epsilon.$$

This is the absolute Siegel lemma used in the argument. Gaudron and Rémond [GR13, Theorem 1.1 and Lemma 3.2] deduced it, with a sharper constant, from Zhang's inequalities [Zha95]. The finite extension is natural here: the point supplied by the height comparison need not be K-rational.

For the period maps, [Alp26] also needs to compare the slopes of a lattice and its image under an integral map. Bost's [Bos01, §4.1.4] method of slopes gives the relevant inequality. For a nonzero integral map $b : M \to N$ of Hermitian $\mathcal{O}_K$-lattices, let b_ν be its realization at a complex place ν, with the induced operator norm. Use the same factor of two in its height:

$$h(b) := \frac{2}{[K : \mathbb{Q}]} \sum_{\nu|\infty} \log||b_\nu||_{\mathrm{op}}^2.$$

Each complex place occurs once. With the induced metrics on the image, the inequality is

$$\mu_{\min}(b(M)) \geq \mu_{\min}(M) - h(b).$$

In particular, bounding the operator norms limits how far the minimum slope can fall under the map. In Section 4.4, [Alp26] uses this inequality after selecting sublattices of the domains of the period maps with the Harder–Narasimhan filtration.

Integral structures and covolumes. There is also an integral question behind the height estimates: how do the Hodge bundles and their sections behave at the finite primes when the weights are changed? Here the relevant p-divisible groups are those of the abelian varieties parametrized by the moduli spaces. Grothendieck–Messing theory, developed in Messing's [Mes72] study of crystals associated with p-divisible groups, describes deformations through lifts of the Hodge filtration over nilpotent divided-power thickenings. In [Alp26], this enters the construction of integral points on the moduli spaces at good primes and the comparison of valuations when the weights are changed. The integral estimates for Hodge–Tate maps in Fargues [Far11, §5.3.2] also belong here: they relate the Tate module, invariant differentials, and duality while controlling the torsion in these comparisons between integral modules. The later theory of Bhatt, Morrow, and Scholze [BMS18] compared integral p-adic cohomology theories through modules with Frobenius. Bhatt and Scholze [BS22] developed prismatic cohomology, and Anschütz and Le Bras [ALB23] used it to classify p-divisible groups over quasisyntomic rings by prismatic Dieudonné crystals. The prismatic input in [Alp26] is used to identify the kernel of the Hodge–Tate map with the annihilator of the kernel of the logarithmic differential under the Tate pairing. This comparison

then enters the estimates of heights and covolumes, where the torsion in the integral modules affects the valuations being estimated.

For each choice of level and weight, [Alp26] first bounds the heights of the compactified components and compares their algebraic degrees with the dimensions of the spaces of sections. It then uses the Hilbert–Samuel estimates to bound the arithmetic degrees of the integral lattices of sections. Adding these degrees over the components and dividing by the total rank gives a lower bound for the mean slope of the resulting lattice of automorphic forms, hence an upper bound for its logarithmic covolume. The estimates retain the losses from changing weights, passing between covers, and treating the boundary. These are the bounds on lattices of automorphic forms that enter Proposition 4.3. To transfer them through the maps b_α to a bound for $\mu(C)$, [Alp26] also uses information about the subspace of $C \otimes K$ spanned by their images.

4.4. **How the maps defined by periods control the mean slope.** There are many maps b_α, but their number alone does not settle the question of spanning: their images could all lie in the same small subspace of $C \otimes K$. Moreover, [Alp26] restricts the maps to sublattices of their domains chosen according to slope. An estimate for the span of the original images would not suffice. The linear algebra below bounds how much of the target can be missed after those restrictions, in terms of the dimensions removed from the domains. The hypotheses measure how nearly the maps preserve lengths and how widely their images span the target.

Let $\mathcal{H}$ be a finite-dimensional complex Hilbert space, and let $B_i : \mathbb{C}^{r_i} \to \mathcal{H}$ be finitely many linear maps. Put $D := \dim_{\mathbb{C}} \mathcal{H} > 0$ and $R := \sum_i r_i > 0$. Write B_i^* for the adjoint and $||B_i||_{HS}^2 := \mathrm{Tr}(B_i^* B_i)$ for the squared Hilbert–Schmidt norm. Choose positive weights γ_i and set $K_i := \gamma_i B_i^* B_i$ and $Q := \sum_i \gamma_i B_i B_i^*$. Assume that $0 \le e \le 1/4$ and $\eta \ge 0$ satisfy both of the following bounds:

$$\sum_i ||K_i - I||_{HS}^2 \le eR, \tag{4.2}$$

$$\mathrm{Tr}(Q^2) \le (1+\eta)\frac{\mathrm{Tr}(Q)^2}{D}. \tag{4.3}$$

Equation 4.2 says that the weighted maps nearly preserve lengths, on average over the maps B_i. To interpret Equation 4.3, observe that Cauchy–Schwarz always gives $\mathrm{Tr}(Q^2) \ge \mathrm{Tr}(Q)^2/D$, with equality precisely when Q is scalar. Equation 4.3 therefore requires Q to be nearly scalar. In particular, the images of the maps B_i cannot all be concentrated in a small part of $\mathcal{H}$.

The equality case suggests a useful analogy with finite tight frames: collections of vectors for which the sum of the associated rank-one operators is a positive scalar multiple of the identity. In their study of such frames, Benedetto and Fickus [BF03] characterized finite normalized tight frames by minimizing the sum of squared inner products of their vectors. The maps B_i

have domains of varying dimensions, and Equations 4.2 and 4.3 allow small errors. Nevertheless, the comparison with a scalar operator serves the same purpose: it measures how evenly the images of the maps B_i occupy $\mathcal{H}$.

Under these hypotheses, restricting the domains of the B_i loses only the following proportion of $\mathcal{H}$.

Lemma 4.2. *Let $E_i \subseteq \mathbb{C}^{r_i}$ be subspaces such that $\sum_i \operatorname{codim} E_i \le \epsilon R$, where $\epsilon \ge 0$. Then*

$$\operatorname{codim}_{\mathcal{H}} \sum_i B_i E_i \le (\eta + 4\epsilon + 4\sqrt{e})D.$$

Proof. The point is to estimate the trace of Q on the part of the target missed by the restricted maps. Write P_U for orthogonal projection onto a subspace U, and put $T := (\sum_i B_i E_i)^{\perp}$, $w := \dim T$, and $t := \operatorname{Tr} Q$. By Cauchy–Schwarz and Equation 4.2, $|t - R| \le R\sqrt{e}$, so $t \ge R/2$. Since $B_i^* T \subseteq E_i^{\perp}$,

$$\begin{aligned}\operatorname{Tr}(P_T Q) &\le \sum_i \operatorname{Tr}(P_{E_i^{\perp}} K_i) \\ &\le \sum_i \operatorname{codim} E_i + \left(\sum_i \operatorname{codim} E_i\right)^{1/2} \left(\sum_i ||K_i - I||_{\mathrm{HS}}^2\right)^{1/2}.\end{aligned}$$

Since $\sum_i \operatorname{codim} E_i \le \epsilon R$, Equation 4.2 now gives $\operatorname{Tr}(P_T Q) \le R(\epsilon + \sqrt{\epsilon e})$. On the other hand, by Equation 4.3,

$$\left|\left| Q - \frac{t}{D} I \right|\right|_{\mathrm{HS}}^2 \le \eta \frac{t^2}{D}.$$

Since P_T has Hilbert–Schmidt norm $\sqrt{w}$, the trace on T consequently satisfies

$$\operatorname{Tr}(P_T Q) \ge \frac{tw}{D} - t\sqrt{\frac{\eta w}{D}}.$$

Since $t \ge R/2$, the two bounds imply

$$w - \sqrt{\eta D w} \le 2D(\epsilon + \sqrt{\epsilon e}).$$

The inequality $\sqrt{\eta D w} \le (w + \eta D)/2$ implies $w \le D(\eta + 4\epsilon + 4\sqrt{\epsilon e})$. For $\epsilon \le 1$, this is the desired estimate; for $\epsilon \ge 1$, the bound $w \le D$ suffices. □

To use Lemma 4.2 for the period maps, [Alp26] chooses the subspaces according to the slopes of the lattices N_α, using a filtration whose construction came from vector bundles on curves. Harder and Narasimhan [HN75] introduced a filtration of a vector bundle whose successive quotients were semistable and had decreasing slopes. For lattices, Arakelov degree plays the role of the degree of a vector bundle. Stuhler [Stu76] carried the construction over to Euclidean lattices, and Grayson [Gra84] developed the corresponding reduction theory over rings of integers of number fields. The successive quotients of the resulting filtration have decreasing mean slopes and are semistable: no saturated sublattice has larger mean slope than the quotient itself.

For a given slope threshold, [Alp26] retains the part of each N_α above that threshold, adjusted by the bound for the height of b_α. Bost's inequality from Section 4.3 then places its image in the part of C above the unadjusted threshold. By Lemma 4.2, the images span all but a bounded codimension in C, so only that proportion of its slopes can lie below the threshold. Varying the threshold and integrating gives a bound for the mean slope of C.

For Hermitian lattices over a totally imaginary field K, use the normalization of slopes and heights in Section 4.3. The same factor of two applies to all the Harder–Narasimhan slopes.

Consider full Hermitian $\mathcal{O}_K$-lattices C and N_i of ranks $d > 0$ and $r_i > 0$, together with integral maps $b_i : N_i \to C$. Assume that $h(b_i) \le h_i$ whenever $b_i \neq 0$, and put $R := \sum_i r_i > 0$. Choose $s \in \mathbb{R}$, $u \ge 0$, and $H \ge 0$ such that

$$\begin{aligned} Rs &\le \sum_i r_i(\mu(N_i) - h_i), \\ \mu_{\max}(N_i) - h_i &\le s + u, \\ s - \mu_{\min}(C) &\le H. \end{aligned}$$

Here s is a lower bound for the mean slope of the domain lattices after the bounds h_i for the heights of the maps are subtracted. The bound $s + u$ limits how far the largest adjusted slope can lie above s; together, these two bounds limit the proportion of adjusted slopes below a given threshold. The number H bounds how far a slope of C can lie below s.

At one archimedean place of K, assume also that, after choosing invertible coordinates, the maps b_i are projections to $\mathbb{C}^d$ of maps $B_i : \mathbb{C}^{r_i} \to \mathbb{C}^d \oplus \mathbb{C}^{d_1}$ satisfying the hypotheses of Lemma 4.2, with positive weights, $0 \le e \le 1/4$, and $\eta \ge 0$. Put $D := d + d_1$ and assume that $(1-\delta)D \le d$ for some $0 \le \delta < 1$. Thus Lemma 4.2 applies in a slightly larger space before projection to the complex realization $\mathbb{C}^d$ of C. Proposition 4.3 bounds the mean slope of C in terms of the slopes of the domain lattices.

Proposition 4.3. *With the lattices, maps, and bounds just described,*

$$\mu(C) \ge s - \frac{4u\log(1 + H/u) + (\eta + 4\sqrt{e})H}{1 - \delta}. \tag{4.4}$$

When $u = 0$, interpret $u\log(1 + H/u)$ as zero.

Proof. List the slopes of all N_i, with rank multiplicities, after subtracting h_i in the ith summand. Their average is at least s, and each is at most $s + u$. For $\tau > 0$, let $k(\tau)$ count the adjusted slopes below $s - \tau$. Since

$$Rs \le k(\tau)(s - \tau) + (R - k(\tau))(s + u),$$

it follows that $k(\tau) \le Ru/(u + \tau)$. If $u = 0$, every adjusted slope equals s, so the same conclusion holds.

Let $N_i(\tau)$ be the Harder–Narasimhan sublattice with slopes at least $s - \tau + h_i$. Since b_i is integral and has height at most h_i, the slope inequality of Bost [Bos01, §4.1.4], applied after quotienting by its kernel, shows that each

nonzero image $b_i N_i(\tau)$ has minimum slope at least $s - \tau$. Thus the images $b_i N_i(\tau)$ lie in the Harder–Narasimhan part $C_{\geq s-\tau}$.

By Lemma 4.2, the images of the complex spans of the $N_i(\tau)$ have corank at most

$$\left(\eta + 4\frac{u}{u+\tau} + 4\sqrt{e}\right) D$$

in $\mathbb{C}^d \oplus \mathbb{C}^{d_1}$. Projection to $\mathbb{C}^d$ cannot increase corank. Since $d \geq (1-\delta)D$, the proportion of slopes of C below $s - \tau$ is therefore at most

$$\frac{1}{1-\delta}\left(\eta + 4\frac{u}{u+\tau} + 4\sqrt{e}\right).$$

Write $\lambda_1, \ldots, \lambda_d$ for the slopes of C. Each is at least $s - H$, so integration recovers the total deficit below s:

$$s - \mu(C) \leq \frac{1}{d}\sum_{j=1}^{d}(s - \lambda_j)_+ = \int_0^H \frac{\#\{j : \lambda_j < s - \tau\}}{d}\, d\tau.$$

Since $\int_0^H u/(u+\tau)\, d\tau = u \log(1 + H/u)$, the bound for the proportion of slopes of C below $s - \tau$ proves Equation 4.4. □

The conclusion now bounds the mean slope of C, and hence its covolume, in terms of the lattices in the domains and the heights of the maps.

Remark 4.4. The logarithm in Equation 4.4 comes from integrating the proportion of slopes that were discarded from the domains. In particular, the estimate uses the distribution of the images of all the maps, beyond what an estimate for any one map could provide.

Applying the inequality to the period maps. For each α, let h_α bound the height of b_α, and let s be the mean of $\mu(N_\alpha) - h_\alpha$, weighted by the ranks of N_α. These are the quantities used for h_i and s in Proposition 4.3. Retain the variable x, the small parameter ε, and the scale $b = \exp(x^\varepsilon)$ from Section 3.

For the periods, [Alp26] proves the two estimates in Lemma 4.2 before making all the restrictions on the automorphic forms. At the distinguished complex place of K used to realize the periods, it first works with the full spaces $\mathcal{U}_\alpha$ of initial automorphic forms on $U(X_0)$ at the chosen levels and weights. For the target, take the finite-dimensional part of $L^2([U(Y_0)])$ selected by the level and the representations at infinity, before the further restrictions defining W. Let $\mathcal{H}$ denote its conjugate Hilbert space. The Petersson pairing identifies $\mathcal{H}$ linearly with the dual of that space of test automorphic forms; the retained part is therefore identified with $W^\vee$.

For each α, an operator $\mathcal{P}_\alpha : \mathcal{U}_\alpha \to \mathcal{H}$ is obtained from the integrations defining the periods. In orthonormal bases $(e_{\alpha,j})$ and (f_a), its matrix has entries $\langle f_a, \mathcal{P}_\alpha e_{\alpha,j}\rangle$. Call this matrix B_α^{full}. The weights γ_α in Lemma 4.2 are fixed by Schur orthogonality for these operators. More precisely, let $\mathcal{K}_f$ be the finite level subgroup of $U(Y_0)(\mathbb{A}_{E^+,f})$, where the subscript f denotes finite adeles. Retain the Bessel unipotent subgroup N from Section 4.1, and put

$N_\infty := N(E^+ \otimes_{\mathbb{Q}} \mathbb{R})$ and $U_\infty = U(X_0)(E^+ \otimes_{\mathbb{Q}} \mathbb{R})$. The finite adelic translate x_α used in the period specifies, with n_f denoting the finite adelic component,

$$\Gamma_\alpha := \{n \in N(E^+) : n_f \in x_\alpha \mathcal{K}_f x_\alpha^{-1}\}.$$

If π_α is the representation of the compact group U_∞ prescribed by the weights, the normalizing constant is

$$c_\alpha = \frac{\mathrm{vol}(\mathcal{K}_f)\mathrm{vol}(\Gamma_\alpha \backslash N_\infty)\mathrm{vol}(U_\infty)}{\dim \pi_\alpha},$$

and $\gamma_\alpha = c_\alpha^{-1}$. The nonzero scalar relating the analytic Bessel pairing to P_α is included in the change of coordinates allowed in Proposition 4.3.

The lattice N_α uses a subspace of $\mathcal{U}_\alpha$ specified at the auxiliary prime T_α, which is chosen to split completely in E^+. Here $f = [E^+ : \mathbb{Q}]$. At each of the f places of E^+ above T_α, let $\rho_{\alpha,v}$ denote the action of the multiplicative group of the residue field on $\mathcal{U}_\alpha$. Each of these residue fields has T_α elements. The source subspace used for N_α is

$$\mathcal{U}_\alpha \cap \bigcap_{v \mid T_\alpha} \ker\left(\sum_{c \in \mathbb{F}_{T_\alpha}^\times} \rho_{\alpha,v}(c)\right).$$

In other words, the component invariant under any of these multiplicative groups is removed. If $r_\alpha^{\mathrm{full}} := \dim \mathcal{U}_\alpha$ and r_α is the retained dimension, the calculation in [Alp26] gives

$$r_\alpha = r_\alpha^{\mathrm{full}}\left(1 - \frac{1}{T_\alpha - 1}\right)^f.$$

For $T_\alpha \geq b \geq 3$, the proportion removed is at most $2f/b$. Choose an isometric inclusion V_α of this subspace in $\mathcal{U}_\alpha$ and set $B_\alpha := B_\alpha^{\mathrm{full}} V_\alpha$. After the changes of coordinates just described, projecting B_α to $W^\vee$ gives b_α.

The estimates distinguish two cases: the interaction of period maps from different sources, and the norms of the period map on a single source. Write D_* for the main term in the dimension of the full space of test automorphic forms. Its factors are an invariant volume, indices of level subgroups, and formal degrees of the representations at infinity. Put $D := \dim \mathcal{H}$. With nonnegative error bounds η_0, η_1, and e_0, [Alp26] proves, for distinct α and β,

$$\gamma_\alpha \gamma_\beta \left|\left|(B_\alpha^{\mathrm{full}})^* B_\beta^{\mathrm{full}}\right|\right|_{\mathrm{HS}}^2 \leq (1 + \eta_0)\frac{r_\alpha^{\mathrm{full}} r_\beta^{\mathrm{full}}}{D_*}.$$

The estimates for a single source and for the total target are

$$\left|\left|\gamma_\alpha (B_\alpha^{\mathrm{full}})^* B_\alpha^{\mathrm{full}} - I\right|\right|_{\mathrm{HS}}^2 \leq e_0 r_\alpha^{\mathrm{full}},$$
$$D \leq (1 + \eta_1) D_*.$$

The first of these two bounds says that each normalized operator nearly preserves lengths. The second compares the actual dimension of $\mathcal{H}$ with the

term D_* used in the estimate for distinct sources. The errors η_0, η_1, and $\sqrt{e_0}$ can be bounded by $\exp(-x^{\varepsilon/8})$.

There are enough choices of α to make $\sum_\alpha r_\alpha^{\mathrm{full}} \geq D_*/\sigma$ for a chosen $0 < \sigma \leq 1/20$. Once η_0, η_1, and $\sqrt{e_0}$ are at most $\sigma/2$, restricting the sources with total proportion of lost dimension at most $\delta_0 \leq 1/4$ yields

$$e = 2e_0 \qquad \text{and} \qquad \eta = 5\sigma + 3\sqrt{\delta_0}$$

in Lemma 4.2. The proportion δ lost from the target is estimated separately. This distinction matters: restricting the sources changes e and η, whereas projecting to $W^\vee$ changes the denominator $1 - \delta$ in Proposition 4.3.

To integrate the proportions of slopes below a given threshold, the proof of Proposition 4.3 also uses a lower bound for $\mu_{\min}(C)$. Since C is the dual of the lattice of test automorphic forms, $\mu_{\min}(C) = -\mu_{\max}(C^\vee)$. The estimates for that lattice give, for a fixed $C_H > 0$,

$$s + \mu_{\max}(C^\vee) \leq bx^{C_H}.$$

Thus $H = bx^{C_H}$ is available when the distribution of slopes is integrated. The number u can be taken to be $\max_\alpha(\mu_{\max}(N_\alpha) - h_\alpha) - s$. The estimates for the initial lattices and for their changes at finite places bound this number as well. The comparison with the norms of the periods now depends on how these bounds vary with the parameters. In particular, the choice of T_α ties the dimensions of the spaces of automorphic forms to the estimate for their periods.

4.5. **How the period estimates recover the main term.** The question is now how these estimates for automorphic periods give a bound for the corrected sums over prime powers. Equation 3.4 expresses the difference $\mathcal{A}(a, t, w) - \mathcal{M}(a, w)$ in terms of the corrected sums G_χ. A bound for $\mathcal{A}(a, t, w)$ alone would not suffice. The comparison of periods in [Alp26] recovers $\mathcal{M}(a, w)$ with exactly the constants in that equation, leaving an error small enough for the proposed proof.

One might wonder why the constants should agree at all. The logarithmic derivatives of L-functions in the corrected sums have a classical connection with periods of CM abelian varieties. For imaginary quadratic points, Chowla and Selberg [CS49] expressed periods through gamma values. Anderson [And82] considered abelian varieties with abelian CM and related their periods, up to algebraic factors, to logarithmic derivatives of Dirichlet L-functions at zero. His proof brought together Shimura's relations between periods, the calculations for Fermat curves by Weil and by Gross and Rohrlich, and Deligne's description of these periods through gamma values. In [Alp26], this connection explains why estimates of automorphic periods can contain the constants already present in the sums over prime powers. The algebraic factors also contribute to the valuations at the finite places.

Accounting for these factors requires more than their algebraicity. The argument also needs to know how the periods change under Galois conjugation. For this, [Alp26] uses reciprocity for elliptic curves with complex multiplication: the conjugate elliptic curve and the action on its torsion are described through the Artin map. This is the elliptic case of the theory of Shimura and Taniyama [ST61]; see Milne [Mil07, Theorem 3.13] for a statement in terms of complex uniformization. A separate issue is to control denominators at finite places. Here [Alp26] uses the integrality of suitable modified values of abelian L-functions at negative integers over totally real fields, proved by Deligne and Ribet [DR80]. The reciprocity law and the integrality of the modified L-values control the algebraic factors in the expressions for periods; identities involving only absolute values would not determine their behavior at finite places.

The leading terms in the dimensions of spaces of automorphic forms also have a geometric meaning. Hirzebruch's proportionality principle related characteristic numbers of compact locally symmetric spaces to those of their compact duals. Mumford [Mum77] extended this comparison to noncompact arithmetic quotients by controlling Hermitian metrics near the boundary. In [Alp26], the resulting relation between intersection numbers and invariant volumes identifies the leading growth of the dimensions of these spaces with a polynomial in their highest weights. This is where the geometry discussed in Section 4.3 meets the dimension formulas.

There is another reason to expect the dimensions of the spaces of automorphic forms to appear here. The comparison of periods averages over these spaces, whose dimensions vary with the highest weights; the multiplicities of the automorphic representations affect the average as well. For these quantities, [Alp26] uses the dimension formulas developed by Weyl [Wey46] and the mass formulas associated with Siegel [Sie35]. In Siegel's theory, weighted counts of lattices were related to products of local densities. Together with Chowla–Selberg, these are the formulas singled out in [Alp26] in its explanation of the cancellation.

To see how the leading terms in [Alp26] cancel, fix one translate u in the finite group H from Section 3 and suppress u from the notation. Retain the positive parameters a_t indexed by $t \in T$, and let $w_{j,\rho} > 0$ be the Harish–Chandra weights at the complex places of E along the m lifts. The indices $\rho \in \Phi$ represent these places in the chosen CM type, and $f = |\Phi| = [E^+ : \mathbb{Q}]$, as in Section 3. Retain $\ell_\kappa = L'(1,\kappa)/L(1,\kappa)$. Let Δ_E be the discriminant of E, and put $\nu := |\Delta_E|^{1/[E:\mathbb{Q}]}/2$, half its root discriminant. With $A := \sum_{t\in T} a_t$, write

$$\log \Omega := \gamma + \log(\pi/\nu) - \ell_\kappa.$$

The lower bounds for the valuations of period values contribute the term D_{div} below. The quantity L_w records the logarithms of the weights:

$$D_{\mathrm{div}} := m \sum_{t \in T} (a_t \log a_t - a_t - a_t \log \Omega),$$

$$L_w := \sum_{j=1}^{m} \frac{1}{f} \sum_{\rho \in \Phi} \log(w_{j,\rho} \nu / \pi).$$

The sum of the main terms from Equation 3.4 is

$$\mathcal{M}_{\mathrm{tot}} := \sum_{j=1}^{m} \sum_{t \in T} a_t \left(\log a_t - 1 - \gamma + \ell_\kappa - \frac{1}{f} \sum_{\rho \in \Phi} \log w_{j,\rho} \right).$$

Since $\log(w\nu/\pi) = \log w + \log \nu - \log \pi$, the contributions from $\log \nu - \log \pi$ cancel against the corresponding terms in $\log \Omega$. Consequently,

$$D_{\mathrm{div}} - A L_w = \mathcal{M}_{\mathrm{tot}}. \tag{4.5}$$

Observe that $\mathcal{M}_{\mathrm{tot}}$ is exactly the main term from Equation 3.4, summed over the theta lifts and the complex places indexed by T. Thus the comparison of periods removes the main term that had to be subtracted from the sum over prime powers.

To see how this identity enters the period estimates, consider first the choice of T_α. This is an auxiliary rational prime in the level of the initial automorphic forms, and [Alp26] chooses its size using the dimensions of the spaces of automorphic forms. Write d_α^* for the product of the mass, the indices of the fixed level subgroups, and the dimensions of the representations at infinity for the initial group. The full space at level T_α, before imposing the additional conditions at this prime, has dimension $d_\alpha^*(T_\alpha^k - 1)^f$. The factor $T_\alpha^k - 1$ counts nonzero vectors over the residue field, and there are f places above T_α in E^+.

Recall the corresponding main term D_* for the dimension of the space of test automorphic forms from Section 4.4. The parameter $\Gamma_P > 1$ controls the size of the source spaces relative to that test space: the prime in [Alp26] is chosen so that

$$\left(\frac{D_*}{d_\alpha^* \Gamma_P^2} \right)^{1/(fk)} \le T_\alpha \le \left(1 + \frac{1}{fk} \right) \left(\frac{D_*}{d_\alpha^* \Gamma_P^2} \right)^{1/(fk)}.$$

The prime also satisfies the prescribed splitting and congruence conditions. Thus the estimates from Weyl's formula and the mass formula determine a bound for $\log T_\alpha$. The term $-A \log T_\alpha$ below comes from the period map at this auxiliary prime.

The other part of the comparison is the change of integral lattice from M_α to N_α. Start with the lattice M_α of initial automorphic forms. Impose the restrictions under which the lower bounds for valuations of period values hold, obtaining M'_α. At a finite prime $\mathfrak{p}$ of K, let $e_\mathfrak{p}$ be the ramification index over the rational prime below $\mathfrak{p}$. Measure local exponents in units in

which that rational prime has valuation 1. For the jth lift, write $g_{j,\mathfrak{p}}$ for its contribution to the lower bound for valuations of period values, $F_{j,\mathfrak{p}}$ for the change from the local choice of Hecke characters, and $E_{\alpha,j,\mathfrak{p}}$ for the correction from normalizing the lift. Set $g_{j,\mathfrak{p}} = 0$ at primes omitted from the gain. All these exponents have finite support. The modified lattice has the form

$$N_\alpha = \left(\prod_{\mathfrak{p}} \mathfrak{p}^{-\left\lfloor e_{\mathfrak{p}} \sum_{j=1}^{m} (g_{j,\mathfrak{p}} + F_{j,\mathfrak{p}} - E_{\alpha,j,\mathfrak{p}}) \right\rfloor} \right) M'_\alpha.$$

The divisibility of the period values permits these negative powers of prime ideals: the values remain integral after the elements of M'_α have been divided by the indicated factors. In particular, N_α remains in the space of initial automorphic forms; the successive lifts determine its integral structure.

Passing from local exponents to slopes means weighting a valuation by the logarithm of the norm of its prime ideal. In the normalization of Section 4.3, put $c_{\mathfrak{p}} := 2e_{\mathfrak{p}} \log(N\mathfrak{p})/[K : \mathbb{Q}]$. If $\tau_{\mathfrak{p}}$ denotes the allowance from restricting M_α at $\mathfrak{p}$, the three contributions are

$$F := \sum_{\mathfrak{p}} c_{\mathfrak{p}} \sum_{j=1}^{m} F_{j,\mathfrak{p}},$$

$$G := \sum_{\mathfrak{p}} c_{\mathfrak{p}} \left(\sum_{j=1}^{m} g_{j,\mathfrak{p}} - \tau_{\mathfrak{p}} \right),$$

$$E_\alpha := \sum_{\mathfrak{p}} c_{\mathfrak{p}} \sum_{j=1}^{m} E_{\alpha,j,\mathfrak{p}}.$$

These local exponents explain where the sums over prime powers enter the construction. Consider a rational prime p unramified in E, of the inert type selected in Section 3, and outside the primes omitted from the gain. The local estimate applies for $p \le |H| x \exp(x^{\varepsilon/3})$. Fix a prime $\mathfrak{p}$ of K above p. Let $T_{\mathfrak{p}} \subseteq T$ consist of the selected places for which the local divisibility conditions hold, and let τ_t be the representative of the place indexed by t in the chosen CM type. For an odd integer ν, the relevant weight is $w_{j,c\mathrm{Frob}_{\mathfrak{p}}^{\nu}\tau_t}$, where c is complex conjugation. The conjugation returns the transported place to the chosen CM type. These are the parameters in the blocks introduced in Section 4.1: $w_{j,\rho} = e_{j-1,\rho}$ for a representative ρ in the CM type. With $R_{\mathfrak{p}}$ collecting the losses in the local estimate, [Alp26] obtains

$$\sum_{j=1}^{m} g_{j,\mathfrak{p}} - \tau_{\mathfrak{p}} \ge \sum_{j=1}^{m} \sum_{t \in T_{\mathfrak{p}}} \sum_{\substack{\nu \ge 1 \\ \nu \text{ odd}}} \left(\frac{a_t}{p^\nu} - w_{j,c\mathrm{Frob}_{\mathfrak{p}}^{\nu}\tau_t} \right)_+ - R_{\mathfrak{p}}.$$

The sum is finite because the weights are at least 1. The local condition on Frobenius accounts for the odd exponents. The weights satisfy $\sum_{\mathfrak{p}|p} c_{\mathfrak{p}} = 2 \log p$, since $\sum_{\mathfrak{p}|p} e_{\mathfrak{p}} \log(N\mathfrak{p}) = [K : \mathbb{Q}] \log p$. After accounting for the places omitted from $T_{\mathfrak{p}}$, the local gain therefore reproduces the terms $(a_t/p^\nu - w)_+$

in Equation 3.2. To keep the losses from $T_{\mathfrak{p}}$ and $R_{\mathfrak{p}}$ small, [Alp26] controls the number of selected places in short segments of Frobenius orbits, as described in Section 4.6.

Retain the bound h_α for the height of b_α, and write L_α for the remaining allowance for a decrease in slope when the lattices are changed. This allowance includes the loss from rounding the exponents in the fractional ideal down to integers. The local normalizations at two auxiliary primes g_0 and g_1 contribute $\ell_g := \log(g_0 g_1)$. The same term occurs in the estimate for $\log T_\alpha$. With nonnegative error bounds E_{div}, E_{sl}, E_{norm}, and E_{count}, the estimates in [Alp26] take the form

$$\mu(C) + h_\alpha - F \le D_{\mathrm{div}} - A \log T_\alpha + E_{\mathrm{div}}, \tag{4.6}$$

$$\mu(N_\alpha) \ge \mu(M_\alpha) + G + F - E_\alpha - L_\alpha, \tag{4.7}$$

$$\mu(M_\alpha) \ge -E_{\mathrm{sl}},$$

$$E_\alpha \le mA\ell_g + E_{\mathrm{norm}}, \tag{4.8}$$

$$\log T_\alpha \ge L_w + m\ell_g - E_{\mathrm{count}}. \tag{4.9}$$

Equation 4.6 compares the size of the period maps with the divisibility of their values, while Equation 4.7 records the effect of restricting and rescaling the initial lattices. The remaining bounds concern the slopes of M_α, the normalization of the theta lifts, and the auxiliary prime T_α selected using the dimension and mass formulas. Notice that the change F in the local normalizations occurs in both of the first two estimates, and that ℓ_g occurs both in E_α and in $\log T_\alpha$. These terms cancel when the inequalities are combined; the remaining expression $D_{\mathrm{div}} - AL_w$ is precisely the main term in Equation 4.5.

Recall that s is the mean of $\mu(N_\alpha) - h_\alpha$, weighted by the ranks of N_α. Let $\overline{L}$ be the corresponding mean of L_α. By Equations (4.6)– (4.9),

$$s \ge G + \mu(C) - \mathcal{M}_{\mathrm{tot}} - (E_{\mathrm{sl}} + E_{\mathrm{norm}} + E_{\mathrm{div}} + AE_{\mathrm{count}}) - \overline{L}.$$

Here Equation 4.5 accounts for the main term, and the two occurrences of $mA\ell_g$ cancel. Write $\mathcal{A}_{\mathrm{tot}}$ for the corresponding sum of the expressions $\mathcal{A}$ in Equation 3.2, and suppose $\mathcal{A}_{\mathrm{tot}} - G \le E_{\mathrm{loss}}$. Since Proposition 4.3 bounds $s - \mu(C)$, the excess of the arithmetic sum over its main term satisfies

$$\begin{aligned} \mathcal{A}_{\mathrm{tot}} - \mathcal{M}_{\mathrm{tot}} &\le E_{\mathrm{sl}} + E_{\mathrm{norm}} + E_{\mathrm{div}} + AE_{\mathrm{count}} \\ &\quad + \overline{L} + E_{\mathrm{loss}} + \frac{4u \log(1 + H/u) + (\eta + 4\sqrt{e})H}{1 - \delta}. \end{aligned} \tag{4.10}$$

The left-hand side is the excess over the main term that occurs in the character identity. Thus the period estimates have produced the required subtraction, and the problem is now to bound the errors in Equation 4.10. Among them is $\overline{L}$, the mean allowance for a possible decrease in slope when the integral lattices of automorphic forms are changed.

4.6. **The bound obtained from the periods.** To use Equation 4.10 in the character identity, [Alp26] chooses the automorphic forms, weights, and complex places of E so that the estimates for divisibility, norms, and slopes hold together. Recall the group H, the sets T and T_2 of selected complex places of E, and the scale b of the weights from Section 3. The chosen CM type supplies representatives for these places. At unramified finite primes, Frobenius acts on the places; in terms of the representatives, the action includes the complex conjugation needed to stay in the chosen CM type. The parameter k is the dimension of the initial Hermitian space X_0. Write $n := |H|$, $t := |T|$, and $\nu_2 := |T_2|$, and let a be an upper bound for the parameters a_t. The coefficient field is still denoted by K.

The choice of T_2 illustrates the competing requirements on the parameters. More places in T_2 give more divisibility of period values, but too many in a short segment of a Frobenius orbit increase the losses in the local estimates. To obtain a large set with few such intersections, [Alp26] uses an idea from additive combinatorics. Salem and Spencer [SS42] studied large sets of integers containing no nontrivial three-term arithmetic progression. Behrend [Beh46] obtained his stronger construction by placing integer vectors on a sphere before encoding them as integers.

The version used in [Alp26] excludes three-term linear relations with bounded nonzero coefficients whose sum is zero. After restricting and translating Behrend's set, [Alp26] uses its elements to select the places in T_2 so that a short segment of a Frobenius orbit meets T_2 in at most two points. Thus a method originally developed to avoid arithmetic progressions also selects enough complex places of E for the period values to acquire the required divisibility, while controlling how often the selected places occur in a short segment of a Frobenius orbit.

For a fixed exponent $c_{\mathrm{err}} \geq 0$, the estimate is

$$\mathcal{A}_{\mathrm{tot}} - \mathcal{M}_{\mathrm{tot}} \leq x^{c_{\mathrm{err}}\varepsilon}\left(ta\left(1 + \frac{m^2}{k}\right) + m\left(2nmbx^{\varepsilon} + \frac{\nu_2 kbx^{2\varepsilon}}{n^2}\right)\right). \tag{4.11}$$

The terms in Equation 4.11 record the errors in comparing the period values and slopes of lattices of automorphic forms as the weights, levels, and selected complex places of E vary.

Remark 4.5. Estimates for the inner products arising from pairs of period maps b_α control e and η in Proposition 4.3. Estimates for the changes in the integral lattices of automorphic forms control $\overline{L}$. The remaining terms account for divisibility of period values at finite places, the slopes of the initial lattices, and the dimensions of the spaces of automorphic forms. For the eventual choice of parameters, the crucial point is that c_{err} is independent of the small parameter ε.

Equation 4.11 is the bound used with the character identity in Equation 3.4. Before choosing the parameters, however, it remains to account for the loss of divisibility when the integral lattices of automorphic forms are changed.

The bound for $\overline{L}$ included in Equation 4.11 uses the Galois cohomology of Section 5.

5. Controlling denominators with Galois cohomology

The term $\overline{L}$ in Equation 4.10 allows for the loss of mean slope when [Alp26] restricts and rescales the initial lattice of automorphic forms to record divisibility of the period values. One difficulty in bounding that loss comes from projection onto chosen Hecke eigenspaces. A decomposition of a vector space need not split an integral lattice inside it, so projection onto a summand can introduce denominators even when the decomposition respects every Hecke operator.

To bound these denominators, [Alp26] uses integral operators that separate the chosen Hecke eigenspaces from the other eigenspaces. The elementary step is recorded in Lemma 5.1; the arithmetic enters in the divisibility of the operators' eigenvalues. In [Alp26], estimates for that divisibility involve congruences between automorphic representations and then extensions of their Galois representations. The number fields, weights, and representations all vary, so the estimates also retain their dependence on those choices.

5.1. **Denominators introduced by projection.** Retain the coefficient field K of the automorphic periods from Section 4.1. Fix a rational prime ℓ and a place $\lambda \mid \ell$ of K. Write K_λ for the completion and $\mathcal{O}_\lambda$ for its ring of integers. The valuation $\mathrm{ord}_{\mathcal{O}_\lambda}$ takes the value 1 on a uniformizer. Consider a full $\mathcal{O}_\lambda$-lattice M in $V = V_1 \oplus V_2$ over K_λ. If e denotes projection onto V_1, then $M \cap V_1 \subseteq eM$; the quotient measures the denominators introduced by the projection. Suppose an endomorphism preserving M acts by a nonzero scalar on V_1 and vanishes on V_2. Dividing by that scalar recovers the projection, and gives the following bound.

Lemma 5.1. *Suppose* $T \in \mathrm{End}_{\mathcal{O}_\lambda}(M)$ *acts by a nonzero scalar* $a \in \mathcal{O}_\lambda$ *on* V_1 *and by zero on* V_2*. Then* $aeM \subseteq M \cap V_1$*, and*

$$\mathrm{length}_{\mathcal{O}_\lambda}(eM/(M \cap V_1)) \leq \dim_{K_\lambda}(V_1)\mathrm{ord}_{\mathcal{O}_\lambda}(a).$$

Proof. For $m \in M$, the equality $aem = Tm$ shows that $aem \in M \cap V_1$. Thus $eM/(M \cap V_1)$ is a quotient of eM/aeM. Since eM is free of rank $\dim_{K_\lambda} V_1$, the desired length bound follows. □

In the automorphic application, distinct systems of Hecke eigenvalues can become congruent modulo λ. An integral Hecke operator that vanishes on one eigensystem then has an eigenvalue divisible by λ on any eigensystem congruent to it. This is why congruence ideals enter the estimate. Suppose the integral Hecke algebra is a local subalgebra $\mathbb{T} \subseteq \mathcal{O}_\lambda^{\mathcal{X}}$ that separates the finite set $\mathcal{X}$ of eigensystems and preserves M. The coordinate t_π of $t \in \mathbb{T}$ is its eigenvalue on the eigenspace V_π corresponding to $\pi \in \mathcal{X}$. Write $V = \bigoplus_{\pi \in \mathcal{X}} V_\pi$ and $m_\pi := \dim_{K_\lambda} V_\pi$. For a subset $Z \subseteq \mathcal{X}$, let e_Z project onto

$V_Z := \bigoplus_{\pi\in Z} V_\pi$. For each $\pi \in Z$, the relevant congruence ideal is

$$\mathfrak{c}_\pi := \{t_\pi : t \in \mathbb{T},\ t_{\pi'} = 0 \text{ for every } \pi' \notin Z\} \subseteq \mathcal{O}_\lambda.$$

The ideal records the divisibility forced on the eigenvalue at π when an integral Hecke operator kills every eigenspace outside Z. The estimate used in [Alp26] includes

$$\operatorname{length}_{\mathcal{O}_\lambda}(e_Z M/(M \cap V_Z)) \le \sum_{\pi\in Z} m_\pi \operatorname{length}_{\mathcal{O}_\lambda}(\mathcal{O}_\lambda/\mathfrak{c}_\pi).$$

At two fixed auxiliary primes, [Alp26] also uses integral operators to retain only a subset of the eigenspaces indexed by Z. If such an operator acts by an integer $A \neq 0$ on the retained eigenspaces and by zero on the other eigenspaces in Z, the bound has an additional $\operatorname{ord}_{\mathcal{O}_\lambda}(A)$ for each retained eigenspace, counted with its dimension.

To use these inequalities, [Alp26] bounds the lengths of the quotients $\mathcal{O}_\lambda/\mathfrak{c}_\pi$. Hida [Hid81] had related congruences between cusp forms to special values of their L-functions. In [Alp26], the Galois representations of Section 4.2 provide another way to measure the congruences: integral lattices can retain extension classes that disappear when the representations are semisimplified over K_λ. To identify the groups whose lengths enter the denominator bound, it helps to write out the coefficients of these extension classes and their restrictions at the finite places.

5.2. **Classes that vanish over the local fields.** Recall that h is the odd integer introduced in Section 4.1 to specify the sizes of the groups in the theta lifts; m is the number of lifts. Use $G_E = \operatorname{Gal}(\overline{E}/E)$ and the cyclotomic character ϵ_ℓ with the conventions of Section 4.2. The Tate twists in the denominator estimate come from an auxiliary theta lift from a unitary group in $r = k+m-1$ variables to a definite unitary group in $r+h$ variables. For a retained Hecke eigensystem π on the larger group, the associated semisimple representation of G_E over K_λ in [Alp26] has the form

$$R_\lambda \simeq \rho_\lambda \oplus \bigoplus_{c=1}^{h} \xi_c,$$
$$\xi_c = \xi_1 \epsilon_\ell^{-(c-1)}.$$

Here ρ_λ has dimension r. The one-dimensional characters ξ_c constitute the additional block in the base change; after division by ξ_1, they are $1, \epsilon_\ell^{-1}, \dots, \epsilon_\ell^{-(h-1)}$. This explains both the negative Tate twists and the range $1 \le j < h$ that will occur below.

The summand ρ_λ also contributes to the bound for denominators. To describe that contribution, choose a Galois-stable lattice T_π in $\rho_\lambda \otimes \xi_1^{-1}$, and set $A_\pi := (T_\pi \otimes_{\mathcal{O}_\lambda} K_\lambda)/T_\pi$. Write $H^i(E, A_\pi)$ for continuous Galois cohomology $H^i(G_E, A_\pi)$, and use the same convention for completions E_w at finite places w of E. The corresponding term in [Alp26] is the length of a Selmer group: the

subgroup of $H^1(E, A_\pi)$ whose restrictions to the local fields satisfy specified conditions. At $w \mid \ell$, take the image in $H^1(E_w, A_\pi)$ of $H^1_f(E_w, T_\pi \otimes_{\mathcal{O}_\lambda} K_\lambda)$. Here H^1_f denotes the finite, or crystalline, local condition of Bloch and Kato [BK90, §3]. Away from ℓ and the prescribed places of ramification, the permitted classes are unramified; at those prescribed places, no local restriction is imposed. The length of this Selmer group thus enters the denominator estimate along with the lengths of the groups with one-dimensional Tate coefficients. Further terms measure the indices between the integral lattices used in the local Galois comparisons, the class number of E, and valuations of differences of roots of unity.

To control the Selmer contribution, [Alp26] varies the auxiliary characters before fixing the lattices. Write s_π for its $\mathcal{O}_\lambda$-length divided by the ramification index $e_\lambda := \operatorname{ord}_{\mathcal{O}_\lambda}(\ell)$. The chosen finite-order twists make this length finite and give $s_\pi \le \nu_\ell + \ell^{-2}$. With ν_ℓ normalized by $\nu_\ell(\ell) = 1$, put

$$\nu_\ell := \min_{g \in G_E} \nu_\ell\Big(\epsilon_\ell(g)^{hr} - 1\Big).$$

The integer ν_ℓ measures how close the hr-th powers in the image of the cyclotomic character are to 1. Retaining $n = |H|$ from Section 4.6, [Alp26] obtains

$$\sum_\ell \nu_\ell \log \ell \le 2 \log n + \log h + \log r.$$

Since $\sum_\ell \ell^{-2} \log \ell$ also converges, [Alp26] can bound the contribution of these Selmer groups after summing over the coefficient primes. The groups with one-dimensional Tate coefficients require a different estimate.

For the one-dimensional terms, [Alp26] isolates classes whose local restrictions vanish. Let S_ℓ be the set of finite places of E above ℓ, and write G_{E,S_ℓ} for the Galois group of the maximal extension unramified away from S_ℓ and the archimedean places. The coefficients $\mathbb{Q}_\ell/\mathbb{Z}_\ell(-j)$ are $\mathbb{Q}_\ell/\mathbb{Z}_\ell$ with G_E acting through ϵ_ℓ^{-j}. For $j \ge 1$, the relevant group is the kernel of

$$H^1(G_{E,S_\ell}, \mathbb{Q}_\ell/\mathbb{Z}_\ell(-j)) \longrightarrow \prod_{w \mid \ell} H^1(E_w, \mathbb{Q}_\ell/\mathbb{Z}_\ell(-j)).$$

Denote this kernel by $\mathcal{S}_\ell(E, j)$ and, when it has finite length, write $s_\ell(E, j) := \operatorname{length}_{\mathbb{Z}_\ell} \mathcal{S}_\ell(E, j)$. By definition, these extension classes disappear after restriction to every E_w with $w \mid \ell$. The denominator estimate consequently depends on a global bound for their orders, as both ℓ and j vary. For a finite ℓ-primary module, its length times $\log \ell$ is the logarithm of its order, which explains the weight in the following sum. For an integer $h' \ge 1$, let $\mathcal{P}$ range over the finite sets of rational primes and put

$$\mathcal{V}(E, h') := \sup_{\mathcal{P}} \sum_{\ell \in \mathcal{P}} \log \ell \sum_{1 \le j < h'} s_\ell(E, j).$$

Set $\mathcal{V}(E, h') = \infty$ if $\mathcal{S}_\ell(E, j)$ has infinite length for any prime ℓ and any $1 \le j < h'$. A finite bound for $\mathcal{V}(E, h')$ therefore bounds the orders of the

kernels $\mathcal{S}_\ell(E, j)$ for all primes ℓ and all $1 \le j < h'$ at once, and includes their finiteness. The denominator estimate uses this sum with $h' = h$.

For the theta lifts, [Alp26] uses $\mathcal{V}(E, h)$ to bound the indices of projected lattices. At a finite place $w \mid \ell$ of the coefficient field K, let $M^{\mathrm{pr}}_{j,w}$ be the lattice obtained by projection at the j-th lift, for $1 \le j \le m$, and let $M^{\mathrm{int}}_{j,w}$ be the corresponding lattice obtained by intersection with the integral automorphic forms before projection. Write $\mathcal{O}_w$ for the ring of integers of K_w, and set $e_w := \mathrm{ord}_{\mathcal{O}_w}(\ell)$ and

$$\eta_{j,w} := \frac{1}{e_w} \mathrm{length}_{\mathcal{O}_w}\Big(M^{\mathrm{pr}}_{j,w}/M^{\mathrm{int}}_{j,w}\Big).$$

Among the tuples of auxiliary Hecke characters in Section 4.1, [Alp26] fixes one tuple μ_0 as a reference. Sum over the finitely many coefficient places where [Alp26] uses auxiliary characters different from the fixed reference choice, and put

$$\eta_j := \frac{1}{[K : \mathbb{Q}]} \sum_w e_w \eta_{j,w} \log Nw.$$

Here Nw is the cardinality of the residue field. Let $Z_j(\mu_0)$ be the image of the auxiliary theta lift at step j, and put $d_j := \dim_{\mathbb{C}} Z_j(\mu_0)$. For the parameter choices in [Alp26], the reference lift is injective, so d_j is also the dimension of its source. These source dimensions are independent of j. For sufficiently large x, [Alp26] obtains, simultaneously for $1 \le j \le m$, the following bound in the parameters of Section 4.6, with $f = [E^+ : \mathbb{Q}]$:

$$\eta_j \le d_j \left(C_2 b x^{2\varepsilon} \left(fm + \frac{|T_2| k (\log x)^2}{n^2} \right) + \mathcal{V}(E, h) \right), \tag{5.1}$$

provided $\mathcal{V}(E, h) < \infty$. The constant C_2 is independent of the varying parameters. The factor e_w restores the ordinary length before the local contributions are summed. For the normalization of mean slope in Section 4.3, a quotient of lattices of rank d_j changes mean slope by twice its normalized logarithmic index divided by d_j. Using Equation 5.1, [Alp26] can therefore bound the change in mean slope caused by these projections. It combines this estimate with the bounds for the remaining local indices to obtain the allowance for denominators in Equation 4.10.

The appearance of Galois cohomology in a bound for denominators also suggests a connection with special values of L-functions. In his conjectures, Lichtenbaum [Lic73] related zeta values to the orders of étale cohomology groups and algebraic K-groups; [Alp26] explicitly points to this perspective. The required estimate is $\mathcal{V}(E, h) \le \exp(x^\varepsilon/2)$ for every field and odd rank difference used in the theta lifts, uniformly for x in one selected long interval. For sufficiently large x, the parameter choices in [Alp26] satisfy $h \le \lfloor x^{401} \rfloor$. Since the lengths $s_\ell(E, j)$ are nonnegative, $\mathcal{V}(E, h) \le \mathcal{V}(E, \lfloor x^{401} \rfloor)$. In Proposition 5.2, [Alp26] obtains the required bound by estimating $\mathcal{V}(E, h')$ at the larger cutoff $h' = \lfloor x^{401} \rfloor$.

5.3. **Choosing fields for the estimate of cohomology.** The Dirichlet character in the sums over prime powers already enters the choice of field E. For a primitive Dirichlet character ψ of conductor N_ψ, write A_ψ for the abelian field cut out by ψ, namely the fixed field of $\ker\psi$ in $\mathbb{Q}(\zeta_{N_\psi})$, where ζ_N is a primitive N-th root of unity. In [Alp26], E contains A_ψ and varies in cyclotomic towers in which the lengths $s_\ell(E,j)$ can be estimated uniformly. To arrange this, [Alp26] enlarges A_ψ by a real cyclotomic layer and two imaginary quadratic fields. More precisely, for $a \in \mathbb{Z}_{\geq 0}$ and primes $q \equiv 3 \pmod 8$ and $d \equiv 7 \pmod{16}$, set

$$E(\psi, a, q, d) := A_\psi \mathbb{Q}(\zeta_{2^{a+2}})^+ \mathbb{Q}(\sqrt{-q})\mathbb{Q}(\sqrt{-d}).$$

The compositum is taken inside a common cyclotomic field; its factors need not be linearly disjoint. The superscript $+$ denotes the maximal real subfield. Write σ_b for the cyclotomic automorphism that raises roots of unity to their b-th powers. Among these automorphisms, those fixing $E(\psi, a, q, d)$ satisfy exactly the following conditions:

(i) $b \equiv \pm 1 \pmod{2^{a+2}}$;
(ii) $\psi(b) = 1$;
(iii) $\left(\frac{b}{q}\right) = 1$;
(iv) $\left(\frac{b}{d}\right) = 1$.

The field A_ψ remains fixed, while a grows with the parameter x in the sums over prime powers.

The congruences imposed on q and d enter both the choice of local automorphic representations and the uniform bounds for $s_\ell(E,j)$. The conditions on their sizes, for $Z > 1$ and $\varepsilon > 0$, are

$$\tfrac{1}{2} Z^{\varepsilon/5} < \log q \leq Z^{\varepsilon/5},$$
$$\tfrac{1}{2} Z^{\varepsilon/5} (\log Z)^2 < \log d \leq Z^{\varepsilon/5} (\log Z)^2.$$

In addition, [Alp26] imposes $q \mid d+1$ and $p \mid d+1$ for every odd prime $p \leq Z^{\varepsilon/5}$. Since both q and d are 3 modulo 4, the fields E are CM. The existence of pairs of primes satisfying all these size and congruence conditions is proved in [Alp26].

The uniform estimate for these fields can now be stated. Fix $0 < \varepsilon < 1/52$, and choose a function $a(x) \in \mathbb{Z}_{\geq 0}$ such that, for every $x \geq 1$,

$$x^{2+8\varepsilon} \leq 2^{a(x)} \leq 2x^{2+8\varepsilon}.$$

In the following estimate from [Alp26], the auxiliary primes q, d remain fixed throughout the interval, while the cyclotomic layer varies with x. This is the uniformity needed for the application to periods.

Proposition 5.2. *For the choices of ψ, ε, and $a(x)$ in Section 5.3, there are $J \geq 1$ and X_0 such that, for every $X \geq X_0$ and every pair of primes q, d satisfying the specified bounds and congruences with $Z = X^{1/J}$, there is*

$Y \geq X^{1/J}$ *with* $Y^2 \leq X$ *for which*

$$\mathcal{V}\Big(E(\psi, a(x), q, d), \lfloor x^{401} \rfloor\Big) \leq \exp\big(\tfrac{1}{2}x^{\varepsilon}\big) \tag{5.2}$$

for every $x \in [Y, Y^2]$.

Remark 5.3. The exponent 401 in [Alp26] leaves room for all the Tate twists arising from the repeated theta lifts; its precise value plays no further role. Notice, though, that the same interval works for all primes. The primes q, d remain fixed while $a(x)$, and hence E, may change; the bounds for $s_\ell(E, j)$ hold throughout that interval as both ℓ and j vary.

5.4. **The two eigenspaces for complex conjugation.** To estimate $s_\ell(E, j)$, [Alp26] separates the two eigenspaces for complex conjugation on $\mathcal{S}_\ell(E, j)$. The two parts lead to different arithmetic arguments, while the prime 2 needs separate treatment because the usual projections involve division by 2. Let c be complex conjugation on E. For odd ℓ, split $\mathcal{S}_\ell(E, j)$ into the eigenspaces on which c acts as $(-1)^j$ and $(-1)^{j+1}$. Denote their lengths by $s_\ell^+(E, j)$ and $s_\ell^-(E, j)$, respectively. The signs therefore include the twist by j. Since 2 is invertible in $\mathbb{Z}_\ell$, the projections $\big(1 \pm (-1)^j c\big)/2$ show that $s_\ell = s_\ell^+ + s_\ell^-$.

For a cutoff $h' \geq 1$, write B_2 for a bound on $\sum_{1 \leq j < h'} s_2(E, j)$, and write $B_\pm$ for bounds on the sums of $s_\ell^\pm(E, j)$ over odd primes ℓ and $1 \leq j < h'$, weighted by $\log \ell$. These bounds imply

$$\mathcal{V}(E, h') \leq B_2 \log 2 + B_- + B_+. \tag{5.3}$$

To estimate $s_\ell^+(E, j)$ and $s_\ell^-(E, j)$, [Alp26] uses global duality to relate the kernels $\mathcal{S}_\ell(E, j)$ to cohomology of positive Tate twists. Poitou [Poi67] and Tate [Tat63] established the duality theorems relating global cohomology to its local restrictions. In the application here, their theorems express $s_\ell(E, j)$ in terms of H^2, with corrections from local and global H^0.

For the finite groups in this comparison, the identity used in [Alp26] is

$$\begin{aligned} &\operatorname{length}_{\mathbb{Z}_\ell} H^2(G_{E,S_\ell}, \mathbb{Z}_\ell(j+1)) + \operatorname{length}_{\mathbb{Z}_\ell} H^0(G_{E,S_\ell}, \mathbb{Q}_\ell/\mathbb{Z}_\ell(-j)) \\ &\qquad = s_\ell(E, j) + \sum_{w \mid \ell} \operatorname{length}_{\mathbb{Z}_\ell} H^0(E_w, \mathbb{Q}_\ell/\mathbb{Z}_\ell(-j)). \end{aligned}$$

The local H^0 lengths are therefore part of the estimate for the kernels $\mathcal{S}_\ell(E, j)$.

For $s_\ell^-(E, j)$, [Alp26] uses generalized Bernoulli values. Write N_E for the conductor of E, the least positive integer with $E \subseteq \mathbb{Q}(\zeta_{N_E})$. The resulting estimate is

$$B_- \leq 3[E : \mathbb{Q}](h')^2 \log(h' N_E) + C \sum_{\substack{\ell \mid [A_\psi : \mathbb{Q}] \\ \ell \text{ odd}}} \ell [E : \mathbb{Q}]^2 (h')^2 \log(h' \ell N_E)^2. \tag{5.4}$$

The constants in the estimates for B_- and B_+ may depend on the fixed character ψ and the fixed arithmetic data. The use of special values in this estimate recalls Stickelberger's [Sti90] construction of annihilators of class groups. In the bound of [Alp26], generalized Bernoulli values control $s_\ell^-(E, j)$

when ℓ does not divide $[E : \mathbb{Q}]$. Taking the norms of these values down to $\mathbb{Q}$ then bounds the sum of their valuations. At primes dividing $[E : \mathbb{Q}]$, the argument combines annihilators with bounds on the numbers of generators of the corresponding eigenspaces of $\mathcal{S}_\ell(E, j)$. These bounds for the numbers of generators contribute the second term in Equation 5.4. Every odd prime dividing $[E : \mathbb{Q}]$ already divides $[A_\psi : \mathbb{Q}]$, since the cyclotomic and quadratic extensions adjoined to A_ψ have degrees that are powers of 2.

For $s_\ell^+(E, j)$, [Alp26] instead uses congruences between automorphic forms and deformations of Galois representations. The connection with cohomology is that cohomology groups also measure the obstructions to deforming a representation. For the fixed character ψ, [Alp26] first sets aside a finite set of exceptional odd primes. At the remaining primes, it decomposes the Galois character modules according to their prime-to-ℓ characters, parity, and congruences involving $\ell - 1$ and j. The estimate obtained in [Alp26] for the resulting cohomology groups takes the form

$$B_+ \leq (N_E h')^C + CB_{\mathrm{exc}}, \tag{5.5}$$

where B_{exc} is a common bound for the contribution to B_+ at each of finitely many exceptional cyclotomic layers, after summing over the twists indexed by j. The constant C includes the number of these layers.

It is worth pausing at Equation 5.5, since its proof in [Alp26] is one of the substantial arithmetic steps in the argument. The bound uses deformations of Galois representations and congruences between automorphic forms on unitary groups in three variables and on GSp_4, together with automorphy lifting and the associated local and global Galois representations. Here GSp_4 is the group of symplectic similitudes in four variables. The representations in Section 4.2 were obtained from automorphic representations. In bounding B_+, [Alp26] also uses automorphy lifting to recognize suitable deformations of Galois representations as arising from automorphic forms. The historical starting point was the problem of lifting a residual Galois representation with prescribed local behavior. Mazur [Maz89] formulated this problem using deformation rings, with tangent spaces and obstructions described by Galois cohomology. In his proof of modularity for semistable elliptic curves, Wiles [Wil95] compared deformation rings with Hecke algebras, with Selmer groups measuring the Galois side of the comparison. The work of Taylor–Wiles [TW95] established the needed complete-intersection property of the Hecke algebras by varying auxiliary primes. Thus the attempt to show that Galois representations come from modular forms also gave a method for relating the size of a cohomology group to congruences between modular forms.

For higher-dimensional Galois representations, the comparison between deformation rings and Hecke algebras was developed by Clozel–Harris–Taylor [CHT08], who extended the method using automorphic forms on unitary groups. For GSp_4, Weissauer [Wei05] had constructed four-dimensional

Galois representations associated to certain cuspidal automorphic representations of GSp_4 over $\mathbb{Q}$. In applying these methods, [Alp26] keeps track of the varying fields and weights. The dependence recorded in Equation 5.5 is needed because the bound on B_+ will be used throughout an interval of values of x.

The remaining terms are B_{exc}, from the exceptional cyclotomic layers, and B_2, from the prime 2. To bound them, [Alp26] uses the freedom to choose the starting point Y in Proposition 5.2.

5.5. **Choosing intervals for the exceptional Tate twists.** To bound the contributions from the exceptional cyclotomic layers and from $\ell = 2$, [Alp26] studies how the lengths $s_\ell(E, j)$ vary through cyclotomic towers.

The variation through a tower brings the argument back to the problem that originally motivated Iwasawa theory, namely the growth of class groups as the field runs up a tower. Iwasawa [Iwa59] studied that growth by treating inverse limits of class groups as modules with a continuous Galois action. For abelian fields, Ferrero–Washington [FW79] proved that the Iwasawa μ-invariant vanishes. In the eventual formula for the ℓ-adic valuation of the class number, μ is the coefficient of ℓ^ν at the layer of degree ℓ^ν over the base field. The argument in [Alp26] uses this vanishing to estimate the lengths $s_\ell(E, j)$, keeping track of the dependence on the base field as the auxiliary primes and the cyclotomic layers vary.

To control the exceptional indices, [Alp26] uses the following estimate along cyclotomic towers. Let L_0 be an abelian number field containing the ℓ-th roots of unity, and also i if $\ell = 2$. Write L_ν for the layer of $L_0(\mu_{\ell^\infty})/L_0$ with $[L_\nu : L_0] = \ell^\nu$. Here μ_{ℓ^∞} denotes all roots of unity of ℓ-power order. There are a positive integer R, at most R elements s_i in an extension of $\mathbb{Q}_\ell$, and constants j_0, c_0 such that, for every $B \in \mathbb{Z}_{\geq 0}$, if $|s_i - j|_\ell \geq \ell^{-B}$ for every i, then

$$s_\ell(L_\nu, j) \leq R(\max(\nu, j_0) + c_0 + B).$$

The values s_i are the finitely many ℓ-adic parameters that the integer j has to avoid for this estimate to apply. The number B sets the lower bound ℓ^{-B} on the distances from these parameters. For $\ell = 2$, the base field in [Alp26] is $L_0 = A_\psi(i, \sqrt{-q}, \sqrt{-d})$. For this base field, [Alp26] bounds $s_2(E(\psi, a, q, d), j)$ by $1 + s_2(L_0(\zeta_{2^{a+2}}), j)$. The field $L_0(\zeta_{2^{a+2}})$ is a layer L_ν with $\nu \leq a + 2$. For each exceptional odd prime ℓ, the base field is $L_0 = \mathbb{Q}(\zeta_{\ell N_\psi})$, and the estimate uses the fixed layer $\nu = \lfloor \log_\ell R \rfloor + 3$. The number R is independent of the layer ν. For $\ell = 2$, [Alp26] also obtains a choice of R independent of q and d, using the congruences imposed on these primes in Section 5.3.

Although the parameters s_i are ℓ-adic, the indices j of the Tate twists are ordinary integers. A bounded interval of integers can therefore contain only a few indices j that are very close to some s_i in the ℓ-adic metric. To see why, fix $Z \geq 1$. For a fixed s_i, the condition $|s_i - j|_\ell < \ell^{-\lfloor \log_\ell Z \rfloor}$ can hold for at most one integer $1 \leq j \leq Z$. Indeed, if two distinct integers

satisfy the condition, then the ultrametric inequality forces their difference to have ℓ-adic valuation greater than $\lfloor \log_\ell Z \rfloor$. This is impossible for a nonzero integer of ordinary absolute value at most Z. Thus there are at most R exceptional indices j between 1 and Z.

At the exceptional indices, [Alp26] has a much coarser bound for $s_\ell(E, j)$, which suffices only when j is small relative to x. It therefore chooses the interval of values of x so that every exceptional index in the range of Proposition 5.2 lies below a small fixed power of x. For a fixed upper endpoint X, every required index lies below X^{401}. It is the exceptional indices in this finite range that enter the following choice of interval. Here is the elementary observation behind that choice of interval. Let $0 < \eta \le C_0$, let $R \in \mathbb{Z}_{\ge 0}$, and let $\mathcal{E} \subseteq \mathbb{N}$ have cardinality at most R. Set $J := 2^{R+2}(C_0/\eta)^R$.

Lemma 5.4. *For every $X > 1$, there is $Y \ge X^{1/J}$ with $Y^2 \le X$ such that, for every $x \in [Y, Y^2]$ and $j \in \mathcal{E}$,*

$$j \le x^{C_0} \quad \Longrightarrow \quad j \le x^{\eta}.$$

Proof. Let $L := \log X$. For $0 \le i \le R$, set $u_i := (L/2)(\eta/(2C_0))^i$. The $R+1$ intervals $(\eta u_i, 2C_0 u_i]$ are pairwise disjoint, since $2C_0 u_{i+1} = \eta u_i$ for $0 \le i < R$. At least one of these intervals contains no point $\log j$ with $j \in \mathcal{E}$. Choose such an i and set $Y := e^{u_i}$. Since $u_i \le L/2$ and

$$u_i \ge \frac{L}{2^{R+1}(C_0/\eta)^R} \ge L/J,$$

the required bounds on Y hold. If $x \in [Y, Y^2]$ and $x^\eta < j \le x^{C_0}$, then $\eta u_i < \log j \le 2C_0 u_i$, contrary to the choice of i. □

For the exceptional indices $j \le x^\eta$, [Alp26] uses a coarse bound for the order of H^2: for suitable totally imaginary fields L containing the required roots of unity,

$$\log \#H^2(G_{L,S_\ell}, \mathbb{Z}_\ell(j+1)) \le \exp\Big((2j[L:\mathbb{Q}])^{C_K} \log(2|\Delta_L|)\Big),$$

Here S_ℓ denotes the places of L above ℓ, Δ_L is the discriminant of L, and C_K is fixed. The restriction $j \le x^\eta$ is useful here because j occurs in the exponent. With η sufficiently small, [Alp26] can keep the sum of these coarse bounds within the estimate sought in Proposition 5.2. For the other indices, the distances from the parameters s_i are large enough to use the bound for $s_\ell(L_v, j)$ that is linear in B, with B logarithmic in the largest index under consideration. Consequently, for the cutoff $h' = \lfloor x^{401} \rfloor$, the contribution B_2 at $\ell = 2$ and the contribution B_{exc} from exceptional cyclotomic layers satisfy the following bounds on the same interval:

$$B_2 \le \exp(x^{4\varepsilon/5}) \qquad \text{and} \qquad B_{\mathrm{exc}} \le \exp(x^{4\varepsilon/5}).$$

The point of choosing Y is therefore to arrange that the coarse estimate is applied only to indices $j \le x^\eta$. For every other index in the required range, [Alp26] uses the estimate linear in B.

5.6. **Recovering the required bound for cohomology.** To recover Proposition 5.2, [Alp26] combines the bounds for B_-, B_+, and B_2. For the application in [Alp26], take $E := E(\psi, a(x), q, d)$ and $h' := \lfloor x^{401} \rfloor$. Since $N_E \leq 2^{a(x)+2} N_\psi q d$, the bounds on $a(x)$, q, d, and h' imply $\log(N_E h') \leq 2x^{\varepsilon/4}$ for sufficiently large x in the selected interval. Moreover, $[E : \mathbb{Q}] \leq \varphi(N_E) \leq N_E$. Here φ is Euler's totient function. The right side of Equation 5.4 is therefore at most $C_1 \exp(8x^{\varepsilon/4})$ for a fixed C_1. Combining this estimate with Equations (5.3) and (5.5) gives

$$\begin{aligned}\mathcal{V}(E, h') &\leq \exp(x^{4\varepsilon/5}) \log 2 + C_1 \exp(8x^{\varepsilon/4}) + (N_E h')^C + C \exp(x^{4\varepsilon/5}) \\ &\leq \exp(x^{4\varepsilon/5}) \log 2 + C_1 \exp(8x^{\varepsilon/4}) + \exp(2Cx^{\varepsilon/4}) + C \exp(x^{4\varepsilon/5}).\end{aligned}$$

Each exponent is $o(x^\varepsilon)$. For fixed $C, C_1, \varepsilon > 0$ and sufficiently large x, the sum is therefore at most $\exp(\frac{1}{2}x^\varepsilon)$. This is Equation 5.2 in Proposition 5.2.

This is the bound for locally trivial extension classes that [Alp26] uses in comparing the integral lattices of automorphic forms. Together with the other local and global estimates for their indices, it bounds $\overline{L}$ in Equation 4.10 and puts the denominators within the error allowed in Equation 4.11. The remaining choice of parameters, including the number of theta lifts, is made in Section 6. With those choices, [Alp26] obtains a bound for $\mathcal{G}_\psi(x)$ with exponent as close to $1/2$ as required.

6. Choosing the parameters to obtain GRH

At this point, [Alp26] has an upper bound for $\mathcal{A}_{\mathrm{tot}} - \mathcal{M}_{\mathrm{tot}}$ from estimates for automorphic periods and integral lattices of automorphic forms. Recall the real linear combination $\mathcal{G}_\psi$ of corrected sums over prime powers in Equation 1.1. The expansion in characters from Section 3 expresses the same quantity in terms of $\mathcal{G}_\psi(x)$. To obtain GRH, [Alp26] chooses the parameters so that this inequality bounds $\mathcal{G}_\psi(x)$ with exponent arbitrarily close to $1/2$.

Increasing the number m of theta lifts has two competing effects on the bound. After division by the coefficient of $\mathcal{G}_\psi(x)$, two of the terms have sizes x/m and m, up to fixed powers of x^ε. Increasing m decreases x/m and increases m. Taking m close to $\sqrt{x}$ keeps both terms close to $\sqrt{x}$, which explains the exponent $1/2$. The remaining terms depend on the dimension of the initial Hermitian space and the sizes of the selected sets of complex places of E. In [Alp26], those parameters are chosen together with m to bound every remaining term by $x^{1/2}$ times a fixed power of x^ε.

6.1. **The number of lifts and the size of the error.** Recall the bound in Equation 4.11. With H, T, and T_2 as in Section 4.6, put $n = |H|$, $t = |T|$, and $\nu_2 = |T_2|$, while m is the number of theta lifts. The number b is the scale of the weights at the complex places of E in T_2, as in Section 3. The parameter $k = \dim_E X_0$ is the dimension of the initial Hermitian space, and a bounds the positive parameters a_t indexed by the places in T. The estimates in [Alp26] allow multiplicative factors that are fixed powers of x^ε, so the parameters

are chosen within ranges. The elementary consequence of those ranges is given in Lemma 6.1. The simultaneous choices in [Alp26] are recorded in Proposition 6.3.

Fix $x \geq 3$, $\varepsilon > 0$, and $C \geq 0$. Suppose n, t, ν_2, m are positive integers, $b, k \in \mathbb{R}_{>0}$, and $a \in \mathbb{R}_{\geq 0}$ such that

$$\begin{aligned} a &\leq 2xb, \\ \frac{n}{2}x^{-\varepsilon/8} &\leq t, \\ nx^{-\varepsilon} &\leq \nu_2 \leq n, \\ x^{1/2-\varepsilon} &\leq m \leq x^{1/2+\varepsilon/2}, \\ xm &\leq k \leq n^2 x^{1/2-\varepsilon}. \end{aligned}$$

Write

$$B := ta\left(1 + \frac{m^2}{k}\right) + m\left(2nmbx^{\varepsilon} + \frac{\nu_2 kbx^{2\varepsilon}}{n^2}\right)$$

for the expression in brackets in Equation 4.11. The calculation in Section 3 leads to a coefficient proportional to $tmb\nu_2/n$ for $\mathcal{G}_\psi(x)$. Dividing the error by this coefficient is what allows the bound to be compared with $\sqrt{x}$. In Lemma 6.1, Z will represent $\mathcal{G}_\psi(x)$ minus a bound for the contributions from the surviving auxiliary characters.

Lemma 6.1. *Assume the parameter bounds in Section 6.1. If $Z \in \mathbb{R}$ satisfies*

$$\frac{tmb\nu_2}{2n}\left(Z - 4x^{1/2+\varepsilon}\right) \leq x^{C\varepsilon}B + \frac{2tmb}{x}, \tag{6.1}$$

then $Z \leq 26x^{1/2+(C+21/8)\varepsilon}$. In particular, if $26 \leq x^{3\varepsilon/8}$, then $Z \leq x^{1/2+(C+3)\varepsilon}$.

Proof. It is convenient to divide first by $Q := tmb$ and to write $d := \nu_2/n$. The effect of increasing m is visible in the summand ta/Q:

$$\frac{ta}{Q} = \frac{a}{mb} \leq \frac{2x}{m} \leq 2x^{1/2+\varepsilon}.$$

The summand $ta(m^2/k)/Q$ stays bounded because $k \geq xm$:

$$\frac{ta(m^2/k)}{Q} = \frac{am}{bk} \leq \frac{2xm}{k} \leq 2.$$

The summand $2nm^2bx^{\varepsilon}/Q$ also involves the ratio $t/n = |T|/|H|$. The lower bound on t/n ensures that

$$\frac{2nm^2bx^{\varepsilon}}{Q} = \frac{2nmx^{\varepsilon}}{t} \leq 4mx^{9\varepsilon/8} \leq 4x^{1/2+13\varepsilon/8}.$$

For the last term of B/Q, the upper bound on k is needed as well:

$$\frac{m\nu_2 kbx^{2\varepsilon}/n^2}{Q} = \frac{\nu_2 kx^{2\varepsilon}}{n^2 t} \leq \frac{2k}{n^2}x^{17\varepsilon/8} \leq 2x^{1/2+9\varepsilon/8}.$$

Since $x \geq 1$, the sum of the four terms is bounded by

$$\frac{B}{Q} \leq 10x^{1/2+13\varepsilon/8}.$$

Dividing Equation 6.1 by $Q/2 > 0$ and using the bound for B/Q gives

$$d\Big(Z - 4x^{1/2+\varepsilon}\Big) \leq 20x^{1/2+(C+13/8)\varepsilon} + \frac{4}{x} \leq 22x^{1/2+(C+13/8)\varepsilon}.$$

The last inequality uses $x \geq 3$ and $x^{1/2+(C+13/8)\varepsilon} \geq 1$. Since $d \geq x^{-\varepsilon} > 0$, it follows that

$$Z \leq 4x^{1/2+\varepsilon} + 22x^{1/2+(C+21/8)\varepsilon} \leq 26x^{1/2+(C+21/8)\varepsilon}.$$

If $26 \leq x^{3\varepsilon/8}$, the factor 26 is absorbed by $x^{3\varepsilon/8}$. □

Remark 6.2. Notice how the parameter bounds in Lemma 6.1 enter its proof. The bounds on the number m of theta lifts keep the terms x/m and m close to $\sqrt{x}$. The bounds on $k = \dim_E X_0$ control the two terms involving k in B/Q. The lower bounds on t/n and ν_2/n ensure that division by the coefficient of Z introduces at most a fixed power of x^{ε}. To satisfy these conditions simultaneously, [Alp26] chooses the Hermitian spaces, highest weights, and selected complex places of E together.

6.2. **Applying the bound to sums over prime powers.** The argument about zeros in Section 2 uses a bound for $\mathcal{G}_\psi(x)$ throughout arbitrarily large intervals $[Y, Y^{3/2}]$. For each such interval, [Alp26] arranges that the number fields, weights, and selected complex places satisfy the hypotheses of Lemma 6.1 at every value of x.

Before applying the lemma, there is one more term to include. The expansion in characters of $\mathcal{A}_{\mathrm{tot}} - \mathcal{M}_{\mathrm{tot}}$ contains surviving auxiliary characters besides those defining $\mathcal{G}_\psi$. For the choices in [Alp26], the sum of the absolute values of these additional contributions is at most $4rx^{1/2+\varepsilon}$, where r is the order of ψ. The following proposition records the uniform choices in [Alp26], with the term $4rx^{1/2+\varepsilon}$ subtracted in the definition of Z.

Proposition 6.3. *Let ψ be a primitive Dirichlet character of order r. There is $C \geq 0$ such that, for every sufficiently small $\varepsilon > 0$, there are arbitrarily large $Y \geq 3$ for which the hypotheses of Lemma 6.1 can be satisfied at every $x \in [Y, Y^{3/2}]$ with this C and*

$$Z = \mathcal{G}_\psi(x) - 4rx^{1/2+\varepsilon}.$$

The constant C in Proposition 6.3 accounts for the errors in the estimates for period values and in the comparisons between integral lattices of automorphic forms. It is fixed before ε is chosen; the remaining parameters may depend on x.

Proof of Proposition 1.2. Fix ψ, $\eta > 0$, and $A \geq 2$. Let r be the order of ψ, and choose C as in Proposition 6.3. Choose $\varepsilon > 0$ small enough that

the proposition applies and $(C+4)\varepsilon \le \eta$. Since Proposition 6.3 permits arbitrarily large Y, choose one with

$$Y \ge \max\Big\{A, 3, 26^{8/(3\varepsilon)}, (1+4r)^{1/\varepsilon}\Big\}.$$

For each $x \in [Y, Y^{3/2}]$, choose the parameters in Proposition 6.3 with the same C. The choice of Y ensures that $26 \le x^{3\varepsilon/8}$, so Lemma 6.1 gives

$$\mathcal{G}_\psi(x) \le x^{1/2+(C+3)\varepsilon} + 4rx^{1/2+\varepsilon}.$$

Since $C+3 \ge 1$ and $1+4r \le x^{\varepsilon}$, it follows that

$$\mathcal{G}_\psi(x) \le (1+4r)x^{1/2+(C+3)\varepsilon} \le x^{1/2+(C+4)\varepsilon} \le x^{1/2+\eta}.$$

This holds throughout the chosen interval, with $Y \ge A$. □

With the interval bound of Proposition 1.2 in place, [Alp26] applies Proposition 2.3 to exclude zeros of the Dirichlet L-functions to the right of the critical line. The functional equation then excludes zeros to the left, completing the proposed argument for Conjecture 1.1.

References

[AB95] Ahmed Abbes and Thierry Bouche. Théorème de Hilbert–Samuel "arithmétique". *Ann. Inst. Fourier (Grenoble)*, 45(2):375–401, 1995.

[ALB23] Johannes Anschütz and Arthur-César Le Bras. Prismatic Dieudonné theory. *Forum Math. Pi*, 11:e2, 1–92, 2023. Erratum, *Forum Math. Pi* **13** (2025), e3.

[Alp26] Levent Alpöge. [Announcement concerning the Riemann hypothesis]. `https://alpo.ge/remark.pdf`, 2026. Announcement of an argument developed with Claude. The Lean formalization is available in fourteen archive parts, to be concatenated in the indicated order: part01, part02, part03, part04, part05, part06, part07, part08, part09, part10, part11, part12, part13, part14. Accessed November 20, 2026.

[AMRT75] Avner Ash, David Mumford, Michael Rapoport, and Yung-Sheng Tai. *Smooth Compactification of Locally Symmetric Varieties*, volume 4 of *Lie Groups: History, Frontiers and Applications*. Math Sci Press, Brookline, Massachusetts, 1975. Second edition, as *Smooth Compactifications of Locally Symmetric Varieties*, Cambridge University Press, 2010, with the collaboration of Peter Scholze.

[And82] Greg W. Anderson. Logarithmic derivatives of Dirichlet L-functions and the periods of abelian varieties. *Compos. Math.*, 45(3):315–332, 1982.

[AS95] Urban Angehrn and Yum-Tong Siu. Effective freeness and point separation for adjoint bundles. *Invent. Math.*, 122(2):291–308, 1995.

[Beh46] Felix A. Behrend. On sets of integers which contain no three terms in arithmetical progression. *Proc. Natl. Acad. Sci. USA*, 32(12):331–332, 1946.

[BF03] John J. Benedetto and Matthew Fickus. Finite normalized tight frames. *Adv. Comput. Math.*, 18(2–4):357–385, 2003.

[BK90] Spencer Bloch and Kazuya Kato. L-functions and Tamagawa numbers of motives. In *The Grothendieck Festschrift, Vol. I*, volume 86 of *Progress in Mathematics*, pages 333–400. Birkhäuser Boston, Boston, MA, 1990.

[BLGGT14] Thomas Barnet-Lamb, Toby Gee, David Geraghty, and Richard Taylor. Local-global compatibility for $l = p$, II. *Ann. Sci. Éc. Norm. Supér. (4)*, 47(1):165–179, 2014.

[BMS18] Bhargav Bhatt, Matthew Morrow, and Peter Scholze. Integral $\mathfrak{p}$-adic Hodge theory. *Publ. Math. Inst. Hautes Études Sci.*, 128:219–397, 2018.

[Bos01] Jean-Benoît Bost. Algebraic leaves of algebraic foliations over number fields. *Publ. Math. Inst. Hautes Études Sci.*, 93:161–221, 2001.

[BS22] Bhargav Bhatt and Peter Scholze. Prisms and prismatic cohomology. *Ann. of Math. (2)*, 196(3):1135–1275, 2022.

[BV89] Jean-Michel Bismut and Éric Vasserot. The asymptotics of the Ray–Singer analytic torsion associated with high powers of a positive line bundle. *Comm. Math. Phys.*, 125(2):355–367, 1989.

[BZ77] I. N. Bernstein and A. V. Zelevinsky. Induced representations of reductive $\mathfrak{p}$-adic groups. I. *Ann. Sci. Éc. Norm. Supér. (4)*, 10(4):441–472, 1977.

[Car12] Ana Caraiani. Local-global compatibility and the action of monodromy on nearby cycles. *Duke Math. J.*, 161(12):2311–2413, 2012.

[Car14] Ana Caraiani. Monodromy and local-global compatibility for $\mathfrak{l} = \mathfrak{p}$. *Algebra Number Theory*, 8(7):1597–1646, 2014.

[CH13] Gaëtan Chenevier and Michael Harris. Construction of automorphic Galois representations, II. *Camb. J. Math.*, 1(1):53–73, 2013.

[CHT08] Laurent Clozel, Michael Harris, and Richard Taylor. Automorphy for some $\mathfrak{l}$-adic lifts of automorphic mod $\mathfrak{l}$ Galois representations. *Publ. Math. Inst. Hautes Études Sci.*, 108:1–181, 2008.

[Clo91] Laurent Clozel. Représentations galoisiennes associées aux représentations automorphes autoduales de GL(n). *Publ. Math. Inst. Hautes Études Sci.*, 73:97–145, 1991.

[CS49] Sarvadaman Chowla and Atle Selberg. On Epstein's zeta function. I. *Proc. Natl. Acad. Sci. USA*, 35(7):371–374, 1949.

[Dir37] Peter Gustav Lejeune Dirichlet. Beweis des Satzes, daß jede unbegrenzte arithmetische Progression, deren erstes Glied und Differenz ganze Zahlen ohne gemeinschaftlichen Factor sind, unendlich viele Primzahlen enthält. *Abh. Königl. Preuss. Akad. Wiss. Berlin*, pages 45–71, 1837.

[DR80] Pierre Deligne and Kenneth A. Ribet. Values of abelian L-functions at negative integers over totally real fields. *Invent. Math.*, 59(3):227–286, 1980.

[Far11] Laurent Fargues. La filtration canonique des points de torsion des groupes $\mathfrak{p}$-divisibles. *Ann. Sci. Éc. Norm. Supér. (4)*, 44(6):905–961, 2011. Avec la collaboration de Yichao Tian.

[FC90] Gerd Faltings and Ching-Li Chai. *Degeneration of Abelian Varieties*, volume 22 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3)*. Springer-Verlag, Berlin, 1990.

[FW79] Bruce Ferrero and Lawrence C. Washington. The Iwasawa invariant μ_p vanishes for abelian number fields. *Ann. of Math. (2)*, 109(2):377–395, 1979.

[Gan23] Wee Teck Gan. Automorphic forms and the theta correspondence. arXiv:2303.14918, 2023.

[GGP12] Wee Teck Gan, Benedict H. Gross, and Dipendra Prasad. Symplectic local root numbers, central critical L-values, and restriction problems in the representation theory of classical groups. In *Sur les conjectures de Gross et Prasad. I*, number 346 in Astérisque, pages 1–109. Société Mathématique de France, Paris, 2012.

[GK75] I. M. Gelfand and D. A. Kazhdan. Representations of the group $\mathrm{GL}(n, K)$ where K is a local field. In I. M. Gelfand, editor, *Lie Groups and Their Representations*, pages 95–118. Halsted Press, New York, 1975.

[GQT14] Wee Teck Gan, Yannan Qiu, and Shuichiro Takeda. The regularized Siegel–Weil formula (the second term identity) and the Rallis inner product formula. *Invent. Math.*, 198(3):739–831, 2014.

[GR13] Éric Gaudron and Gaël Rémond. Minima, pentes et algèbre tensorielle. *Israel J. Math.*, 195:565–591, 2013.

[Gra84] Daniel R. Grayson. Reduction theory using semistability. *Comment. Math. Helv.*, 59(4):600–634, 1984.

[GS88] Henri Gillet and Christophe Soulé. Amplitude arithmétique. *C. R. Acad. Sci. Paris Sér. I Math.*, 307:887–890, 1988.

[Gui48] A. P. Guinand. A summation formula in the theory of prime numbers. *Proc. London Math. Soc. (2)*, 50:107–119, 1948.

[Had96] Jacques Hadamard. Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques. *Bulletin de la Société Mathématique de France*, 24:199–220, 1896.

[Har26a] Michael Harris. Mathematics as gift economy. Silicon Reckoner, November 20, 2026. `https://siliconreckoner.substack.com/p/mathematics-as-gift-economy`.

[Har26b] Michael Harris. Now all mathematicians can be Mechanical Turks. Silicon Reckoner, October 8, 2026. `https://siliconreckoner.substack.com/p/now-all-mathematicians-can-be-mechanical`.

[Hen00] Guy Henniart. Une preuve simple des conjectures de Langlands pour GL(n) sur un corps p-adique. *Invent. Math.*, 139(2):439–455, 2000.

[Hid81] Haruzo Hida. Congruences of cusp forms and special values of their zeta functions. *Invent. Math.*, 63(2):225–261, 1981.

[Hir64] Heisuke Hironaka. Resolution of singularities of an algebraic variety over a field of characteristic zero. I, II. *Ann. of Math. (2)*, 79:109–203, 205–326, 1964.

[HN75] Günter Harder and Mudumbai Seshachalu Narasimhan. On the cohomology groups of moduli spaces of vector bundles on curves. *Math. Ann.*, 212(3):215–248, 1975.

[How79] Roger Howe. θ-series and invariant theory. In Armand Borel and William Casselman, editors, *Automorphic Forms, Representations and* L*-Functions, Part 1*, volume 33 of *Proceedings of Symposia in Pure Mathematics*, pages 275–285. American Mathematical Society, Providence, RI, 1979.

[HT01] Michael Harris and Richard Taylor. *The geometry and cohomology of some simple Shimura varieties*, volume 151 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2001. With an appendix by Vladimir G. Berkovich.

[Iwa59] Kenkichi Iwasawa. On Γ-extensions of algebraic number fields. *Bull. Amer. Math. Soc.*, 65(4):183–226, 1959.

[JS81a] Hervé Jacquet and Joseph A. Shalika. On Euler products and the classification of automorphic representations. I. *Amer. J. Math.*, 103(3):499–558, 1981.

[JS81b] Hervé Jacquet and Joseph A. Shalika. On Euler products and the classification of automorphic forms. II. *Amer. J. Math.*, 103(4):777–815, 1981.

[Kaw82] Yujiro Kawamata. A generalization of Kodaira–Ramanujam's vanishing theorem. *Math. Ann.*, 261(1):43–46, 1982.

[Kot92] Robert E. Kottwitz. On the λ-adic representations associated to some simple Shimura varieties. *Invent. Math.*, 108(3):653–665, 1992.

[Kud84] Stephen S. Kudla. Seesaw dual reductive pairs. In Ichirô Satake and Yasuo Morita, editors, *Automorphic Forms of Several Variables (Katata, 1983)*, volume 46 of *Progress in Mathematics*, pages 244–268. Birkhäuser, Boston, MA, 1984.

[Kud94] Stephen S. Kudla. Splitting metaplectic covers of dual reductive pairs. *Israel J. Math.*, 87(1–3):361–401, 1994.

[Lab11] Jean-Pierre Labesse. Changement de base CM et séries discrètes. In Laurent Clozel, Michael Harris, Jean-Pierre Labesse, and Bao Châu Ngô, editors, *On*

the Stabilization of the Trace Formula, volume 1, pages 429–470. International Press, Somerville, MA, 2011.

[Lan05] Edmund Landau. Über einen Satz von Tschebyschef. *Math. Ann.*, 61:527–550, 1905.

[Lic73] Stephen Lichtenbaum. Values of zeta-functions, étale cohomology, and algebraic K-theory. In Hyman Bass, editor, *Algebraic* K*-Theory, II: "Classical" Algebraic* K*-Theory, and Connections with Arithmetic*, volume 342 of *Lecture Notes in Mathematics*, pages 489–501. Springer, Berlin, 1973.

[Mas26] Alessandro Maria Masullo. The Music of the Primes. `https://math216.wordpress.com/wp-content/uploads/2026/11/the-music-of-the-primes.pdf`, 2026. Draft released November 20, 2026. Announced by Ravi Vakil on November 20, 2026, on the Math 216 course website: `https://math216.wordpress.com/2026/11/20/hurried-announcement-alessandro-masullo-grh/`.

[Maz89] Barry Mazur. Deforming Galois representations. In Yasutaka Ihara, Kenneth A. Ribet, and Jean-Pierre Serre, editors, *Galois Groups over* $\mathbf{Q}$, volume 16 of *Mathematical Sciences Research Institute Publications*, pages 385–437. Springer, New York, 1989.

[Mes72] William Messing. *The Crystals Associated to Barsotti–Tate Groups: With Applications to Abelian Schemes*, volume 264 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1972.

[Mil07] James S. Milne. The fundamental theorem of complex multiplication. arXiv:0705.3446, 2007.

[Mor09] Atsushi Moriwaki. Continuity of volumes on arithmetic varieties. *J. Algebraic Geom.*, 18(3):407–457, 2009.

[Mum77] David Mumford. Hirzebruch's proportionality theorem in the non-compact case. *Invent. Math.*, 42:239–272, 1977.

[MW89] Colette Mœglin and Jean-Loup Waldspurger. Le spectre résiduel de $\mathrm{GL}(\mathfrak{n})$. *Ann. Sci. École Norm. Supér. (4)*, 22(4):605–674, 1989.

[Nad90] Alan Michael Nadel. Multiplier ideal sheaves and Kähler–Einstein metrics of positive scalar curvature. *Ann. of Math. (2)*, 132(3):549–596, 1990.

[Pil84] Adolf Piltz. *Über die Häufigkeit der Primzahlen in arithmetischen Progressionen und über verwandte Gesetze*. A. Neuenhahn, Jena, 1884. Habilitationsschrift, Universität Jena.

[Poi67] Georges Poitou, editor. *Cohomologie galoisienne des modules finis*. Number 13 in Travaux et Recherches Mathématiques. Dunod, Paris, 1967. Séminaire de l'Institut de mathématiques de Lille, 1962–1963.

[Ral87] Stephen Rallis. L*-Functions and the Oscillator Representation*, volume 1245 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1987.

[Rie59] Bernhard Riemann. Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse. *Monatsber. Königl. Preuss. Akad. Wiss. Berlin*, pages 671–680, 1859.

[Sch73] Wolfgang M. Schmidt. Zur Methode von Stepanov. *Acta Arith.*, 24(4):347–367, 1973.

[Seg59] Irving Ezra Segal. Foundations of the theory of dynamical systems of infinitely many degrees of freedom. I. *Mat.-Fys. Medd. Danske Vid. Selsk.*, 31(12):1–39, 1959.

[Seg63] Irving Ezra Segal. Transforms for operators and symplectic automorphisms over a locally compact abelian group. *Math. Scand.*, 13:31–43, 1963.

[Sha62] David Shale. Linear symmetries of free boson fields. *Trans. Amer. Math. Soc.*, 103(1):149–167, 1962.

[Shi11] Sug Woo Shin. Galois representations arising from some compact Shimura varieties. *Ann. of Math. (2)*, 173(3):1645–1741, 2011.

[Sie35] Carl Ludwig Siegel. Über die analytische Theorie der quadratischen Formen. *Ann. of Math. (2)*, 36(3):527–606, 1935.

[Siu93] Yum-Tong Siu. An effective Matsusaka big theorem. *Ann. Inst. Fourier (Grenoble)*, 43(5):1387–1405, 1993.

[SS42] Raphaël Salem and Donald C. Spencer. On sets of integers which contain no three terms in arithmetical progression. *Proc. Natl. Acad. Sci. USA*, 28(12):561–563, 1942.

[ST61] Goro Shimura and Yutaka Taniyama. *Complex Multiplication of Abelian Varieties and Its Applications to Number Theory*, volume 6 of *Publications of the Mathematical Society of Japan*. Mathematical Society of Japan, Tokyo, 1961.

[Sta26] Stark units for real quadratic fields and reciprocity laws, 4 October 2026. Available at `https://www.danrad.net/papers/stark-reciprocity.pdf`.

[Ste69] S. A. Stepanov. On the number of points of a hyperelliptic curve over a finite prime field. *Math. USSR-Izv.*, 3(5):1103–1114, 1969.

[Sti90] Ludwig Stickelberger. Ueber eine Verallgemeinerung der Kreistheilung. *Math. Ann.*, 37(3):321–367, 1890.

[Stu76] Ulrich Stuhler. Eine Bemerkung zur Reduktionstheorie quadratischer Formen. *Arch. Math. (Basel)*, 27(6):604–610, 1976.

[Tat63] John Tate. Duality theorems in Galois cohomology over number fields. In *Proceedings of the International Congress of Mathematicians (Stockholm, 1962)*, pages 288–295, Djursholm, 1963. Institut Mittag-Leffler.

[Tur49] Pál Turán. On a new method in the analysis with applications. *Časopis pro pěstování matematiky a fysiky*, 74(2):123–131, 1949.

[Tur50] Paul Turán. On the remainder-term of the prime-number formula, I. *Acta Math. Acad. Sci. Hungar.*, 1:48–63, 1950.

[TW95] Richard Taylor and Andrew Wiles. Ring-theoretic properties of certain Hecke algebras. *Ann. of Math. (2)*, 141(3):553–572, 1995.

[TY07] Richard Taylor and Teruyoshi Yoshida. Compatibility of local and global Langlands correspondences. *J. Amer. Math. Soc.*, 20(2):467–493, 2007.

[vH51] Léon van Hove. Sur certaines représentations unitaires d'un groupe infini de transformations. *Acad. Roy. Belgique Cl. Sci. Mém. Coll. in-8° (2)*, 26(6):1–102, 1951.

[Vie82] Eckart Viehweg. Vanishing theorems. *J. Reine Angew. Math.*, 335:1–8, 1982.

[Wei52] André Weil. Sur les "formules explicites" de la théorie des nombres premiers. *Communications du Séminaire Mathématique de l'Université de Lund*, pages 252–265, 1952. Tome supplémentaire.

[Wei64] André Weil. Sur certains groupes d'opérateurs unitaires. *Acta Math.*, 111:143–211, 1964.

[Wei72] André Weil. Sur les formules explicites de théorie des nombres. *Math. USSR-Izv.*, 6(1):1–17, 1972.

[Wei05] Rainer Weissauer. Four dimensional Galois representations. In Jacques Tilouine, Henri Carayol, Michael Harris, and Marie-France Vignéras, editors, *Formes automorphes (II): le cas du groupe* GSp(4), volume 302 of *Astérisque*, pages 67–150. Société Mathématique de France, 2005.

[Wey46] Hermann Weyl. *The Classical Groups: Their Invariants and Representations*, volume 1 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, second edition, 1946. First published in 1939.

[Wil95] Andrew Wiles. Modular elliptic curves and Fermat's last theorem. *Ann. of Math. (2)*, 141(3):443–551, 1995.

[Yua08] Xinyi Yuan. Big line bundles over arithmetic varieties. *Invent. Math.*, 173:603–649, 2008.

[Zha95] Shouwu Zhang. Positive line bundles on arithmetic varieties. *J. Amer. Math. Soc.*, 8(1):187–221, 1995.

Institut des Hautes Études Scientifiques
Email address: rzhang@ihes.fr