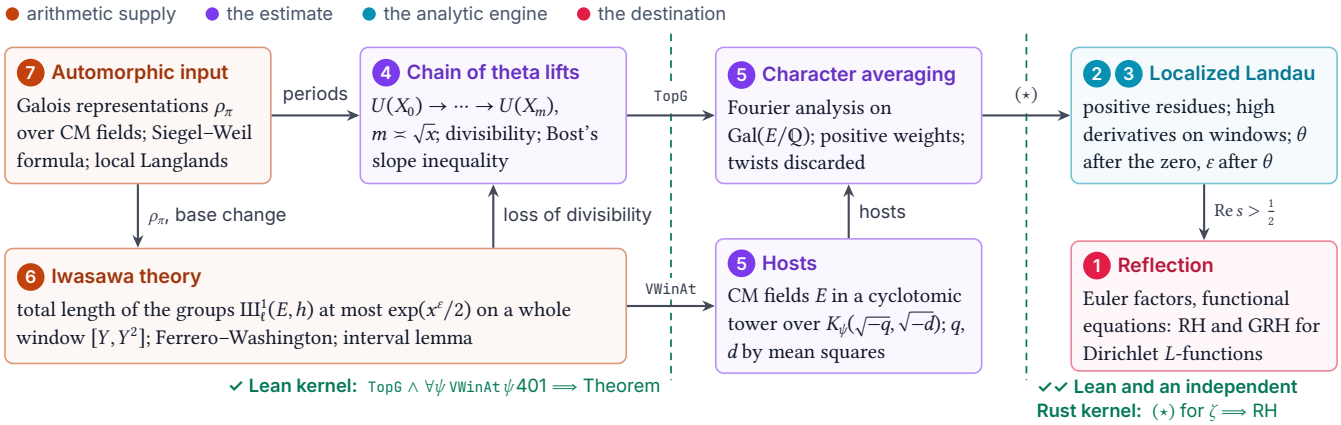
The proof of the Riemann hypothesis announced by L. Alpöge

Its architecture: the statements, the flow, and why each idea calls for the next

One hard estimate and two classical transformations of it. Arithmetic geometry bounds a smoothed sum over primes *from above*, by a square root, on sparse but long intervals. A localized form of Landau's theorem turns this one-sided, sparse information into a zero-free half-plane, and functional equations carry the half-plane to the critical line. The text reads the proof backwards from the theorem; the diagram shows the same flow forwards.



The flow of the argument; circled numbers refer to the sections below. Everything to the right of a dashed line has been machine-checked as an implication from the statements that cross the line.

Theorem. *Let χ be a Dirichlet character of modulus $N \geq 1$. Then $L(s,\chi) \neq 0$ for $\mathrm{Re}\, s > \frac{1}{2}$, except at the pole $s = 1$ of a principal character. Consequently every zero of $L(s,\chi)$ with $0 < \mathrm{Re}\, s < 1$ lies on the critical line $\mathrm{Re}\, s = \frac{1}{2}$; for $N = 1$, where $L(s,\chi) = \zeta(s)$, this is the Riemann hypothesis.*

LEAN `LFunction_ne_zero_of_half_lt_re` · `LFunction_eq_zero_re` · `riemannHypothesis`, a term of Mathlib's `RiemannHypothesis`

Lines marked LEAN link to the Lean archive, the formal proof; Lean and all notation are explained at the end.

The shape of the argument: Stepanov over $\mathbb{Z}$

RH is square-root cancellation among the primes. For curves over a finite field $\mathbb{F}$, Stepanov obtained it from an auxiliary function vanishing to high order at the rational points: comparing zeros with the degree bounds the number of points from *above*, and upper bounds over all finite extensions of $\mathbb{F}$ suffice.

The announcement calls its argument "a descendant of Stepanov's", each theta lift "mimicking multiplicity", with $\sqrt{X}$ lifts for the von Mangoldt sum of length X.

STEPANOV, FOR CURVES OVER $\mathbb{F}$		THE ANNOUNCED PROOF, OVER $\mathbb{Z}$
vanishing to order m	⟶	divisibility of an m-fold iterated theta lift (§4)
counting degrees	⟶	Bost's slope inequality (§4)
the finite extensions of $\mathbb{F}$	⟶	the layers of a cyclotomic tower of fields (§4)
"upper bounds suffice"	⟶	a theorem of Landau type (§3)

1 Reflection

Let χ be a Dirichlet character modulo N and χ^* the primitive character that induces it. Then

$$L(s,\chi) = L(s,\chi^*)\prod_{p\mid N}\bigl(1-\chi^*(p)\,p^{-s}\bigr),$$

the product running over the primes p dividing N, and no factor of the product vanishes for $\operatorname{Re} s > 0$.

For primitive χ let $\Lambda(s,\chi)$ be the completed L-function, that is, $L(s,\chi)$ multiplied by its Gamma factor and a power of the modulus, and let $W(\chi)$ be the root number, a complex number of absolute value 1. With $\bar\chi$ the complex conjugate character, the functional equation

$$\Lambda(s,\chi) = W(\chi)\,\Lambda(1-s,\bar\chi)$$

sends a zero in the critical strip $0 < \operatorname{Re} s < 1$ with $\operatorname{Re} s < \tfrac12$ to a zero of $L(\,\cdot\,,\bar\chi)$ with $\operatorname{Re} s > \tfrac12$. So it suffices to prove $L(s,\psi) \neq 0$ for $\operatorname{Re} s > \tfrac12$ and every primitive character ψ, the trivial one, for which $L = \zeta$, included.

LEAN all characters `imprimitive` · reflection `Claim.strip` · reflection for ζ `rh_of_right`

2 The statistic

Let $\Lambda(n)$ be the von Mangoldt function, equal to $\log p$ if n is a power of the prime p and to 0 otherwise, and let γ be Euler's constant. For a real number $y \geq 1$ and a primitive character $\chi \neq 1$ put

$$G(y) = \sum_{n\le y}\Lambda(n)\Bigl(\frac{y}{n}-1\Bigr) - y(\log y - 1 - \gamma),$$

$$g_\chi(y) = \sum_{n\le y}\Lambda(n)\,\chi(n)\Bigl(\frac{y}{n}-1\Bigr) + y\,\frac{L'}{L}(1,\chi).$$

The weight $(y/n-1)_+$, where $(u)_+ := \max(u,0)$, is a *ramp*: it vanishes at $n = y$ and grows as n decreases. The last term of each line is the *centring*; it is what makes the transform below regular at $s = 1$.

The *Mellin transform* of a function f on $[1,\infty)$ is the function $s \mapsto \int_1^\infty f(y)\,y^{-s-1}\,dy$. That of the ramp $(y/n-1)_+$ is $n^{-s}/(s(s-1))$, so summation over n gives, for $\operatorname{Re} s > 1$,

$$\int_1^\infty g_\chi(y)\,y^{-s-1}\,dy = -\frac{1}{s(s-1)}\Bigl[\frac{L'}{L}(s,\chi) - s\,\frac{L'}{L}(1,\chi)\Bigr],$$

and a similar formula for G, with ζ in place of $L(\,\cdot\,,\chi)$. These transforms continue to $\operatorname{Re} s > 0$, where their only singularities are simple poles at the zeros ρ of the L-function: if ρ is a zero of multiplicity $m_\chi(\rho)$, the residue is $-m_\chi(\rho)/(\rho(\rho-1))$. Read backwards, this is the explicit formula

$$G(y) = \log 2\pi - \sum_\rho \frac{y^\rho}{\rho(\rho-1)} + O(y^{-2}),$$

in which ρ runs over the zeros of ζ in the critical strip, with multiplicity. Hence RH gives $|G(y) - \log 2\pi| \leq 0.0462\,\sqrt{y} + O(y^{-2})$, the constant being $2 + \gamma - \log 4\pi$. Numerically one finds $|G(y) - \log 2\pi| < 0.023\sqrt{y}$ for $10 \leq y \leq 10^7$.

The proof needs far less: an upper bound, on windows.

LEAN the two sums `Deduction.G` · transform of the ramp `MellinRamp` · poles and residues `poleData`

3 A localized Landau theorem

Landau's theorem says that the Mellin transform of a function of constant sign has a singularity at the real point of its abscissa of convergence. Applied to $Cy^\theta - G(y)$ it shows that a bound $G(y) \le Cy^\theta$ *for all large* y would exclude zeros of ζ with $\operatorname{Re} s > \theta$. The arithmetic of §4–6 yields such a bound only on certain intervals, and only for several characters together. So the theorem has to be localized, and the zeros of different L-functions must be kept from cancelling.

A *positive family* F is a finite set of Dirichlet characters, each either non-principal or the trivial character of modulus one, with weights $c_\chi > 0$. Put

$$G_F(y) = \sum_{\chi \in F} c_\chi \operatorname{Re} g_\chi(y), \qquad \text{with } g_1 := G \text{ for the trivial character.}$$

A *window* is an interval $[Y, Y^{1+\delta}]$ with $\delta > 0$.

Proposition. *Let F be a positive family, $0 < \theta < 1$ and $\delta > 0$. Suppose that some $L(s,\chi)$ with $\chi \in F$ has a zero with $\operatorname{Re} s > \theta$. Then for every real number C there is Y_0 such that every window $[Y, Y^{1+\delta}]$ with $Y \ge Y_0$ contains a point y with $G_F(y) > C\,y^\theta$.*

Thus a zero shows itself, with a definite sign, in every late window. Three ideas make this work.

(i) *Zeros cannot cancel.* By §2, and because $\operatorname{Re} g_\chi = \frac{1}{2}(g_\chi + g_{\bar\chi})$, the transform of G_F has at a zero ρ the residue $-S(\rho)/(\rho(\rho-1))$, where

$$S(\rho) = \tfrac{1}{2} \sum_{\chi \in F} c_\chi \big(m_\chi(\rho) + m_{\bar\chi}(\rho) \big) > 0.$$

As all weights are positive, the contributions of the different characters add up. At a real zero the residue is positive, which covers possible real zeros close to 1 (Siegel zeros).

(ii) *High derivatives localize.* Put $g(u) = G_F(e^u)$ and let $\hat g(s) = \int_0^\infty g(u)\, e^{-su}\, du$ be its Laplace transform; this is the transform of §2 in the variable $u = \log y$. For an integer $k \ge 0$,

$$A_k(s) := \frac{(-1)^k}{k!}\, \hat g^{\,(k)}(s) = \int_0^\infty g(u)\, \frac{u^k}{k!}\, e^{-su}\, du.$$

Write $\beta = \operatorname{Re} s$. The weight $u^k e^{-\beta u}/k!$ is a bump centred at $u = k/\beta$, of relative width $k^{-1/2}$. Taking $k = \lceil (1+\delta/2)\,\beta \log Y \rceil$ places it on the window, where $\log Y \le u \le (1+\delta)\log Y$.

(iii) *One side suffices.* With $g^+ = \max(g, 0)$ one has $|g| + g = 2g^+$. So an upper bound $g(u) \le Ce^{\theta u}$ on the window gives, for every $\eta > 0$ and a constant C_η,

$$|A_k(s)| + A_k(\beta) \;\le\; C_\eta\, Y^\eta (\beta - \theta)^{-k}.$$

On the other hand a pole of $\hat g$ closer to s than $\beta - \theta$ makes the left side grow faster in k than $(\beta-\theta)^{-k}$: through the first term if the pole is not real, and through the second, with the right sign by (i), if it is. For large k the two statements are incompatible.

The order of choices. Let ψ be a primitive character of order o, that is, $\psi^o = 1$ and no smaller power is trivial. Let $P(\psi)$ be the positive family formed by the primitive characters

inducing the o powers of ψ, with weight 1 on the trivial character and $1/(3o+1)$ on each of the others; write $G_\psi = G_{P(\psi)}$. The reason for these weights is in §5. The arithmetic half of the proof is:

> **Statement** $(\star)$. *There is $\varepsilon_1 > 0$ such that for every $0 < \varepsilon < \varepsilon_1$ there are arbitrarily large Y with*
>
> $$G_\psi(x) \le x^{1/2+9\varepsilon} \quad \textit{for all} \quad x \in [Y, Y^{3/2}].$$

Suppose $L(s,\psi)$ had a zero ρ with $\operatorname{Re}\rho > \frac{1}{2}$. Choose θ with $\frac{1}{2} < \theta < \operatorname{Re}\rho$, and then ε with $\frac{1}{2} + 9\varepsilon \le \theta$. By $(\star)$ there are arbitrarily late windows $[Y, Y^{3/2}]$ on which $G_\psi(x) \le x^\theta$. This contradicts the Proposition for the family $P(\psi)$, with $\delta = \frac{1}{2}$ and $C = 1$.

So θ is chosen after the zero and ε after θ. No bound at $\varepsilon = 0$ and no uniformity in ε is needed, and the windows may depend on everything.

LEAN the Proposition `zeroFree_of_rpow_window` · the order of choices `level_of_tiedWin`, `zeroFreeAll_of_topG`

4 The estimate, I: a slope inequality for a chain of theta lifts

The principle, in its simplest case. Let K be an imaginary quadratic field and ϑ a theta function on an elliptic curve with complex multiplication by K, with period Ω and invariant derivation D. The announcement records (entry 2) that for every prime p inert in K and every integer $a \ge 1$ the number $D^a\vartheta(0)/\Omega^a$ is divisible by

$$p^{a/p + a/p^3 + \cdots - O(\log a)},$$

much as $a!$ is divisible by $p^{\lfloor a/p\rfloor + \lfloor a/p^2\rfloor + \cdots}$. Now a non-zero algebraic integer that is divisible by many primes cannot be small: by the product formula, the sum of $\log p$ times these exponents, over the inert primes, is at most the logarithm of its size at the infinite places. The left side of this inequality is a sum over primes weighted by a ramp, as in §2; the right side can be computed. So one obtains an *upper* bound for a prime sum. The rest of this section makes the error terms small enough: the single derivative becomes a chain of m theta lifts, and the product formula becomes a slope inequality.

The setting. Fix a primitive character ψ of order o, a small $\varepsilon > 0$ and a *scale* x, a large real number: the aim is $(\star)$ at the point x. Choose two auxiliary primes q and d, subject to conditions given in §5, and put $K = \mathbb{Q}(\sqrt{-d})$. Let χ_{-d} be the quadratic Dirichlet character of K: $\chi_{-d}(p) = 1$ if the prime p splits in K, -1 if it is inert, and 0 if $p = d$. Define χ_{-q} likewise for $\mathbb{Q}(\sqrt{-q})$.

The host field. Let e be an integer with $x^{2+8\varepsilon} \le 2^e \le 2x^{2+8\varepsilon}$. The *host* is the abelian number field

$$E = \mathbb{Q}(\zeta_{2^{e+2}})^+ \cdot K_\psi \cdot \mathbb{Q}(\sqrt{-q}, \sqrt{-d}).$$

Here ζ_n is a primitive n-th root of unity, the superscript $+$ denotes the maximal real subfield, and K_ψ is the field cut out by ψ, that is, the fixed field of the kernel of ψ; it is cyclic of degree o over $\mathbb{Q}$. The field E is a CM field: a totally imaginary quadratic extension of a totally real field E^+, with complex conjugation c. As x grows, the first factor climbs the tower $\mathbb{Q} \subset \mathbb{Q}(\zeta_8)^+ \subset \mathbb{Q}(\zeta_{16})^+ \subset \cdots$. This cyclotomic tower plays the part of the finite extensions of $\mathbb{F}$ in Stepanov's proof.

The chain. Let $X_0, X_1, \dots, X_m$ be Hermitian spaces over E, for the conjugation c, that are positive definite at every real place of E^+ ("totally definite" in the announcement), with

$$\dim X_j = n_0 + j(t+1), \qquad n_0 \asymp x^{9/2+10\varepsilon} \text{ even}, \qquad t \text{ odd}, \;\; n_0 < t \le x^{400}.$$

Let $U(X_j)$ be the unitary group of X_j. Because X_j is definite, $U(X_j)$ is compact at the infinite places, and an automorphic form on it, of fixed level and weight, is simply a function f on a finite set: the classes M of Hermitian lattices in a genus, each counted with the weight $1/|\mathrm{Aut}\, M|$.

The *theta correspondence* for the pair $U(X_j)$, $U(X_{j+1})$, together with a Hecke character ξ_j of E, provides a linear map from forms on $U(X_j)$ to forms on $U(X_{j+1})$, the *theta lift*

$$(\Theta_j f)(M') = \sum_M \frac{\vartheta_j(M, M')}{|\mathrm{Aut}\, M|}\, f(M),$$

where M and M' run over the classes of lattices in X_j and X_{j+1}, and ϑ_j is a theta kernel twisted by ξ_j. The *chain* is the composite $\Theta = \Theta_{m-1} \circ \cdots \circ \Theta_0$. In terms of automorphic representations each lift adjoins to a representation π a block attached to ξ_j; the announcement writes $\pi \mapsto \pi \boxplus \mu_j[t]$ (entry 8, with μ_j for our ξ_j), where $\boxplus$ is the isobaric sum. The number m of lifts is the parameter to be optimized.

Three computations enter the slope inequality.

(a) *Finite places: divisibility.* The forms with integral values make up a lattice in each space of forms. The theta kernels are integral and the factors $1/|\mathrm{Aut}\, M|$ have bounded denominators, so the chain maps integral forms to integral forms after multiplication by an explicit integer. The point is that the image is far more divisible than that, at the primes inert in K, as in the simplest case above. The local source of this divisibility is a statement on minors: if $\varpi^2 = p$ in a commutative ring, then ϖ divides a certain 2×2 minor formed from two rows of the lifted data.
Each prime being weighted by its logarithm, the divisibility gained at *height* $a > 0$, the analogue of the order a of the derivative, is

$$\mathrm{ASK}(a) = \sum_{n \ge 1} \Lambda(n)\,\big(1 - \chi_{-d}(n)\big)\Big(\frac{a}{n} - w_n\Big)_+ .$$

The factor $1 - \chi_{-d}(n)$ equals 2 at the odd powers of primes inert in K, 1 at the powers of d, and 0 at all other prime powers. The *weights* are the values of a positive function w on $\Phi := \mathrm{Gal}(E/K)$, and w_n is its value at an element determined by the Frobenius class of n (for $n = p^i$, the i-th power of the Frobenius element of p in $\mathrm{Gal}(E/\mathbb{Q})$). The weights that matter are close to a common value $b := \exp(x^\varepsilon)$, and the heights that occur satisfy $a/b \approx x$. Since $(a/n - w)_+ = w(y/n - 1)_+$ for $y = a/w$, each term is w times the ramp of §2 at a point y close to x.

(b) *The product formula: slopes.* A *Euclidean lattice* $\mathscr{L}$ is a free abelian group of finite rank with a positive definite quadratic form. Its *slope* is $\mu(\mathscr{L}) = -\log \mathrm{covol}(\mathscr{L})/\,\mathrm{rk}(\mathscr{L})$, and $\mu_{\max}(\mathscr{L})$ is the largest slope of a non-zero sublattice. For an injective homomorphism $\varphi\colon \mathscr{L} \to \mathscr{L}'$ of Euclidean lattices, with operator norm $\|\varphi\|$, Bost's slope inequality reads

$$\mu(\mathscr{L}) \le \mu_{\max}(\mathscr{L}') + \log\|\varphi\|.$$

It is applied to the lattices of integral forms and to the map between them given by the chain. What the map gains at the finite places, through divisibility, it must pay at the infinite ones, through its norm.

(c) *Infinite places: periods.* The size of the chain at the infinite places is computed exactly. The announcement names the tools (entries 9, 12 and 18): a seesaw identity with the Rallis inner product formula for the norms of the lifts, Siegel's mass formula, Weyl's formula, and the Chowla–Selberg formula for the periods at the points with complex multiplication. For an elliptic curve with complex multiplication by K and its suitably normalized period Ω, the last one reads

$$\log \Omega = \gamma + \log \frac{2\pi}{\sqrt{d}} - \frac{L'}{L}(1, \chi_{-d}).$$

Together with the growth $a \log a - a$ of derivatives of order a (Stirling's formula), this produces the main term

$$\mathrm{MAIN}(a) = a\Big(\log a - 1 - \gamma + \frac{L'}{L}(1, \chi_{-d}) - \overline{\log w}\Big),$$

where $\overline{\log w}$ is the mean value of $\log w$ over Φ.

The outcome. Applied to the chain, divided by the common weight b and averaged over the m lifts, the slope inequality becomes

$$\frac{1}{b}\Big(\mathrm{ASK}(a) - \mathrm{MAIN}(a)\Big) \;\le\; x^{O(\varepsilon)}\Big(\frac{x}{m} + m + \sqrt{x}\Big).$$

On the right, the errors of the whole chain are shared among its m lifts, which gives x/m; each lift has a cost that grows with the length of the chain, which gives m; the third term does not depend on m. A short chain gains too little and a long one costs too much: the optimum is $m \asymp \sqrt{x}$. In the archive $\sqrt{x\delta_1} \le m \le 2\sqrt{x\delta_1}$ with $\delta_1 = x^{-\varepsilon/8}$.

On the left, suppose for a moment that all weights are equal to one number w, and put $y = a/w$. The definitions of §2 then give at once

$$\mathrm{ASK}(a) - \mathrm{MAIN}(a) = w\big(G(y) - g_{\chi_{-d}}(y)\big).$$

The centring of G and of the g_χ, Euler's constant included, is the archimedean side of a slope inequality. So the outcome is an upper bound of size about $\sqrt{x}$, at $y \approx x$, for a combination of the centred sums of §2. It is not yet $(\star)$: the combination involves χ_{-d} instead of ψ, and with a minus sign. Unequal weights repair this, in §5.

In the archive the outcome of this section is a single proposition, `TopG`: whenever hosts exist on arbitrarily late windows and satisfy the cohomological bound of §6, they carry chains for which the inequality holds, with the exponents hidden in $x^{O(\varepsilon)}$ fixed before ε is chosen.

LEAN hosts `IsChiField` · integrality `ChainInt` · local divisibility `key_minor_dvd` · slopes `slopeInequality_of` · periods `chowla_selberg_period` · the two sides `ASK, MAIN` · sizes `RankVal` · the outcome `TopG`

5 The estimate, II: from the slope inequality to $(\star)$

The characters of the host. The extension $E/\mathbb{Q}$ is abelian, and

$$\mathrm{Gal}(E/\mathbb{Q}) \cong (\mathbb{Z}/2)^2 \times A, \qquad A = \mathrm{Gal}\big(\mathbb{Q}(\zeta_{2^{e+2}})^+ K_\psi/\mathbb{Q}\big),$$

the two factors $\mathbb{Z}/2$ being the Galois groups of $\mathbb{Q}(\sqrt{-q})$ and of K. The characters of $\mathrm{Gal}(E/\mathbb{Q})$ are Dirichlet characters: the products $\chi_{-q}^{\,i}\,\chi_{-d}^{\,j}\,\omega$ with $i, j \in \{0, 1\}$ and ω a character of A. Let $A_0 \subset A$ be the subgroup that fixes K_ψ. The characters of A that are trivial on A_0 are exactly the o powers of ψ.

What unequal weights do. For general weights the difference ASK − MAIN is a finite Fourier sum of centred sums. Write w_σ for the weight attached to an element σ of $\mathrm{Gal}(E/\mathbb{Q})$ outside Φ. Then

$$\mathrm{ASK}(a) - \mathrm{MAIN}(a) = \frac{1}{|\Phi|} \sum_{\chi} \sum_{\sigma \notin \Phi} \overline{\chi(\sigma)}\, w_\sigma\, g_\chi\Big(\frac{a}{w_\sigma}\Big),$$

where χ runs over all characters of $\mathrm{Gal}(E/\mathbb{Q})$, with $g_1 = G$. For equal weights only $\chi = 1$ and $\chi = \chi_{-d}$ survive, and this is the formula of §4. In general the weights decide with which coefficient each character occurs.

The construction of §4 comes as a family. Its members are indexed by the elements τ of a finite set $T \subset A$, called tops in the archive, and by translations by the elements of A_0. Passing from one member to another moves the weights along the group, and so multiplies the coefficient of each character χ by a value of χ. Three steps turn the family of inequalities into $(\star)$.

(i) *Averaging over A_0.* Adding the inequalities over all translations by A_0 annihilates every character that is non-trivial on A_0, by the orthogonality of characters. There remain $4o$ characters: the powers of ψ and their products with χ_{-q}, χ_{-d} and $\chi_{-q}\chi_{-d}$.

(ii) *Positivity by design.* The o powers of ψ are the characters wanted: the primitive characters inducing them form $P(\psi)$. After averaging over $\tau \in T$, a power ψ' of ψ carries the coefficient $|T|^{-1}\sum_{\tau\in T}\psi'(\tau)$. The set T is chosen so that the proportion of its elements lying in A_0 is $4/(3o+1)$, and the proportion in each of the other $o-1$ cosets of A_0 is $3/(3o+1)$, up to a negligible error. The coefficient is then 1 for $\psi' = 1$ and $1/(3o+1)$ for $\psi' \neq 1$. These are the positive weights of $P(\psi)$ in §3.

(iii) *Twisting.* The other $3o$ characters occur with coefficients of no definite sign. They are removed one at a time by means of the bound $|g_\chi(x)| \le x^{1/2+\varepsilon}$. For a given character such a bound is as hard as the theorem itself; but the primes q and d are ours to choose. An averaging argument over the candidate primes (a mean-square bound) shows that for every window some pair (q,d) satisfies the bound for all $3o$ twisted characters, at every point of the window.

The auxiliary primes. For a window beginning at Y put $z = Y^{\varepsilon/5}$. The primes q and d are required to satisfy $\frac{z}{2} < \log q \le z$, $\frac{z}{2}(\log Y)^2 < \log d \le z(\log Y)^2$,

$$q \equiv 3 \pmod 8, \qquad d \equiv 7 \pmod{16}, \qquad q \mid d+1,$$

and every odd prime $p \le z$ must divide $d+1$. These conditions make 2, q and every odd prime up to z split in K, so none of them occurs in ASK. As the pair (q,d) depends on the window, so does the host E: hosts are rebuilt at every scale.

The thick set. The family involves one more finite set, a subset T_2 of A_0 that the archive calls thick. It must contain at least the proportion $x^{-\varepsilon}$ of A and be free of short linear relations. Such sets come from Behrend's construction, which uses the integer points on a sphere: for integers $Q \ge 1$ and $H \ge 2$ there are at least $Q\exp(-\kappa\sqrt{\log Q \log H})/H$ non-negative integers below $Q/(4H)$, with an absolute constant κ, no three distinct ones s_1, s_2, s_3 of which satisfy

$$j(s_2 - s_1) = i(s_3 - s_1) \qquad \text{with } i \neq j \text{ and } 0 < |i|, |j| \le H.$$

For $(i,j) = (1,2)$ this excludes three-term arithmetic progressions. It is applied with $Q = |A_0|$ and H of the size of $\log x$. As a consequence, the orbit of a point under a Frobenius element meets a translate of T_2 at most twice in any H consecutive steps.

Bookkeeping. In units of ε, the exponent lost is $\frac{13}{8}$ in the sizes of m and of the family, 5 in the error terms of the slope inequality, 1 in the density of the thick set, $\frac{3}{8}$ in constants, and 1 in removing the twisted characters. The sum is $\frac{13}{8}+5+1+\frac{3}{8}+1=9$, the 9 of $(\star)$.

LEAN Fourier identity `ch_balance` · weights `Frame.Shares` · twisted characters `BoundedCharsPointwise` · auxiliary primes `QH.pairEx` · hosts at every scale `supplyChi` · the thick set `hostThickExists` · conclusion `core_gm`

6 The price of divisibility: Iwasawa theory

The divisibility of §4(a) holds only up to a loss at each lift, and the loss is governed by Galois cohomology. For a prime ℓ and an integer h, let $\mathbb{Q}_\ell/\mathbb{Z}_\ell(-h)$ be the group $\mathbb{Q}_\ell/\mathbb{Z}_\ell$ on which Galois groups act through the $(-h)$-th power of the ℓ-adic cyclotomic character, and let $G_{E,\ell}$ be the Galois group of the maximal extension of E unramified outside ℓ. Put

$$Ш^1_\ell(E,h)=\ker\Big(H^1\big(G_{E,\ell},\mathbb{Q}_\ell/\mathbb{Z}_\ell(-h)\big)\longrightarrow\prod_{v\mid\ell}H^1\big(E_v,\mathbb{Q}_\ell/\mathbb{Z}_\ell(-h)\big)\Big),$$

the group of classes that are trivial at every place v of E above ℓ, where E_v is the completion of E at v. Write len for the length of a $\mathbb{Z}_\ell$-module. By Poitou–Tate duality $Ш^1_\ell(E,h)$ has the same length as the analogous group $Ш^2$ with coefficients $\mathbb{Z}_\ell(h+1)$.

Vandiver's conjecture, that ℓ does not divide the class number of $\mathbb{Q}(\zeta_\ell)^+$, is a statement on the vanishing of groups of this kind; the announcement speaks of "a significant weakening of Vandiver". What is used is not a vanishing but a bound for the total length, on a whole window:

Statement (`VWinAt` ψ_{401}). *For every $0<\varepsilon<\frac{1}{52}$ there is $J\ge 1$ such that for all sufficiently large Z, and every pair of primes (q,d) satisfying the conditions of §5 with $Z^{1/J}$ in place of Y, there is Y with $Z^{1/J}\le Y\le Z^{1/2}$ such that for every $x\in[Y,Y^2]$ the host E at scale x satisfies*

$$\sum_{\ell}\log\ell\sum_{1\le h<x^{401}}\operatorname{len}Ш^1_\ell(E,h)\;\le\;\exp\big(x^{\varepsilon}/2\big),$$

the outer sum running over all primes ℓ.

Why such a window exists. Fix ℓ. Along a cyclotomic tower these lengths are the lengths of the cokernels of an operator, twisted by h, on an Iwasawa module; by the theorem of Ferrero and Washington the module is finitely generated over $\mathbb{Z}_\ell$. Let R bound its rank. Then there are at most R numbers $c_1,\dots,c_R$ in $\mathbb{Z}_\ell$, the *centres*, with

$$\operatorname{len}Ш^1_\ell(E,h)\;\le\;\sum_{i=1}^{R}v_\ell(c_i-h)\;+\;(\text{terms depending only on the layer}),$$

where v_ℓ is the ℓ-adic valuation. So a twist h is expensive only if it is ℓ-adically very close to a centre. For a bound $\ell^k\le Z^{401}<\ell^{k+1}$, call an integer $1\le h\le Z^{401}$ *exceptional* if $v_\ell(c_i-h)>k$ for some i. Two exceptional integers at the same centre differ by a multiple of ℓ^{k+1}, which exceeds Z^{401}, so they are equal. Each centre therefore accounts for at most one exceptional twist: there are at most R of them, whatever Z.

An exceptional twist h does no harm at a scale x where it is out of range, $h>x^{401}$, or small, $h\le x^{\nu}$ for a fixed small $\nu>0$. In the variable $\log x$ it therefore forbids one interval $[\lambda,B\lambda)$, with $\lambda=\frac{1}{401}\log h$ and $B=401/\nu$.

Interval lemma. *Let $B \geq 1$ and $V > 0$. For any R intervals $[\lambda_i, B\lambda_i)$ there is an interval $[V', 2V']$ disjoint from all of them, with* $\dfrac{V}{2^{R+2}B^R} \leq V'$ *and* $2V' \leq V$.

With $V = \log Z$ and $Y = \exp(V')$, the interval $[Y, Y^2]$ contains no forbidden scale and $Z^{1/J} \leq Y \leq Z^{1/2}$ for $J = 4\,(2B)^R$. This is the window.

LEAN the window `VWinAt` · interval lemma `window1`

7 The automorphic foundation

Both branches of the diagram start from Galois representations attached to automorphic forms: the cohomological bound of §6 is deduced from them, by base change, and so are inputs of §4. The archive constructs them in the following generality. Let E be any CM field, $n \geq 1$, $\mathbb{A}_E$ the ring of adèles of E, and π a cuspidal automorphic representation of $\mathrm{GL}_n(\mathbb{A}_E)$ that is

- *regular algebraic*: at each infinite place its infinitesimal character is that of an irreducible algebraic representation of GL_n, so that its exponents are distinct and lie in $\frac{n-1}{2} + \mathbb{Z}$;
- *polarized*: it is conjugate self-dual up to a twist, $\pi^c \cong \pi^\vee \otimes (\chi \circ N_{E/E^+})$, for a Hecke character χ of E^+ of the appropriate parity, where N_{E/E^+} is the norm.

Fix a prime ℓ and an isomorphism $\iota : \overline{\mathbb{Q}}_\ell \cong \mathbb{C}$. Then there is a continuous semisimple representation

$$r_\pi : \ \mathrm{Gal}(\overline{E}/E) \longrightarrow \mathrm{GL}_n(\overline{\mathbb{Q}}_\ell)$$

which, at all but finitely many places v of E, is unramified and satisfies

$$\iota \ \mathrm{charpoly}\ r_\pi(\mathrm{Frob}_v^{-1}) = \prod_{\alpha} \big(X - q_v^{(n-1)/2}\alpha\big).$$

Here Frob_v is a Frobenius element at v, q_v is the number of elements of the residue field, and α runs over the Satake parameters of the local component π_v, the n complex numbers that determine an unramified π_v. In addition, r_π corresponds to π_v under the local Langlands correspondence at the places $v \nmid \ell$, in the cases specified there; and at the places $v \mid \ell$ it is de Rham, and crystalline where π_v is unramified, with Hodge–Tate weights given by the infinitesimal character of π (these are the conditions of p-adic Hodge theory satisfied by representations that come from geometry). These are the properties known from the work of Harris–Taylor, Shin, Chenevier–Harris, Caraiani and others; the archive contains its own formal proof of the statement it uses.

The construction starts from two inputs, a geometric one resting on a holomorphic Lefschetz fixed-point formula, and a theorem on theta lifts. It proceeds by congruences and limits of traces (pseudocharacters). It first treats CM fields containing an imaginary quadratic field, under further conditions, and then removes these: π is twisted to adjust the polarization and transferred by base change to infinitely many quadratic CM extensions $E(\sqrt{-p})$, and the representations obtained there are patched into r_π by comparing traces.

The final assembly of the theorem has seven groups of hypotheses, the “seven inputs”: facts on the unitary group at an auxiliary place, local theta correspondences, the agreement

of theta integrals, a local matching statement, a covolume bound, the existence of local Langlands parameters, and the cohomological window of §6.

LEAN the Galois representations `Row0b`, `galoisLocalGlobalStrong_recParameter` · the final assembly `every_character_of_literature_of_vWin`

Why each idea calls for the next

2 3 A proof by upper bounds needs a statistic whose transform sees every zero with one sign: hence ramps, centring and positive families.

4 Geometry yields upper bounds through slopes, and slopes see primes through divisibility. Divisibility in bulk needs multiplicity, which iterated theta lifts supply at a price that fixes their number at $\sqrt{x}$, and whose archimedean side reproduces the centring.

5 Theta lifts live on unitary groups of CM fields, so the bound arrives for a whole Galois group of characters at once: hence averaging, positivity by design, and twisting.

6 7 Divisibility has a cohomological defect, tamed by Iwasawa theory in the very tower that carries the averaging; both rest on the Galois representations of the forms involved.

3 Such a construction exists only at special scales, so the analytic theorem must be local in y.

What is machine-checked records of 20 November 2026, 23:51 UTC

Lean is a proof assistant: a proof is a formal object, and a small program, the kernel, checks it. Mathlib is Lean's library of mathematics; `RiemannHypothesis` is its statement of RH. NanoDA is an independent kernel, written in Rust.

✓✓ $(\star)$ for $\psi = 1 \Rightarrow$ `RiemannHypothesis`: Lean kernel and NanoDA (61,697 declarations).

✓ `TopG` (§4) and `VWinAt` ψ `401` for every ψ (§6) $\Rightarrow$ Theorem: Lean kernel (1,421 modules).

✓✓ The five theorems themselves, in the Lean archive of L. Alpöge (with Claude): two independent kernels accepted the complete proof. In runs of the reading project the Rust kernel NanoDA accepted all 1,484,733 declarations of the exported proof, and the five statements matched an independently written Mathlib challenge. S. DeDeo and A. Sutherland each report a Lean kernel replay of all 175,187 modules. No axioms beyond `propext`, `Classical.choice`, `Quot.sound`, the three standard axioms of Lean.

◦ §§4–7 are read here from the Lean source and from the announcement, not re-derived.

✓ Checked in this repository: the logical skeleton of the guide (`lean/Atlas/Architecture.lean`, no axioms) and §1 for ζ against Mathlib at the archive's pin (`lean-mathlib/AtlasZeta.lean`).

Notation

The symbols of the text, in the order of their first appearance. Bound variables of sums and products (n, p, i, j) are omitted; a letter in brackets is the one used in the archive or in the announcement.

s, y complex variable, $\operatorname{Re} s$ its real part; real variable $y \geq 1$

$O(\cdot)$, $\asymp$, $\approx$ at most a constant times; of the same order; close to

$\mathbb{F}$, X a finite field; in the announcement's sentence on page 1, the length of the prime sum

χ, $\bar{\chi}$, χ^* a Dirichlet character, its complex conjugate, the primitive character inducing it; N its modulus

$L(s,\chi)$, $\zeta(s)$ Dirichlet L-function; the case $N = 1$

$\Lambda(s,\chi)$, $W(\chi)$ completed L-function, root number (§1)

ψ, o the primitive character under study; its order

$\Lambda(n)$, γ von Mangoldt function; Euler's constant (§2)

$(u)_+$ $\max(u, 0)$

$G(y)$, $g_\chi(y)$ centred ramp sums of ζ and of χ (§2)

ρ, $m_\chi(\rho)$ a zero; its multiplicity as a zero of $L(s,\chi)$

F, c_χ, G_F positive family, its weights, its sum (§3) [weights a in the archive]

θ, δ, Y, C exponent under test; width and left end of a window $[Y, Y^{1+\delta}]$; a constant

$S(\rho)$ weighted multiplicity of ρ in a family (§3)

$g(u)$, $\hat{g}$, A_k, β $G_F(e^u)$; its Laplace transform; its normalized k-th derivative; $\operatorname{Re} s$ (§3)

η, C_η an arbitrarily small exponent and its constant (§3)

$P(\psi)$, G_ψ the family of the powers of ψ; its sum (§3)

ε, ε_1 small exponent, chosen after θ; its upper limit

x the scale: the point at which $(\star)$ is proved

$(\star)$ the arithmetic statement (§3)

ϑ, D, Ω theta function, invariant derivation, period of a CM elliptic curve (§4)

a height: order of derivative, then the parameter of ASK and MAIN

q, d, z auxiliary primes (§4, §5); $z = Y^{\varepsilon/5}$, the size of $\log q$

K, Φ, χ_{-d}, χ_{-q} $\mathbb{Q}(\sqrt{-d})$; $\mathrm{Gal}(E/K)$; the quadratic characters of K and of $\mathbb{Q}(\sqrt{-q})$

e integer fixing the layer: $2^e \asymp x^{2+8\varepsilon}$ (in e^u it is the usual constant)

E, E^+, c host field, its maximal real subfield, complex conjugation (§4)

ζ_n, K_ψ primitive n-th root of unity; field cut out by ψ

X_j, $U(X_j)$ Hermitian spaces over E and their unitary groups (§4)

n_0, t, m, δ_1 $\dim X_0$ [k in the archive]; odd integer with $\dim X_{j+1} - \dim X_j = t + 1$; number of lifts; $x^{-\varepsilon/8}$

M, f class of Hermitian lattices; automorphic form

ξ_j, ϑ_j, Θ_j, Θ Hecke characters $[\mu_j]$, theta kernels, theta lifts, the chain

π, π^c, $\pi^\vee$, $\boxplus$ automorphic representation, its conjugate, its contragredient; isobaric sum

ϖ an element with $\varpi^2 = p$ (§4a)

ASK, MAIN the finite and the archimedean side of the slope inequality (names of the archive)

w, w_n, w_σ, b weights; their common size $\exp(x^\varepsilon)$

$\mathscr{L}$, μ, $\mu_{\max}$, $\|\varphi\|$ Euclidean lattice, slope, maximal slope, operator norm (§4b)

A, A_0, ω, σ $\mathrm{Gal}(\mathbb{Q}(\zeta_{2^{e+2}})^+ K_\psi/\mathbb{Q})$; its subgroup fixing K_ψ; a character of A; an element of $\mathrm{Gal}(E/\mathbb{Q})$ (§5)

T, τ, T_2 the set of tops, one of its elements; the thick set (§5)

Q, H, κ size and relation bound in Behrend's construction, its constant (§5)

ℓ, h, v a prime; a Tate twist [e in the archive]; a place of E

$G_{E,\ell}$, E_v, H^1 Galois group unramified outside ℓ; completion; Galois cohomology

$Ш^1_\ell(E,h)$, len locally trivial classes (§6); length over $\mathbb{Z}_\ell$

Z, J large parameter; exponent with $Z^{1/J} \leq Y$ (§6)

c_i, R, v_ℓ centres, a bound for their number, ℓ-adic valuation

ν, B, λ_i, V, V' quantities of the interval lemma (§6)

n, $\mathbb{A}_E$ rank; adèles of E (§7)

r_π, ι Galois representation of π; fixed isomorphism $\overline{\mathbb{Q}}_\ell \cong \mathbb{C}$

Frob_v, q_v, α Frobenius element, size of the residue field, Satake parameters

`TopG` the outcome of §4 as one proposition of the archive

`VWinAt` ψ `401` the cohomological window of §6