Levent Alpöge & Claude's Proof of the Riemann Hypothesis

An overview of the proof and its formalization in Lean

Mike Arpaia

Every nontrivial zero of the Riemann zeta function has real part $\frac{1}{2}$. The proof establishes the same statement for every Dirichlet L-function, and it does so without ever locating a zero or exhibiting cancellation in an oscillating sum. It proves one inequality, in one direction, about sums over primes, and that inequality is a descendant of the oldest argument in the subject: Chebyshev's observation that primes are forced to divide binomial coefficients, which are not too large.

This document tells that story from beginning to end. It opens with how the proof came to light and a plain account of what is new. It then shows why a single one-sided inequality is enough, and follows the construction of the inequality step by step. Each step runs into an obstacle, and each obstacle is removed by a new idea; there are ten. The document states what each idea is and why the argument needs it. It does not reproduce the proofs.

Contents

What Happened **3**

What Is New, in Plain Terms **5**

It proves an upper bound and nothing more 5

It goes back to Chebyshev, with a better object 6

Divisibility is manufactured, prime by prime 6

The square root comes from a balance 6

The helpers are chosen at random, and only briefly needed 7

An abelian problem, solved with non-abelian tools 7

What it adds up to 7

The Theorem **8**

The Argument in Brief **9**

Part I: One Direction Is Enough **10**

The sums 10

Landau's principle 10

Part II: Chebyshev's Blueprint **11**

Part III: Ten Ideas **12**

1. The slope inequality is a Chebyshev inequality 12
2. A p-adic theta kernel forces the divisibility 13
3. Hosts give every character a CM field 14
4. A coset design makes the weights positive 15
5. The archimedean constants match exactly 15
6. Integral structures cross the theta correspondence 16
7. Tied chains produce the exponent one half 17
8. Random prime twists behave on a window 18
9. Landau's theorem holds on windows 18
10. Tate–Shafarevich groups are uniformly bounded 19

Part IV: The Proof Assembled **21**

How the ideas fit together 21
What the argument is applied to 21
The foundations 22

The Character of the Proof **24**

Positivity instead of cancellation 24
Exactness instead of estimation 24
Non-abelian methods for an abelian problem 24
Everything is chosen at each scale 25
Where the proof sits in mathematics 25

Overview of the Lean Source **26**

What has to be trusted 26
Following the argument down 28
How the argument is written 29
Where the seventeen million lines go 30
Statements off the main line 31
A reading order 32

Reference: where each part of the argument lives 33
Reference: the libraries in detail 33

References **36**

What Happened

On 20 November 2026 Levent Alpöge, a mathematician at Anthropic, posted a four-page note [1]. In it Alpöge writes that, working with Claude, Anthropic's AI model, they found a proof of the Riemann Hypothesis; that the method extends to the Generalized Riemann Hypothesis; and that "Claude and I have also formalized the arguments without assumptions." The note does not contain the proof. Its one footnote points to it instead: fourteen archive files, `lean.tar.gz.part01` to `lean.tar.gz.part14` on the same site, which together hold the complete Lean source of the formalization.

Publishing only Lean was a deliberate choice. The argument, the note says, is "incredibly long at the moment." A fully detailed informal write-up can be generated mechanically, and Alpöge reports producing one of more than 250,000 pages, but judged it inappropriate to post. A formal proof does not need such a document in order to be checked: Lean's kernel verifies it, and anyone can rebuild it. Alpöge presents the release as a first step, intends to write the argument up properly, and expects it to be "massively simplified, likely immediately."

The note also sketches how the proof was found. Alpöge describes guiding the search to the right setting, after which Claude carried much of the technical work, especially on automorphic forms on higher-rank groups. The starting point was an analogy with Stepanov's elementary proof [10] of the Riemann Hypothesis for curves over finite fields, which forces an auxiliary function to vanish to high multiplicity. Over the integers multiplicity is hard to find, "corresponding to something like differentiating whole numbers." The proof recovers it by chaining theta lifts, each of which plays the role of one order of vanishing, and for a sum over primes up to X the chain has length $\sqrt{X}$.

A timeline in the note runs from early October to 16 November 2026, when the generalized hypothesis was completed. Its entries match the architecture described in this document: divisibility by $p^{a/p+a/p^3+\cdots}$ at inert primes on a CM curve, the switch to the smoothed sum G_ζ, iterated theta lifts between definite unitary groups, the Chowla–Selberg and Siegel mass formulas for the exact constants, Bost's slope inequality, and a weakened form of Vandiver's conjecture for the Galois cohomology classes that twisting could not remove. Alpöge cautions that parts of this account paraphrase summaries of work done by Claude, and defers to the formalization: "the Lean should suffice to clarify everything."

Checking began the same day. Andrew Sutherland at MIT compiled all 175,187 modules and reported, in a status page updated through 21 November [3], that the five main theorems depend only on Lean's three standard axioms, and that the exported proof is accepted by Lean's kernel replay and by the independent checkers nanoda, con-leche and lean4lean. Sutherland told *Scientific American* that "the chance that anything is fishy here is approximately zero" [5].

On 21 November the Lean and Mathlib communities published a joint statement, endorsed by the Lean FRO and the Mathlib Initiative, reporting the outcome of a day of community verification [4]. It checks the release against four criteria. The statement proved is Mathlib's own `RiemannHypothesis`,

which experts in Lean and number theory have vetted. Dozens of people compiled the library without errors and reproduced the same proof export, and the library depends directly only on Mathlib. A separate file that isolates the statement and imports only Mathlib was compared with the proved theorem by the `comparator` tool and by `con-leche --compare`, and both confirm that the two are exactly the same statement. The exported proof was accepted by Lean's official kernel, by the checkers `con-leche` and `con-ron`, which carry a formal proof relative to set theory that they cannot accept a false proof, and by several independent checkers written in other programming languages. All of these ran on an inert export, out of reach of any code in the proof sources. The statement concludes: "We have full confidence that Alpöge's proof of the Riemann Hypothesis is correct."

This document is a companion to that release. It is written from a reading of the Lean source, not from Alpöge's account, and it is meant to be read before the code: to give a mathematician the shape of the argument, the new ideas it rests on and where each of them lives in the formalization, so that seventeen million lines of Lean can be approached with a map. It is not a proof and does not replace one. Every mathematical claim in it is a reading of the statement of some Lean declaration, and where the two differ, the declaration is the authority. Where it describes the *role* of a construction, it follows the structure and names of the source; the overview of the Lean source notes where the compiled proof uses a construction more cheaply than its name suggests. It can be read at three depths. For the big picture, read the plain account of what is new that follows. For the architecture of the argument, read Parts I to IV. To work through the source itself, turn to the overview of the Lean source and its reading order. Its aim is to make the formalization legible now, while the proper write-up that Alpöge intends is still to come.

What Is New, in Plain Terms

For a century and a half the Riemann Hypothesis has been understood as a statement about randomness. The primes should behave like coin flips, and a sum of N coin flips has size about $\sqrt{N}$. Proving the hypothesis was supposed to mean proving that cancellation: showing that some wildly oscillating sum over primes is as small as chance would make it. Nobody knew how.

The proof does not do this. It never shows that anything cancels. Here is what it does instead.

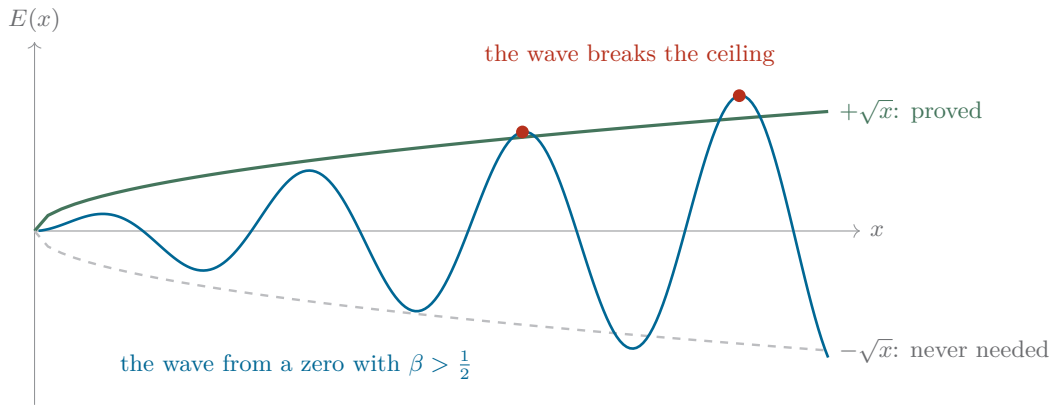
It proves an upper bound and nothing more

Count the primes up to x, giving each prime p the weight $\log p$ and counting prime powers as well. The total, written $\psi(x)$, is very close to x, and the Riemann Hypothesis is equivalent to the statement that the error $E(x) = \psi(x) - x$ stays within about $\sqrt{x}$ of zero, up to a logarithmic factor:

$$-\sqrt{x} \lesssim E(x) \lesssim +\sqrt{x}.$$

That is two inequalities. The error must not be too far above zero, and it must not be too far below. The proof establishes only the first. It shows that the primes are never much *more* plentiful than expected, and it says nothing at all about whether they might be much *less* plentiful.

An old theorem of Landau, from 1905, says that this is enough [7]. The error $E(x)$ is built out of the zeros of ζ. A zero at $\beta + i\gamma$ contributes a wave, roughly $x^\beta \cos(\gamma \log x + \text{phase})$, that swings up and down with an amplitude growing like x^β. A zero off the critical line has $\beta > \frac{1}{2}$, and its wave would rise far above $\sqrt{x}$ again and again. So once the error is known never to rise much above $\sqrt{x}$, no such wave can be present, and no such zero can exist. An oscillation that never goes high cannot be large, so it cannot go low either: the bound from below comes for free. Zeros on the real axis do not oscillate, and a separate check of signs disposes of them.



A zero off the critical line contributes a wave whose amplitude grows like x^β, faster than $\sqrt{x}$. Sooner or later it rises above the ceiling (red), so an upper bound alone rules the zero out. The floor is never checked.

This changes the kind of mathematics required. An upper bound on a count of primes is exactly what divisibility provides: if many primes divide a number, and the number is not too big, there cannot be too many of those primes. A lower bound would require showing that primes exist, which is a different and much harder kind of argument. Cancellation, the thing everyone expected to have to prove, is not needed at all. The first new idea is a decision about where to aim: stop trying to prove that the primes are random, and prove only that they are never too plentiful.

It goes back to Chebyshev, with a better object

There is a classical way to show that a count of primes is not too large, and it is older than Riemann's paper. In 1852 Chebyshev [6] looked at the binomial coefficient $\binom{2n}{n}$. Every prime between n and $2n$ divides it, and yet it is smaller than 4^n. A number cannot be divisible by more primes than its size allows, so there cannot be too many primes between n and $2n$.

Chebyshev's argument has two weaknesses, and they are the reason it was never thought to be a route to the Riemann Hypothesis. It is *lossy*: the size of a binomial coefficient is known only roughly, so the answer is off by a constant factor, where the hypothesis needs an error of square-root size. And it is *blind*: an integer is divisible by a prime or it is not, and it cannot tell a prime that is 1 modulo 4 from one that is 3 modulo 4.

The heart of the proof is a replacement for the binomial coefficient that has neither weakness. It is not a number. It is a map between two lattices that come from the geometry of abelian varieties with complex multiplication.

- **It can see.** Whether a prime divides this map depends on how the prime behaves in a number field. That is exactly the information carried by a Dirichlet character, so the argument reaches every Dirichlet L-function and not only ζ.
- **Its size is known exactly.** The size of the map is a period, one of the numbers that measure the shape of an abelian variety, and for complex multiplication there are exact formulas for periods, going back to Chowla and Selberg. Where Chebyshev had an estimate, the proof has an identity.

Exactness is what closes the gap between Chebyshev and Riemann. An argument that loses a constant factor can never detect a square root. An argument in which the main terms agree identically can.

Divisibility is manufactured, prime by prime

Chebyshev got divisibility for free: a prime between n and $2n$ appears in the numerator of $\binom{2n}{n}$ and not in the denominator. The proof has to build it. At each prime p it constructs, by p-adic analysis, an operator that is provably divisible by a specific power of p. The classical formula for the power of p dividing $a!$ adds up $a/p + a/p^2 + a/p^3 + \cdots$. The new construction produces the same sum with only the *odd* powers $a/p + a/p^3 + a/p^5 + \cdots$, and that pattern is precisely the signature of a quadratic character. This is the step at which characters enter the count of primes, and it is new.

The square root comes from a balance

The exponent $\frac{1}{2}$ in the Riemann Hypothesis is usually explained by randomness. In the proof it arises differently. A single lattice map gives a useless inequality, with an error as large as the answer. Stacking m of them spreads one fixed cost thinly, so that it shrinks like x/m, while a second cost is paid at every level and grows like m. One falls and one rises, and they meet at $m = \sqrt{x}$. The square root is the solution of an optimization problem. No zero of any L-function is ever located.

The helpers are chosen at random, and only briefly needed

The construction cannot be carried out for a character on its own. It needs two helper primes to build the number field in which everything lives, and the helpers bring their own L-functions, which must behave as though the Riemann Hypothesis were already true. That looks circular. It is not, for two reasons. The helpers are needed only on one stretch of the number line at a time, and the proof is free to choose them. Averaged over all choices, the relevant sums behave like sums of random signs, so most choices are good; this is a statement about averages and requires no hypothesis.

Because the helpers work only on stretches, the inequality is obtained only on stretches, and so Landau's theorem itself had to be strengthened to accept an inequality that holds on a sequence of long windows and says nothing in between.

An abelian problem, solved with non-abelian tools

Dirichlet L-functions are the simplest L-functions there are. The proof nevertheless leaves their world immediately. Divisibility can be computed on one kind of group and sizes on another, and the lattice has to be carried between them by the theta correspondence, one of the central mechanisms of the Langlands programme. The bookkeeping errors this introduces are controlled by Galois representations attached to automorphic forms on larger groups. The simplest case of the Riemann Hypothesis turned out to be reachable only from above.

What it adds up to

	What was expected	**What the proof does**
the goal	show that the error in the prime count is small in both directions	show only that primes are never too plentiful: an upper bound, with no lower bound
the mechanism	randomness of the primes	divisibility against size, as in Chebyshev
the key object	the zeros themselves, perhaps as a spectrum	a map of lattices on CM abelian varieties
where $\frac{1}{2}$ comes from	a random walk of length N has size $\sqrt{N}$	two costs, x/m and m, balance at $m = \sqrt{x}$
the level of precision	estimates	exact formulas for every main term
the setting	GL_1 and class field theory	unitary groups, theta lifts, GL_3

The positivity underneath all of this is the humblest in number theory: a nonzero integer has absolute value at least 1. The slope inequality, which drives the proof, is that fact for lattices. It had long been hoped that the Riemann Hypothesis would follow from positivity in arithmetic geometry, as Weil's theorem for curves over finite fields does. That hope was right. The surprise is how elementary the positivity is; the difficulty lies in building an object precise enough to apply it to. The rest of this document makes each of these statements exact.

The Theorem

The Riemann zeta function begins life as a sum over the positive integers,

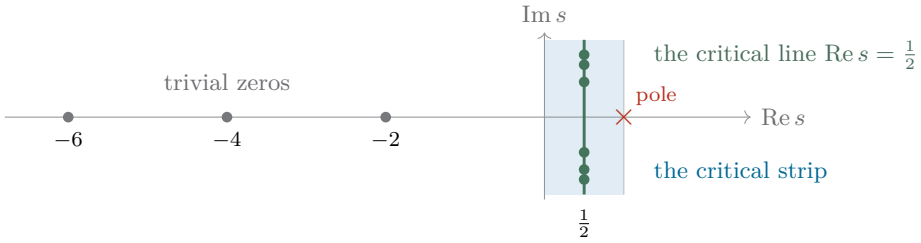
$$\zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \cdots,$$

which converges when $\operatorname{Re} s > 1$. Euler noticed that unique factorization lets the same sum be rewritten as a product with one factor for each prime,

$$\zeta(s) = \prod_{p \text{ prime}} \frac{1}{1 - p^{-s}},$$

and this is why ζ knows about primes. In 1859 Riemann extended ζ to the whole complex plane, apart from a pole at $s = 1$, and showed that its zeros govern the primes: each zero adds an oscillation to the prime count, larger the further right the zero lies.

The negative even integers $-2, -4, -6, \ldots$ are zeros forced by a symmetry of ζ, and are called trivial. All the others lie in the *critical strip* $0 < \operatorname{Re} s < 1$, symmetric about its centre line $\operatorname{Re} s = \frac{1}{2}$.



The zeros of ζ. The Riemann Hypothesis says that every zero in the shaded strip lies on the green line.

Riemann conjectured that every zero in the strip lies on that line, and more than ten trillion have been checked by computer. Stated with Mathlib's definition of ζ, as in the formal proof, the conjecture reads:

Theorem 1 (Riemann Hypothesis)**.** If $\zeta(s) = 0$, s is not a negative even integer, and $s \neq 1$, then $\operatorname{Re} s = \frac{1}{2}$.

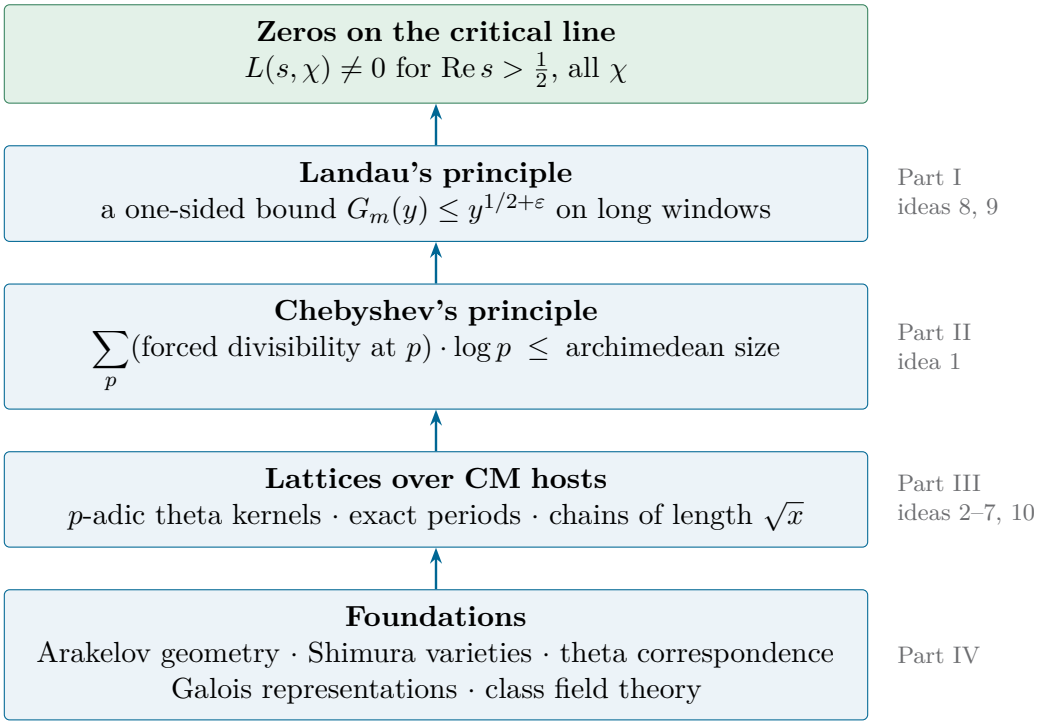
Riemann's function is the first of a family. Dirichlet attached to each *character* χ modulo q, a function on residues that respects multiplication, the series $L(s, \chi) = \sum_n \chi(n) n^{-s}$; ζ is the L-function of the trivial character modulo 1. The proof cannot treat ζ alone, because its central inequality controls ζ only together with the L-functions of a whole group of characters. What it establishes is the generalized Riemann Hypothesis, of which Theorem 1 is the case $N = 1$:

Theorem 2 (Generalized Riemann Hypothesis for Dirichlet characters)**.** Let χ be a Dirichlet character modulo N and let $\operatorname{Re} s > \frac{1}{2}$. If χ is nontrivial or $s \neq 1$, then $L(s, \chi) \neq 0$. Consequently every zero of $L(s, \chi)$ in the strip $0 < \operatorname{Re} s < 1$ lies on the line $\operatorname{Re} s = \frac{1}{2}$.

The Argument in Brief

In outline, the proof is a chain of five reductions, and the parts of this document follow them in order.

1. **One direction is enough.** If a smoothed sum over primes, with its main term removed, is bounded *above* by $y^{1/2+\varepsilon}$, then the corresponding L-functions have no zeros to the right of the critical line. No lower bound is needed. This is Landau's principle.
2. **Upper bounds on prime sums come from divisibility.** If primes are forced to divide an object, and the object is not too large, then a sum over those primes is bounded above. This is Chebyshev's principle.
3. **The object is a map between lattices.** Chebyshev's integer is replaced by an integral map of Hermitian lattices over a CM field. Its forced divisibility is computed by p-adic analysis, and its size by exact formulas for periods.
4. **The exponent comes from a balance.** One such map gives an error as large as the main term. Adding $\sqrt{x}$ synchronized maps reduces the error to $x^{1/2+\varepsilon}$.
5. **Everything auxiliary is chosen at each scale.** The CM field, two auxiliary primes and the lattices are chosen afresh for each x, and averaging arguments show that good choices exist.



Arrows point from hypothesis to conclusion. The document reads from the top down: each part explains how the box it describes is obtained from the one below it.

Part I: One Direction Is Enough

The sums

Let Λ be the von Mangoldt function: $\Lambda(n) = \log p$ if n is a power of the prime p, and 0 otherwise. The proof works with a smoothed Chebyshev sum and its remainder,

$$S(x) = \sum_{n \le x} \Lambda(n)\Big(\frac{x}{n} - 1\Big), \qquad G_\zeta(x) = S(x) - x\log x + (1+\gamma)x,$$

where γ is Euler's constant, and for a nontrivial character χ with the twisted remainder

$$G_\chi(x) = \sum_{n \le x} \Lambda(n)\ \mathrm{Re}\,\chi(n)\Big(\frac{x}{n} - 1\Big) + x \cdot \mathrm{Re}\,\frac{L'}{L}(1,\chi).$$

The weight $x/n - 1$ is chosen for its Mellin transform: for $\mathrm{Re}\, s > 1$,

$$\int_1^\infty S(x)\, x^{-s-1}\, dx = -\frac{\zeta'(s)}{\zeta(s)} \cdot \frac{1}{s(s-1)},$$

and the terms $x\log x$ and $(1+\gamma)x$ in G_ζ are exactly what removes the pole at $s = 1$. Each zero ρ of ζ contributes a term of size $x^{\mathrm{Re}\,\rho}/|\rho(\rho-1)|$ to $G_\zeta(x)$, so the Riemann Hypothesis is essentially the assertion $|G_\zeta(x)| \le Cx^{1/2}$.

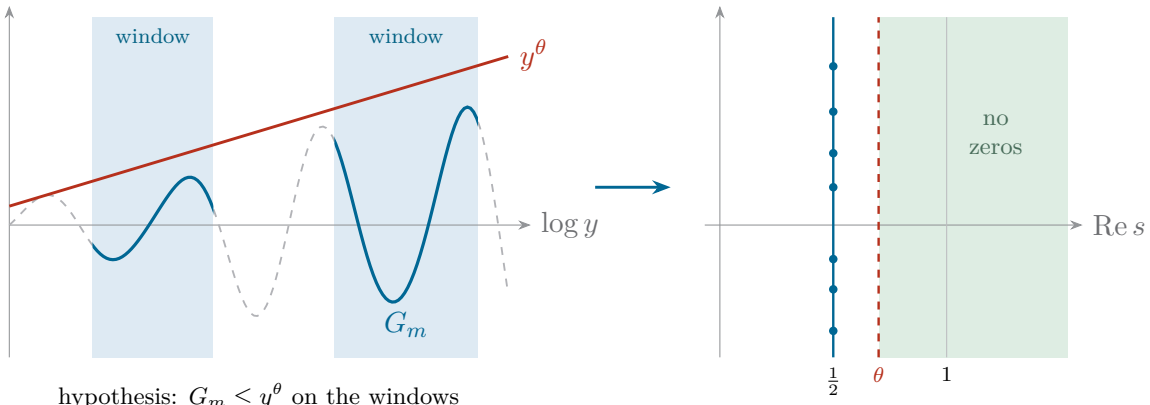
Fix a primitive character ψ, let o be the order of the cyclic group it generates, and let K be the set of primitive characters inducing its powers. The quantity the proof bounds is the positive combination

$$G_m(y) = G_\zeta(y) + \frac{1}{3o+1} \sum_{\chi \in K,\ \chi \ne 1} G_\chi(y).$$

Landau's principle

> **Landau's principle**
>
> Suppose $0 < \theta < 1$ and $G_m(y) \le y^\theta$ for all large y. Then ζ and every $L(s,\chi)$ with $\chi \in K$ are nonzero for $\mathrm{Re}\, s > \theta$.

Here is why. Put $P(s) = \zeta(s)^{3o+1} \prod_{\chi \in K, \chi \ne 1} L(s,\chi)$. Up to the factor $3o+1$, the Mellin transform of G_m is $-P'(s)/\big(P(s)\, s(s-1)\big)$ plus terms that are holomorphic in the region of interest. The hypothesis says that $H(y) = Cy^\theta - G_m(y)$ is nonnegative, and Landau's theorem says that the Mellin transform of a nonnegative function has a singularity at the *real* point of its abscissa of convergence. A real zero $\beta \in (\theta, 1)$ of P is excluded by a sign: it would give the transform of H a residue of the wrong sign for a nonnegative integrand. So the transform converges, and P'/P is holomorphic, throughout $\mathrm{Re}\, s > \theta$. A function with a holomorphic logarithmic derivative has no zeros, and since no factor of P has a pole there, no factor has a zero.



hypothesis: $G_m \le y^\theta$ on the windows

conclusion: zero-free for $\operatorname{Re} s > \theta$

An upper bound on G_m forces a zero-free half-plane. Outside the shaded windows (dashed) nothing is assumed about G_m; the form of the principle that allows this is idea 9.

Two features of the principle shape everything that follows.

- **Only an upper bound is needed.** A method that can only produce inequalities in one direction suffices.
- **A positive combination suffices.** Positivity comes from the hypothesis $H \ge 0$, not from the Dirichlet coefficients, so ζ and its twists may be mixed with any positive weights.

Taking $\theta = \frac{1}{2} + \varepsilon$ for every $\varepsilon > 0$ gives Theorem 2 for $\operatorname{Re} s > \frac{1}{2}$. The functional equation reflects a zero with $\operatorname{Re} s < \frac{1}{2}$ to one with $\operatorname{Re} s > \frac{1}{2}$ for $\bar{\chi}$, which gives the critical line, and passing from primitive to arbitrary characters changes only finitely many Euler factors, none of which vanish for $\operatorname{Re} s > 0$. The entire problem is now to prove

$$G_m(y) \le y^{1/2+\varepsilon}.$$

Part II: Chebyshev's Blueprint

Chebyshev bounded sums over primes without any analysis of ζ. The central binomial coefficient $\binom{2n}{n}$ is an integer; every prime p with $n < p \le 2n$ divides it; and it is at most 4^n. Therefore

$$\sum_{n<p\le 2n} \log p \;\le\; \log \binom{2n}{n} \;\le\; n \log 4.$$

The argument has a *local* half and a *global* half. Locally, at each prime, arithmetic forces a multiplicity. Globally, an archimedean estimate bounds the size. Since the logarithm of a positive integer is the sum of $\log p$ over its prime factors, the two halves combine into a one-sided inequality on a prime sum, which is precisely the kind of statement Landau's principle consumes. The identity $\log n! = \sum_m \Lambda(m)\lfloor n/m \rfloor$ and Stirling's formula $\log n! = n \log n - n + O(\log n)$ are the same blueprint applied to a factorial, and the sum $S(x)$ with its main term $x \log x - x$ is their continuous version.

Chebyshev's inequality has an error of order n. The Riemann Hypothesis needs $n^{1/2+\varepsilon}$, and it needs the inequality for characters, which integers cannot see. The proof keeps the blueprint and replaces every component.

	Chebyshev	**This proof**
the object	the integer $\binom{2n}{n}$	an integral map $\phi\colon N \to R^\vee$ of Hermitian lattices over a CM field
local half	p divides it for $n < p \le 2n$	ϕ is divisible by p^{g_p} at inert p, with g_p given by a p-adic theta kernel
global half	it is at most 4^n	the height of ϕ and the slope of $R^\vee$, given by exact period formulas
what combines them	$\log N = \sum_p v_p(N) \log p$	the slope inequality
which primes are seen	all primes alike	primes by Frobenius class, hence characters
error	order n	order $n^{1/2+\varepsilon}$

Part III: Ten Ideas

The right-hand column of the table is the proof. This part builds it one entry at a time. Each step begins with the difficulty left over from the step before.

1 The slope inequality is a Chebyshev inequality

The obstacle. An integer has a size and a factorization, and nothing else. It cannot distinguish a prime p from another prime in a different residue class, so it can never produce the twisted sums G_χ.

A *Hermitian lattice* over the ring of integers of a number field E is a projective module with a Hermitian metric at each archimedean place. It has an Arakelov *slope* μ, the negative of its normalized log covolume. A linear map ϕ between two such lattices has a *height* $h(\phi)$, the mean over the archimedean places of its log operator norm. These are the analogues, for lattices, of the logarithm of an integer, and the analogue of $\log N = \sum_p v_p(N) \log p$ is the slope inequality.

The slope inequality

Let $\phi\colon N \to R^\vee$ be an integral map of Hermitian lattices over $\mathcal{O}_E$. If ϕ is divisible by p^{g_p} at each prime p, then

$$\sum_p g_p \log p \;\le\; h(\phi) + \mu(R^\vee) + (\text{lower-order terms}).$$

An integral map cannot be both small at the archimedean places and divisible at the finite ones. The two halves of Chebyshev's argument are the finite and infinite places of this single statement, so a one-sided bound on a prime sum is the same thing as a slope inequality, provided a map can be found whose divisibility at p carries the right weight. A lattice over a CM field has enough structure for that: its divisibility at p can depend on the Frobenius class of p, and its slope has an exact

formula where an integer has only a size.

In the formalization. The global half of this inequality is applied to a height that is *defined* to equal the explicit main-term formula, and the two inequalities it must satisfy are proved for that formula from operator-norm, covolume and valuation estimates on matrices, not from the degree of a line bundle.

Lineage. Bost's slope method in Diophantine geometry [11], where the inequality proves algebraicity and transcendence statements, and the Harder–Narasimhan theory of Euclidean lattices of Stuhler and Grayson. The method had not been used to bound sums over primes.

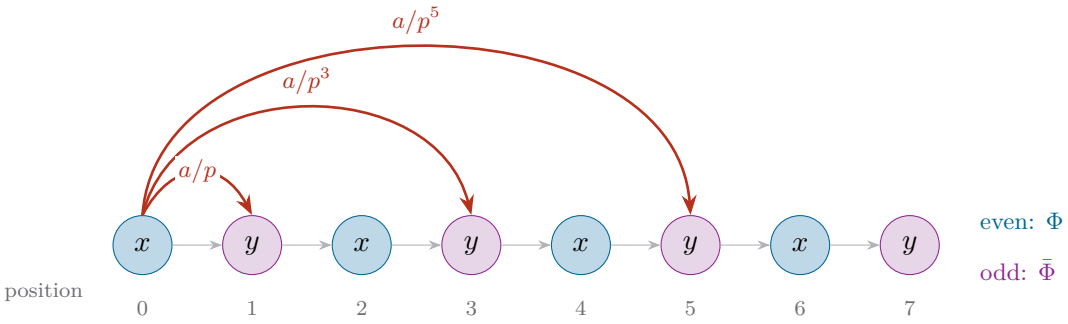
2 A p-adic theta kernel forces the divisibility

The obstacle. The map must be divisible at p by exactly the weight $\Lambda(n)(x/n-1)$ of Part I, and only at the prime powers $n=p^e$ where a chosen character equals -1. The classical source of forced divisibility, Legendre's formula $v_p(a!)=\sum_e \lfloor a/p^e \rfloor$, has the right shape but treats every exponent e alike.

The map is built from algebraic modular forms on a definite unitary group, with values in an integral representation of weight $a \approx x \cdot b$. At a prime p that is inert in the relevant subfield, the local field is unramified of even degree and Frobenius σ moves a form through a cycle of *positions*; the even positions make up one half Φ of the CM type and the odd positions the conjugate half $\bar{\Phi}$. The proof constructs integral lifting operators between unitary groups of increasing rank, each given by a p-adic theta kernel

$$\Theta = \exp(-\Phi), \qquad \Phi = \sum_e \beta(x_e, y_e), \qquad \log_F(x) = \sum_{j\geq 0} \frac{x^{p^{2j}}}{p^j},$$

where β is a σ-sesquilinear pairing built from the formal group with logarithm $\log_F$, of height two. That Θ has integral coefficients is Dwork's lemma.



A Hermitian pairing joins Φ to $\bar{\Phi}$, so a weight a at an even position reaches only odd positions, $2e+1$ steps away, and arrives divided by p^{2e+1}.

Because the variables x sit in even positions and the variables y in odd ones, the kernel couples a weight only across an odd number of Frobenius steps. Comparing the divided-power lattice of the representation with its dual then forces each lifting operator to be divisible by p^g, where

$$g \;=\; \sum_{e \text{ odd}} \max\Big(0,\; \frac{a}{p^e} - c\Big) \;\approx\; b \sum_{e \text{ odd}} \Big(\frac{x}{p^e} - 1\Big)^{+}$$

and $c \approx b$ is a capacity set by the target weight. This is the weight of Part I at $n = p^e$, restricted to odd e. At an inert prime the quadratic character takes the value $(-1)^e$ on p^e, so the odd exponents are exactly the prime powers where it equals -1.

Lineage. Legendre's formula; the integral theory of Weyl modules and divided powers; the formal groups of Lubin–Tate and Honda; Kudla's integral theta kernels. The closed formula for the divisibility, and its term-by-term agreement with a smoothed Chebyshev sum, are new.

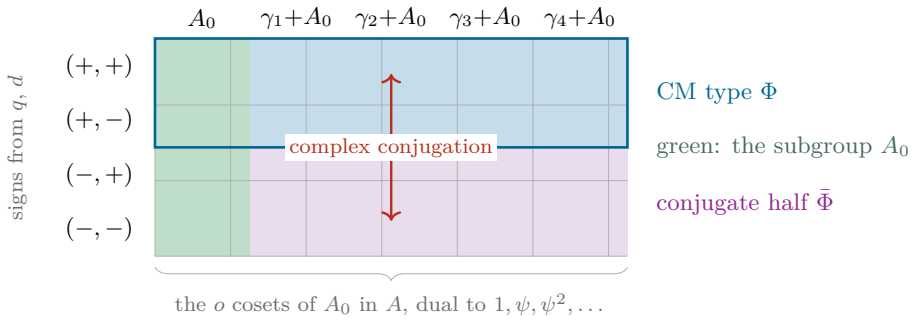
3 Hosts give every character a CM field

The obstacle. Hermitian lattices, CM types and Frobenius positions require a CM field. The target is an arbitrary Dirichlet character ψ, even or odd, with no CM structure of its own.

For each ψ and each scale, the proof constructs a *host*: an explicit abelian CM field E inside a cyclotomic field, with

$$\mathrm{Gal}(E/\mathbb{Q}) \;\cong\; (\mathbb{Z}/2)^2 \times A.$$

The two sign coordinates come from two auxiliary primes $q \equiv 3 \pmod 8$ and $d \equiv 7 \pmod{16}$, and they supply the CM structure: complex conjugation is $(-,-)$, and the elements whose first sign is $+$ form a CM type. The finite abelian group A is reserved for ψ: it contains a cyclic subgroup A_0 of index o, the order of ψ, and the quotient A/A_0 is dual to the group generated by ψ.



The Galois group of a host, drawn for $o = 5$. Rows are the sign classes contributed by the auxiliary primes; columns are the cosets of A_0. The upper half is a CM type; the first column is the subgroup A_0.

The congruence conditions have a purpose. Requiring that q and every small odd prime divide $d+1$ forces all small primes to split in $\mathbb{Q}(\sqrt{-d})$, which keeps them out of the inert set where divisibility is forced and where they would overwhelm the count.

Lineage. The classical use of an auxiliary imaginary quadratic field to study a real character. Here the auxiliary data is chosen afresh at every scale, and its own L-functions have to be controlled separately; that is idea 8.

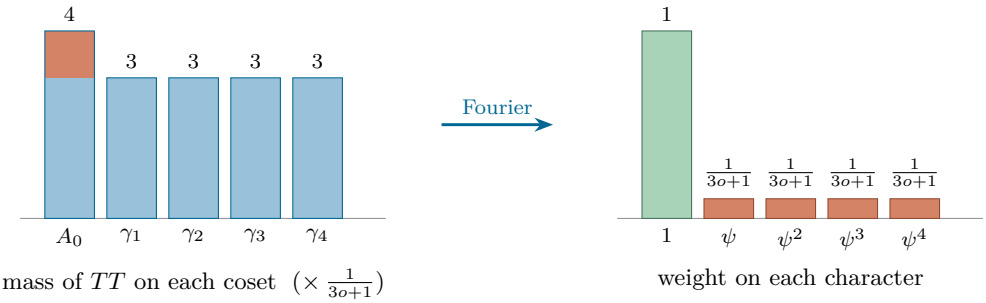
4 A coset design makes the weights positive

The obstacle. The divisibility at p depends on whether Frobenius lies outside the CM type. Expanding that indicator in characters gives a combination of odd characters with coefficients of both signs, but Landau's principle demands a *positive* weight on every L-function whose zeros are to be controlled.

The inequality is not used for one CM type. It is summed over a set $TT \subset A$ of translates of the type, and TT is distributed among the o cosets of A_0 in a prescribed way: mass 3 on every coset, and one extra unit on the trivial coset,

$$\mu(A_0) = \frac{4}{3o+1}, \qquad \mu(\gamma + A_0) = \frac{3}{3o+1} \quad (\gamma \notin A_0), \qquad \text{total } \frac{4+3(o-1)}{3o+1} = 1.$$

The uniform part of μ has Fourier transform supported at the trivial character, and the point mass has constant transform. So the transform of μ on A/A_0 is 1 at the trivial character and exactly $\frac{4-3}{3o+1} = \frac{1}{3o+1}$ at every other one.



The design for $o = 5$. The one extra unit of mass (red, left) becomes an equal positive weight on every power of ψ (red, right).

Fourier inversion on the Galois group therefore turns the summed inequality into a bound on G_ζ with weight 1 plus each G_{ψ^k} with the same positive weight $\frac{1}{3o+1}$. That combination is G_m, and this is where its definition comes from.

Lineage. The averaging over CM types in Colmez's conjecture and its averaged form. There the average removes the dependence on the type; here it is weighted so as to leave a controlled positive residue on a chosen cyclic group of characters.

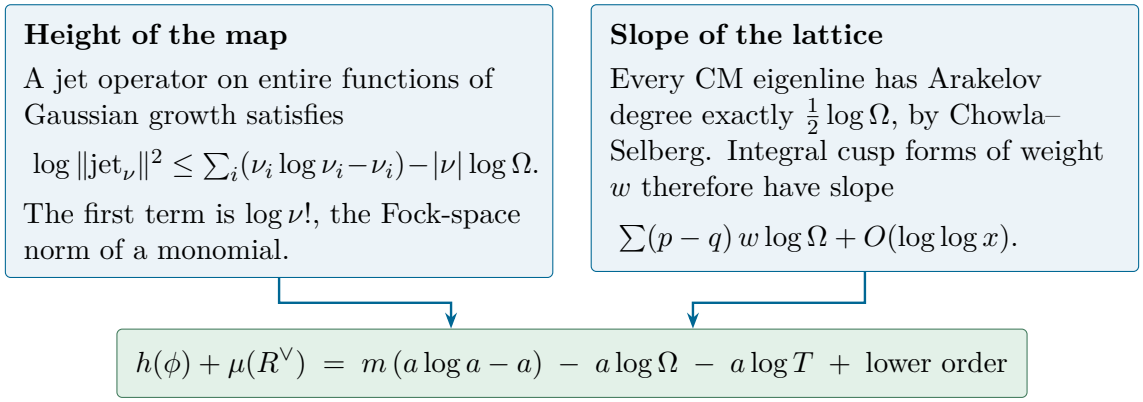
5 The archimedean constants match exactly

The obstacle. The target error is $x^{1/2+\varepsilon}$ against a main term of size $x \log x$. A general positivity theorem bounds the archimedean side only up to a constant factor, and any such loss destroys the result.

The proof never estimates a leading term. The same CM period Ω, with

$$\log \Omega = \gamma + \log(\pi/v) - \frac{L'}{L}(1, \chi_E),$$

enters both the height of the map and the slope of the lattice, and it enters through equalities.



Both sides of the slope inequality are evaluated by formulas, and the Gamma factors and hook products they contain cancel. Stirling's term $a \log a - a$ is the main term $x \log x - x$ of the Chebyshev sum, and the period Ω is the source of the constants $\frac{L'}{L}(1, \chi)$ in G_χ.

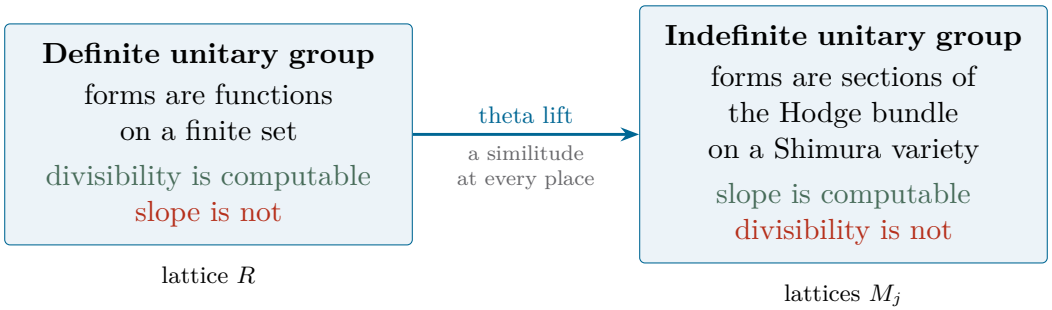
The exponent $\frac{1}{2}$ is available only because the leading constant is an identity. Inequalities are used solely for terms that are already of lower order.

Lineage. The Chowla–Selberg formula [9] and Kronecker's limit formula; Colmez's formula for the Faltings height of a CM abelian variety; Moret-Bailly's formule clé for the theta line; the arithmetic Hilbert–Samuel theorem. The sharp jet-norm bound and the organization of these formulas into one cancellation are new.

6 Integral structures cross the theta correspondence

The obstacle. The two halves live on different groups. Divisibility was computed on a *definite* unitary group, where forms are functions on a finite set. Slopes can be computed only where there is geometry, on the Shimura variety of an *indefinite* unitary group.

The proof carries the lattice from one side to the other. Theta lifts take the lattice R on the definite side into lattices M_j of integral cusp forms on the indefinite side, which are sections of powers of the Hodge bundle with the Petersson metric. The lifts are similitudes at every place, so covolumes multiply up to explicit local indices.



Neither group supports both halves of Chebyshev's argument. The theta correspondence joins them, and the slope of $R^\vee$ is bounded by the covolumes of the M_j.

The outcome is a bound $\mu(R^\vee) \le \sum_j \mathrm{lcv}(M_j)/D + (\text{similitude factors}) + (\text{local indices})$, in which each

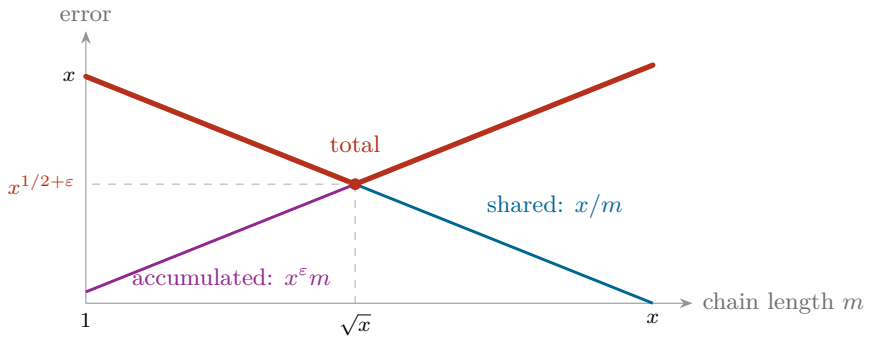
log covolume $\mathrm{lcv}(M_j)$ is controlled by arithmetic intersection theory on a toroidal compactification. An identity between theta integrals for signatures $(a, b+2)$ and (a, b) lets the construction descend through signatures.

Lineage. The Siegel–Weil formula and Howe duality, which compare automorphic forms on the two members of a dual pair as complex vector spaces, and the arithmetic theta lifts of Kudla's programme. Here the comparison is of lattices with metrics, tracked place by place.

7 Tied chains produce the exponent one half

The obstacle. One lattice gives the inequality at one sample point, with a fixed loss at the top of the construction that is as large as the main term.

The inequality is produced on a *chain*: m successive lifts of the lattice data, across all cosets of A_0. A chain is *tied* when its members share the same frame, the same auxiliary primes and the same scale constants. On a tied chain the slope inequalities at the individual nodes add, and two errors pull in opposite directions. The fixed loss at the top is shared among the m nodes and contributes x/m to each. A boundary loss is paid once at every node and contributes $x^{\varepsilon} \cdot m$ in total.



Both axes are logarithmic. The total error (red) is the larger of the two and is smallest at $m \approx \sqrt{x}$.

A bound of quality x^{θ} requires $m \geq x^{1-\theta}$ and $m \leq x^{\theta+\varepsilon}$, and both hold only if $\theta \geq \frac{1}{2}$. The exponent $\frac{1}{2}$ is the balance point of a cost that is amortized against a cost that accumulates. It does not come from the location of any zero.

Lineage. Chebyshev's device of combining several factorials, $\frac{(30n)!\,n!}{(15n)!\,(10n)!\,(6n)!}$, to sharpen his constants. In the proof the number of terms combined grows with x.

8 Random prime twists behave on a window

The obstacle. The host field has more characters than the powers of ψ: it also carries their twists by the quadratic characters of q and d. The slope inequality bounds a combination involving all of them, so the twists must be controlled independently, with the full strength of the Riemann Hypothesis. Assuming that would be circular.

The twists are needed only on a bounded range, and the auxiliary primes can be chosen. For every modulus M, every ε and A, and all large Y, the proof finds primes q, d with $\log q \approx Y^{\varepsilon/5}$ such that every twist χ of a character modulo M by the quadratic characters of q and d satisfies

$$|G_\chi(y)| \le y^{1/2+\varepsilon} \quad \text{for all } y \in [Y, Y^A].$$

The reason is a mean square over prime moduli,

$$\sum_{q\in P}\Big|\sum_{n\le Z}\Lambda(n)\min(1,\tfrac{y}{n})\kappa(n)\Big(\frac{n}{q}\Big)\Big|^2 \;\le\; C\,y\log y\cdot|P|,$$

in which the diagonal terms $n_1n_2 = \square$ give $y\log y$ and the off-diagonal terms are negligible because the symbols (q/s) are equidistributed as q runs over primes in a progression. The decisive point is one of scale: the conductor $q \approx \exp(Y^{\varepsilon/5})$ is enormous compared with the length y of the sum. A sum that is short relative to its conductor behaves, for most q, like a sum of random signs, and for random signs square-root cancellation is a statement about second moments. A zero-density estimate removes the few primes whose twists have a zero near $\operatorname{Re} s = 1$: at most $CQ^{1/4}$ of them, among at least $U^{9/10}$ candidates.

The Riemann Hypothesis for these twists is never used. Only their behaviour on the one window for which q and d were chosen is.

Lineage. Heath-Brown's large sieve for real characters and the zero-density estimates of Montgomery and Jutila. The prime number theorem in progressions is needed beyond the Siegel–Walfisz range and is supplied by Gallagher's theorem with the Deuring–Heilbronn phenomenon. An exceptional real character is not assumed absent: Siegel's theorem bounds its effect, and a single exceptional modulus is excluded from the average.

9 Landau's theorem holds on windows

The obstacle. The good primes of idea 8 exist for each Y separately and give control only on $[Y, Y^A]$. The arithmetic can therefore deliver the bound $G_m(y) \le y^\theta$ only on a sequence of windows, never for all large y, and Landau's principle as stated in Part I asks for all large y.

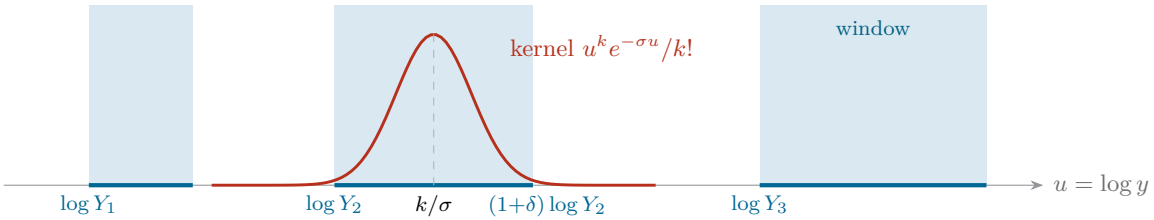
The analytic principle is adapted to what the arithmetic can deliver.

Landau's theorem on windows

Fix $\delta > 0$. If for arbitrarily large Y the bound $G_m(y) \le Cy^\theta$ holds on the window $[Y, Y^{1+\delta}]$, then the L-functions of the family have no zeros with $\operatorname{Re} s > \theta$.

The proof looks at the k-th Taylor coefficient of the Laplace transform of $G_m(e^u)$ at a real point $\sigma > 2$. That coefficient is the integral of $G_m(e^u)$ against the kernel $u^k e^{-\sigma u}/k!$, which is concentrated

near $u = k/\sigma$ with relative width $k^{-1/2}$. For suitable k a Chernoff bound puts all but $C\sigma^{-k}$ of its mass inside one window.



Choosing the order k of the derivative places the kernel inside a window, where the one-sided hypothesis holds. What happens between windows is never seen.

On the window the hypothesis bounds the coefficient by $Ce^{\eta U}(\sigma-\theta)^{-k}$. A zero ρ with $\operatorname{Re}\rho > \theta$ would make it at least $c\,|\sigma-\rho|^{-k}$, and since $|\sigma-\rho| < \sigma-\theta$, geometric growth in k defeats the subexponential slack.

Lineage. Landau's theorem and Pringsheim's theorem on power series with nonnegative coefficients; the localized oscillation theorems of Kaczorowski and Pintz, which use Turán's power sums. Here the localization is by concentration of a high-order derivative kernel.

10 Tate–Shafarevich groups are uniformly bounded

The obstacle. The theta transfer of idea 6 compares lattices across multiplicity spaces, and the local indices between them are congruence modules. Their total, weighted by $\log \ell$ over all primes ℓ, is one more error term. It must be bounded for every ℓ simultaneously and uniformly in the family, which no theorem about a single prime provides.

The indices are bounded by rank times the logarithm of the order of certain Tate–Shafarevich groups, and the proof establishes

$$\sum_{\ell} \log \ell \sum_{1 \le e < x^{401}} \operatorname{length}_{\mathbb{Z}_\ell} \operatorname{Sha}^1\big(G_{E,S},\, \mathbb{Q}_\ell/\mathbb{Z}_\ell(-e)\big) \;\le\; \exp\big(\tfrac{1}{2}x^{\varepsilon}\big)$$

for the host fields E, summed over all primes ℓ including 2. The requirement is modest in size, since the bound need only be smaller than $\sqrt{b}$, but severe in uniformity. The group splits under complex conjugation, and the two parts are bounded by quite different means.

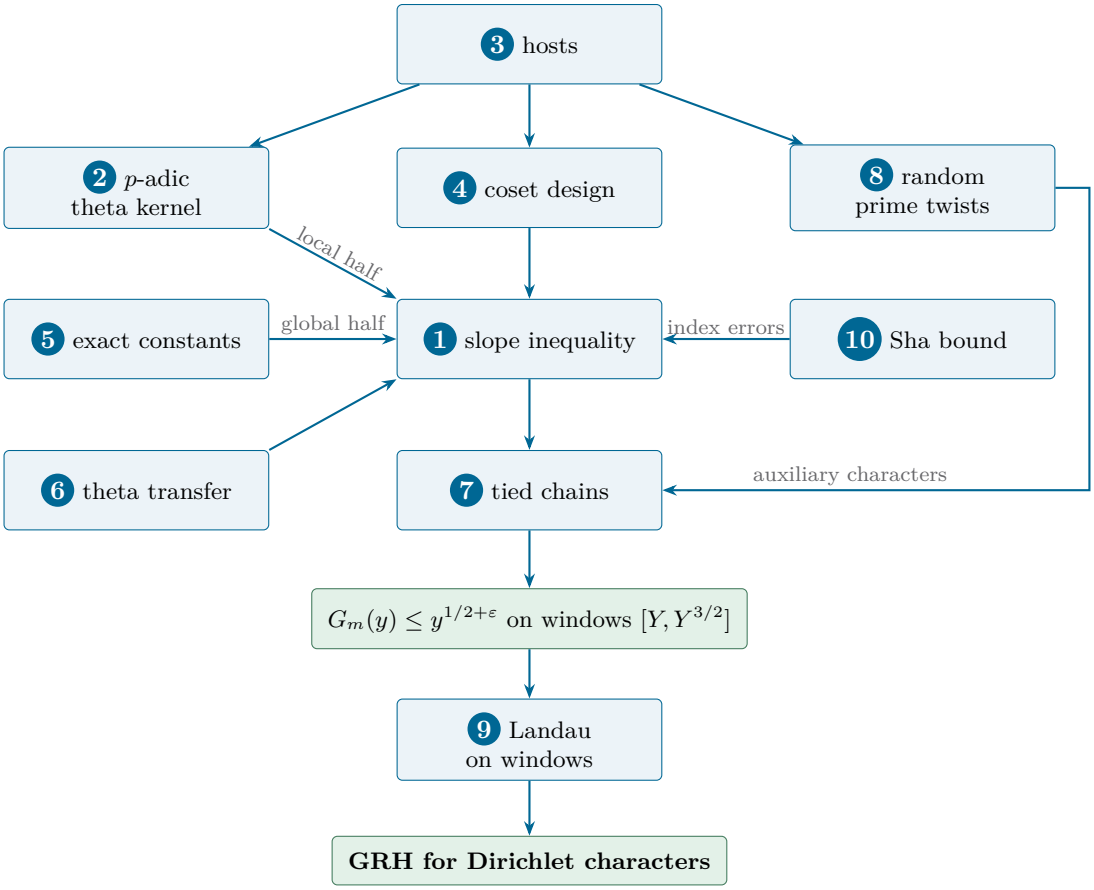
Part	How it is bounded
minus part	Each character eigenspace has length at most the ℓ-adic valuation of a generalized Bernoulli number, a value $L(-e, \chi)$. This is the divisibility in the main conjecture of Iwasawa theory.
plus part	The length is at most the depth of congruences between cuspidal three-dimensional Galois representations ρ over an imaginary quadratic field and the reducible representation $\Psi \oplus 1 \oplus \Psi^{-1}$. One nonzero integer of controlled height, a congruence number for automorphic forms on U(3), absorbs these valuations for every ℓ at once.
$\ell = 2$ and irregular primes	Cyclotomic towers, using the vanishing of the μ-invariant.

This is where Galois representations of automorphic forms and the local Langlands correspondence enter a proof about abelian L-functions: the Tate–Shafarevich groups of an abelian field are controlled through automorphic forms on GL_3.

Lineage. The Mazur–Wiles theorem; Ribet's construction of unramified extensions from Eisenstein congruences and its higher-rank forms due to Bellaïche–Chenevier and Skinner–Urban; Kato's Euler system; the vanishing theorems for adjoint Selmer groups from the Taylor–Wiles method.

Part IV: The Proof Assembled

How the ideas fit together



Dependencies among the ten ideas. Hosts make the other constructions possible; four ideas feed the two halves of the slope inequality; chains, the design and the control of auxiliary characters turn it into the bound on G_m.

Read from the top, the diagram is the proof. Given a character ψ and a scale x, choose auxiliary primes for which the twists behave (8) and build the host (3). On a definite unitary group over the host, the theta kernel forces divisibility at every inert prime (2). Transfer the lattice to a Shimura variety (6), where its slope and the height of the map are given by exact formulas (5), with the index errors bounded by Tate–Shafarevich groups (10). The slope inequality (1) is then an upper bound on a prime sum by Frobenius class. Summing over a tied chain of length $\sqrt{x}$ (7) and over the coset design (4) gives $G_m(y) \le y^{1/2+\varepsilon}$ on a window, and Landau's theorem on windows (9) concludes.

What the argument is applied to

The construction is carried out not for one host but for a *family*: for each scale x in an unbounded set of windows, a collection of hosts with shared ranks, shared Galois data and weights confined to narrow intervals, with scales tied by $L = x^{\varepsilon}$, $b = e^{L}$, $a = x \cdot b$. The family is what makes every error term uniform. The formal proof combines a short list of statements, each holding uniformly over the family.

Link	Ideas	Content
existence	3, 8	large families of hosts exist in every class
local gain	2	prime sums are at most the forced divisibilities at inert places
global size	5, 6	slope plus height is at most the Stirling–period main term plus errors
congruence exponents	10	lattice-index sums are at most rank times (size + $\log \#\text{Sha}$)
irreducibility		a defect term is nonnegative and exponentially small
near-orthogonality		Gram matrices of the relevant vectors are diagonal up to $\exp(-x^{\varepsilon/8})$
rank		the lattice ranks agree across levels

The foundations

Beneath the ten ideas lies a large body of established theory, developed inside the formalization instead of being assumed. By volume it is almost the entire development. By logic it is a supplier: it delivers a handful of uniform inequalities to the links above.

Role in the proof	Theory
slopes of cusp-form lattices	Arakelov geometry: arithmetic Hilbert–Samuel, bigness and nefness of Hermitian line bundles, arithmetic volumes, heights of cycles, Poincaré–Lelong
the ambient spaces	unitary Shimura varieties and Siegel moduli spaces with toroidal compactifications; the Hodge bundle $\det \omega$; semi-abelian degenerations
controlling singularities	resolution and principalization in characteristic zero; Kawamata–Viehweg and Nadel vanishing; L^2 extension
building lattices and test vectors	Weil representation, theta correspondence and Siegel–Weil for unitary groups; Howe duality; the doubling method
the real place	discrete series, Harish-Chandra modules, Casselman–Wallach, Bessel models
the Sha bound	Galois representations of regular algebraic polarized cuspidal representations; local Langlands for GL_n; Poitou–Tate duality; Iwasawa theory of cyclotomic towers
CM theory	class field theory, Chowla–Selberg, Moret-Bailly's key formula, p-adic Hodge theory of abelian varieties

The proof also draws on techniques from areas with no previous bearing on the zeros of L-functions.

Technique	Use in the proof
Harder–Narasimhan filtrations and slope inequalities for Euclidean lattices	the framework in which the global half is stated
Minkowski's second theorem and his bound on finite subgroups of $\mathrm{GL}_n(\mathbb{Q})$	successive minima of the lattices; triviality of finite stabilizers
Strassmann's theorem and p-adic interpolation determinants	counting p-adic zeros in windows of an arithmetic progression
Weyl modules, divided powers and Capelli identities	the two integral structures whose comparison forces divisibility; eigenvalues of radial operators
the Yangian of $\mathfrak{gl}_2$ and its RTT relation	recursions for the radial parts of Bessel models at the real place
Herglotz–Bochner positivity	positivity of orbital integrals on unitary groups
stationary phase with symbol classes	decay of archimedean matrix coefficients
Green's parametrization of the cuspidal characters of $\mathrm{GL}_d(\mathbb{F}_q)$	the irreducibility link

The Character of the Proof

Seen whole, the proof has four traits that set it apart from earlier work on the problem.

Positivity instead of cancellation

Approaches to the Riemann Hypothesis divide broadly into two kinds. *Cancellation* approaches try to show that an oscillating sum, such as $\sum_{n\le x}\mu(n)$ or $\sum_\rho x^\rho/\rho$, is small because its terms interfere. *Positivity* approaches try to exhibit the statement as the nonnegativity of some quantity.

This proof is of the second kind, and it has a clear ancestor. Weil proved the Riemann Hypothesis for curves over finite fields [8] by realizing the zeros as eigenvalues of Frobenius and deducing the bound from the positivity of an intersection pairing on a surface. The long-standing hope has been that an arithmetic variety, with Arakelov's metrics at the infinite place standing in for the missing fibre, could play the same role for ζ itself.

	Weil, curves over $\mathbb{F}_q$	**This proof**
geometric object	the surface $C \times C$	Hermitian lattices of modular forms on unitary Shimura varieties
source of positivity	Hodge index theorem	the slope inequality for integral maps
how primes enter	fixed points of Frobenius	forced divisibility of theta kernels at inert primes
what is bounded	$\lvert\#C(\mathbb{F}_{q^n}) - q^n - 1\rvert$	a positive combination of G_ζ and G_χ
direction of the bound	two-sided	one-sided; Landau's principle supplies the rest

The last row is the decisive structural difference. Geometry over $\mathbb{Z}$ yields inequalities in one direction far more readily than it yields two-sided control. Landau's principle shows that one direction is enough, provided it applies to a positive combination, and the coset design produces exactly such a combination.

Exactness instead of estimation

A positivity argument ordinarily loses constants. This one cannot afford to: an error of relative size $x^{-1/2+\varepsilon}$ requires the main term on each side to be an identity. So every quantity of leading order is given by a formula: Stirling's for the Fock norm, Chowla–Selberg for the period, Moret-Bailly's for the theta line, the mass formula for the lattices in a genus, the arithmetic Hilbert–Samuel formula for the volume. The proof resembles an exact computation in which a single step, the slope inequality itself, is an inequality.

Non-abelian methods for an abelian problem

Dirichlet L-functions belong to GL_1. Their arithmetic is class field theory, and the classical approaches stay within it. This proof leaves it at once: the lattices live on unitary groups of growing

rank, the comparison of integral structures runs through the theta correspondence, and the error terms are controlled by automorphic forms on GL_3 and their Galois representations. The abelian L-function appears only at the very end, as a character of the Galois group of a host acting on Frobenius classes. Information about GL_1 turns out to be most accessible from larger groups, which is the guiding principle of the Langlands programme.

Everything is chosen at each scale

No single object carries the proof. The CM field, the auxiliary primes, the level, the chain length and the set of translates are all chosen afresh for each scale x, and the bound is obtained only on a window around that scale. This is what lets each difficulty be met by averaging: good primes exist because most primes are good, the design exists because translates can be distributed as required, and the family exists because large families exist in every class. Landau's theorem on windows is what makes scale-by-scale information sufficient. Varying auxiliary arithmetic data in order to prove a statement about a fixed L-function is familiar from amplification and from the Taylor–Wiles method; here it organizes the whole argument.

Where the proof sits in mathematics

Three areas meet at the slope inequality. **Arakelov geometry** provides the inequality and the exact formulas for slopes. **p-adic analysis** provides the dictionary between primes and lattices: formal groups and divided powers convert the position of Frobenius into a forced divisibility. **The Langlands programme** provides the lattices and the uniformity: the theta correspondence moves integral structures between groups, and Galois representations bound the indices that the comparison introduces. Analytic number theory enters at two points only, Landau's theorem on windows and the mean-square selection of auxiliary primes, and in neither is a zero of an L-function located.

Overview of the Lean Source

A formal proof is read differently from a paper. A paper asks its reader to check every step. A formal proof asks the reader to check only two things, what was proved and what was assumed, and leaves the steps to the machine. This section is organized around that difference. It first isolates the small part of the development that has to be read in order to trust the result. It then follows the chain of declarations that carries the argument, explains how the argument is written down, and shows where its seventeen million lines go. It closes with a suggested reading order, followed by reference tables. Paths are relative to the `src` directory of the development.

What has to be trusted

The development is large: about 175,000 files and 17.4 million lines of Lean 4, with roughly 813,000 theorems, built with Lean 4.33.1 against a pinned commit of Mathlib. Almost none of it has to be read to believe the result. A theorem accepted by Lean's kernel is correct relative to three things: the kernel itself, the axioms the theorem depends on, and the definitions that appear in its statement. Each of the three is small here.

The trusted base

- **The statement.** `Statements.lean` has 16 lines and `Main.lean` has 12. The final theorem has type `RiemannHypothesis`, which is Mathlib's definition, built on Mathlib's `riemannZeta`; the Dirichlet statements use Mathlib's `DirichletCharacter.LFunction`. None of these names is redefined or shadowed anywhere in the development.
- **The axioms.** The source contains no `sorry` and declares no axioms. `#print axioms riemannHypothesis` reports only Lean's three standard axioms, `propext`, `Classical.choice` and `Quot.sound`.
- **The kernel.** Every declaration passes through it, including those produced by metaprograms.

The rest of the configuration does not enlarge this base. Metaprogramming appears in three places: a tactic idiom that reverts the local context before applying a lemma, a tactic that repackages structures using the kernel's own definitional-equality check, and one command, in `Formalism/HermitianCurvature/HermitianCurvatureProofs/CutCptCopy.lean`, that adds copies of declarations with renamed constants. All three produce terms that the kernel checks like any other. The build raises Lean's heartbeat limits, globally to four million and further in about 54,000 individual files; heartbeats bound how long elaboration may run and have no bearing on soundness. The five files of `BuildAid` only hide internal namespaces from Lean's library-suggestion index.

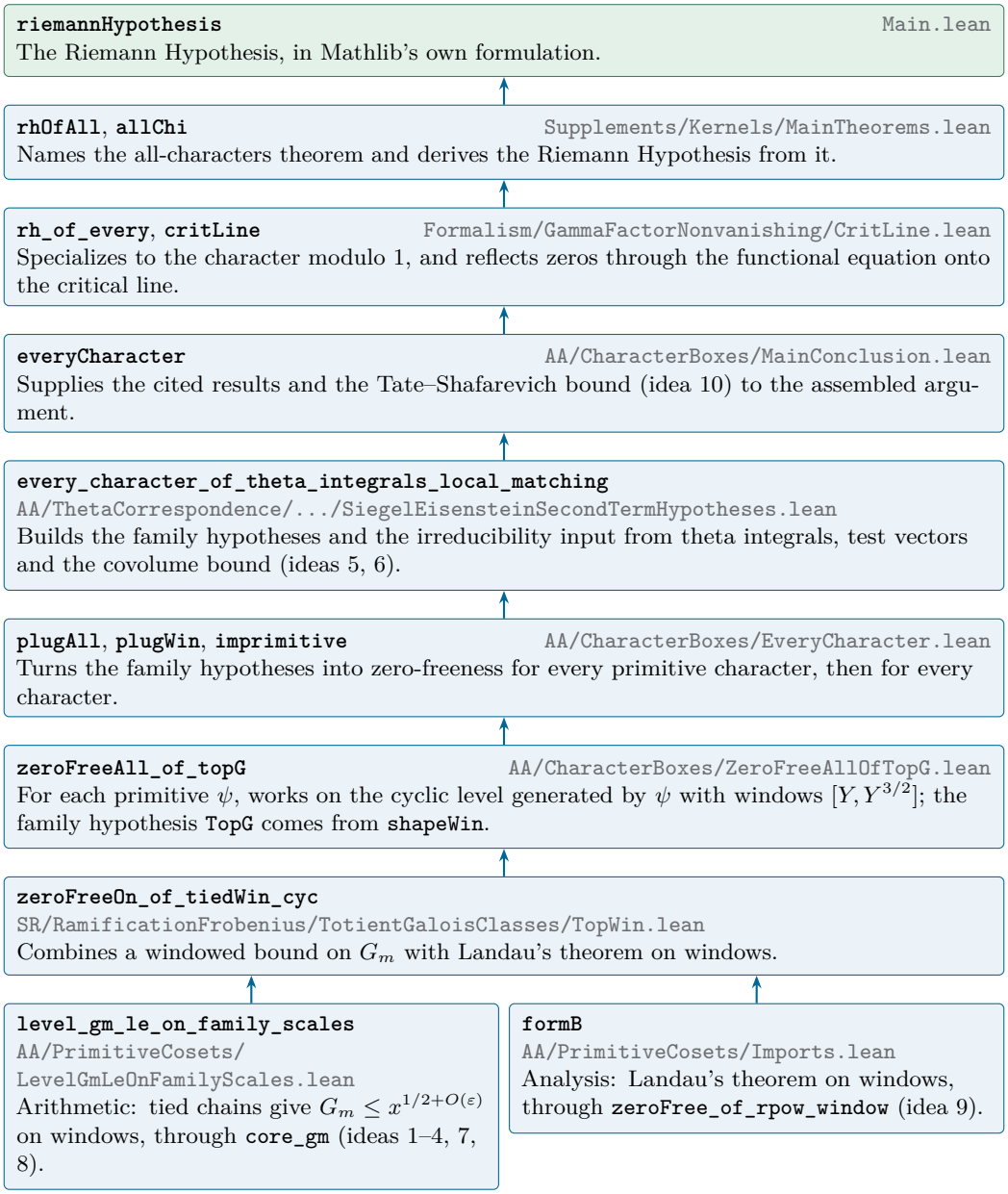
One thing the trusted base does not provide is a guarantee about intermediate statements. Many foundational results are formalized in the form the argument consumes rather than in full generality: the local–global compatibility of Galois representations covers the abelian and tame cases that are used, and the Arakelov layer models arithmetic varieties by multihomogeneous prime ideals of polynomial rings over $\mathbb{Z}$, with Hermitian metrics as positive homogeneous functions. This matters when citing an intermediate theorem as a formalization of a named result. It does not matter for

the final theorem, whose statement uses only Mathlib's definitions.

The names of intermediate structures are also a weaker guide than their statements. As Nielsen points out [2], some of them are satisfied more cheaply than their names suggest. The predicate of geometric conditions is instantiated by `anyGeometry`, which is `True` (`OverIx.lean`). The divisor datum is built from a degenerate model (`Inhabited.lean`). The height in the global half of the slope inequality is defined to be the explicit main-term formula `hExpl6`, so the equation relating the two holds by `rfl` (`SetUnipotentTranslate.lean`). None of this affects the final theorem, since a true intermediate statement, however weak, cannot lead to a false conclusion. It does mean that the content of the global half lies in the operator-norm, covolume and valuation estimates collected in the record `DivInputsR`, not in the geometry of line bundles that the names describe. Everything after this point is for understanding the proof, not for believing it.

Following the argument down

To understand the proof, start where the kernel ends: at the final theorem. A short chain of declarations leads from it to the place where the arithmetic and the analysis meet.



The spine of the formal proof, read from the bottom up. Each box names the declarations at one step and the file that contains them; AA abbreviates Development/AutomorphicArithmetic and SR abbreviates Formalism/SharpRestriction.

The first three steps are bookkeeping. rhOfAll names the statement for all characters; rh_of_every specializes it to the trivial character modulo 1, whose L-function is ζ; and critLine uses the functional equation of the completed L-function to move zeros from the left half of the strip to the right, where they are excluded. Nothing in these steps is deep, but they are where the formal statement meets Mathlib's definitions, so they are worth reading once.

The next two steps are where the large inputs enter. `everyCharacter` supplies the cited theorems and the Tate–Shafarevich bound, both through records proved elsewhere in the development. The theta step consumes the geometry: theta integrals, test vectors and the covolume bound go in, and the family hypotheses come out. This is where the material on Shimura varieties, Arakelov positivity and the theta correspondence is designed to reach the main line; as noted above, its geometric conditions are instantiated trivially on the route the final theorem takes.

The remaining steps are the argument of Parts I to III in compressed form. `plugAll` reduces arbitrary characters to primitive ones. `zeroFreeAll_of_topG` fixes a primitive ψ, works on the cyclic level that ψ generates, and asks for the bound on windows $[Y, Y^{3/2}]$. `zeroFreeOn_of_tiedWin_cyc` is the meeting point: it takes the windowed bound on G_m from the arithmetic branch and Landau's theorem on windows from the analytic branch. The analytic branch is short and self-contained. The arithmetic branch is where the rest of the development lives.

Every arithmetic idea reaches that branch through one of three statements in `SR/GlobalTheoremInputs.lean`. `LinkG_at` is the local half of the slope inequality, the forced divisibility at inert primes. `LinkDIV_at` is the global half, the bound on slope and height by the Stirling–period main term. `LinkN` is the Tate–Shafarevich error. These three statements are the formal counterparts of the inequality at the centre of the proof, and they are the best place to begin reading its arithmetic. On the route the final theorem takes, the divisor link is proved for a height assigned to equal the formula `hExpl6`, from the matrix estimates collected in the record `DivInputsR`; for the global half, those estimates are where the content lies.

How the argument is written

The development is organized the way a long paper would be if every lemma had to be stated precisely before anything could use it.

The vocabulary comes first. `Formalism` defines the objects and propositions of the argument: hosts, families, levels, the sums G_m, the links. Most of its top-level files are named `Statement*.lean`, and they contain definitions of propositions rather than proofs. `Development` proves those propositions, and in its two largest libraries the material is sorted into `Definitions`, `Statements`, `Proofs` and `PublishedTheorems`. Individual results are commonly split the same way, across a file whose name ends in `Statement`, which fixes the exact form of the result under a root-namespace name, and a file that proves it; about 48,000 file names follow this convention.

The main argument is then written against *records*: structures whose fields are the results it needs. A record states the contract and a separate theorem fulfils it, so the assembly reads as a list of hypotheses in their exact form, and each proof can be read as discharging one of them.

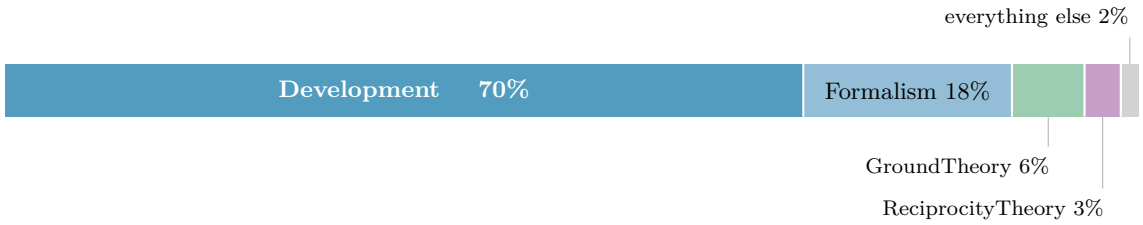
Record	Contents
`OpenLiteratureStatements`	three fields: the existence of Galois representations with local–global compatibility (twice, under two names) and the local Langlands correspondence for GL_n. It is inhabited in `MainConclusion.lean` by the theorems `galoisLocalGlobalStrong_recParameter` and `recTExists`, both proved in the development. About 150 lemmas named `*_of_literature` derive cited results from it; nearly all of them refer directly to in-repo proofs
`Hyps`, `HypsAlt`	about thirty fields over the families of hosts: row and twist conditions, size bounds, a class-number formula, point counts, decay of archimedean matrix coefficients, the mass of a definite Hermitian genus
`UnitaryGroupThird-PlaceHypotheses`	about sixty named results on unitary groups: Siegel–Weil, the doubling method, Bernstein–Zelevinsky, Casselman–Wallach, the Weil representation, local Tate duality
`ThetaLocalInputs`	sixteen local results on theta lifts and their test vectors
`PublishedHodgeBundleTheorems`	the geometric results on Shimura varieties and the Hodge bundle: toroidal compactification, Kawamata–Viehweg and Nadel vanishing, L^2 extension, arithmetic Hilbert–Samuel, Chow forms

The record named for the literature deserves a word. Its fields are the deepest results the proof uses, the Galois representations attached to automorphic forms and the local Langlands correspondence, and the argument is written so that it could take them as hypotheses. It does not: the record is filled by theorems proved in the development, and so the final theorem depends on no external result.

The top-level declarations live in a few namespaces. `UpperTheory` holds the assembly, with `UpperTheory.Top` for the final reduction, `UpperTheory.FinalTheorems` for the closing theorems and `UpperTheory.LiteratureTheorems` for the record of cited results. `GlobalTheory` holds the families and links, and `Formalism.HC` holds the characters, levels and zero-free statements of the analytic endgame.

Where the seventeen million lines go

The ten ideas are not where the volume is. The assembly occupies a few dozen files, and the statements of the argument fit in `Formalism`, a fifth of the whole. Most of the development proves foundations: resolution of singularities, the arithmetic Hilbert–Samuel theorem, toroidal compactifications, the local Langlands correspondence, Galois representations attached to automorphic forms. A paper would cite each of these in a line. A formal proof cannot cite, and Mathlib does not yet contain them, so the development proves them.



Share of the 17.4 million lines by library. The final theorem and its assembly occupy a few dozen files; everything else is the material they rest on.

Library	Files	Lines	Contents
`Main`	2	28	`Statements.lean`, which defines the headline statements, and `Main.lean`, which proves the final theorems
`Supplements`	3	22	the kernel files that connect `Main.lean` to the assembled argument
`Formalism`	18,254	3.20M	statements of the argument's central objects and inequalities; the analytic endgame; the families, hosts, links and local lattice machinery
`Development`	141,928	12.23M	the bulk of the proofs: abelian varieties and CM theory, automorphic arithmetic and the final assembly, automorphic spectra, Arakelov positivity, étale cohomology, resolution of singularities
`GroundTheory`	8,507	1.12M	heights and theta functions, class field theory and global fields, profinite and Tate cohomology, local fields, p-adic Hodge theory
`ReciprocityTheory`	3,845	0.56M	reductive groups and Arakelov geometry, discrete series, L^2 estimates, geometry of numbers, Chebotarev
`CompanionTheory`	2,027	0.19M	representation theory at the real place: spectral geometry, Bessel models, Harish-Chandra modules
`PositivityTheory`	449	55k	the outright forms of the arithmetic Hilbert–Samuel and Moriwaki bigness theorems
`Lemmata`	167	9k	general-purpose lemmas shared across libraries
`BuildAid`	5	20	build configuration: excludes five internal namespaces from Lean's library-suggestion index

Statements off the main line

Alongside the argument, the development contains several conditional reductions to the Riemann Hypothesis that the final theorem does not use. Each states, as a proposition, that some estimate implies the hypothesis, and each sits beside the material it reduces to. None of them is reached from the chain of declarations above, so a reader tracing the proof can set them aside:

- `Formalism/HalfLinks/Statement.lean`, a version of the argument over a fixed Galois group with CM types and Chebyshev-node averaging;
- the `RiemannHypothesisOf...` propositions in `Formalism/SizeWindowBrackets/`, which reduce the hypothesis to growth bounds for G_ζ alone;
- two weighted-sieve reductions, in `Formalism/ThickAverageLoss/` and `Formalism/WideWeightTheorem/`;
- the form of Landau's principle that assumes the bound for all large y, `zero_free_of_Chebyshev_growth`.

A reading order

The most efficient path through the development follows the structure of this section.

1. `Statements.lean` and `Main.lean`, to see exactly what is proved.
2. The chain of declarations, from `CritLine.lean` down to `TopWin.lean`.
3. The file `ExplicitFormulaErrorOfHostConditions.lean`, for the definitions of levels, G_m and the host conditions.
4. `SR/GlobalTheoremInputs.lean` and `SR/FamCore.lean`, for the families and the three links.
5. `Formalism/LocalRingCardinality/DataArith.lean`, for Hermitian lattices, slopes and heights.
6. Then, idea by idea, the files listed in the reference tables that follow.

After the first two items a reader knows what is proved and how it is assembled. After the next three, they know the statements the whole arithmetic is built to satisfy. Everything else is the proof of those statements.

Reference: where each part of the argument lives

The table below locates every part of this document in the source. The abbreviations are `AA` for `Development/AutomorphicArithmetic`, `AV` for `Development/AbelianVarietyTheory` and `SR` for `Formalism/SharpRestriction`.

Topic	Declarations and files
The Theorem	`riemannHypothesis` in `Main.lean`; statements in `Statements.lean`
Part I, the sums	`Sz`, `Gz`, `Gchi` in `Formalism/StatementOC.lean`; `Level.Gm` in `SR/RamificationFrobenius/TotientGaloisClasses/ExplicitFormulaErrorOfHostConditions.lean`
Part I, reflection	`critLine`, `rh_of_every` in `Formalism/GammaFactorNonvanishing/CritLine.lean`
Idea 1, slopes	`HLat`, `lcv`, `muDual`, `DivData` in `Formalism/LocalRingCardinality/DataArith.lean`; `LinkDIV_at` in `SR/GlobalTheoremInputs.lean`; `Formalism/HarderNarasimhan/`, `Formalism/LatticeSlopes/`
Idea 2, theta kernel	`LinkG_at` in `SR/GlobalTheoremInputs.lean`; `SR/GrowthBounds/FormalGroupSystem.lean`, `DataG.lean`, `DataGChain.lean`; `Formalism/GradedWeylDivisibility/`
Idea 3, hosts	`HostRaw`, `HostHyp`, `Frame` in `SR/RamificationFrobenius/TotientGaloisClasses/`; `D7` in `Formalism/LocalRingCardinality/RhoDetPrimeEqMuTildeMulChiPow.lean`
Idea 4, coset design	`Frame.lam`, `Frame.Shares`; `tops_sum_gen` in `AA/PrimitiveCosets/TopsSum.lean`; `AskMainBalance` in `Formalism/ThickAverageLoss/`
Idea 5, constants	`BlockBound` in `AV/BlockTheory/DivisorBase/PaddedLastEntry.lean`; `AV/Proofs/CubicalLineBundles/`; `GroundTheory/HeightsAndTheta/Heights/ChowlaSelberg/`
Idea 6, transfer	`IndefSide.boundsMu` in `Formalism/LocalToGlobal/IndefiniteForms/IndefUse.lean`; `CovolumeBoundOfX` in `AA/ThetaCorrespondence/ThetaPairLifts/ThirdPlaceSeries/LiftLambdaBound.lean`
Idea 7, tied chains	`ChainTied` in `SR/GlobalSchurBounds/TieGlobal.lean`; `core_gm` in `AA/PrimitiveCosets/CoreGm.lean`; `RestSize` in `Formalism/SizeWindowBrackets/SChain.lean`
Idea 8, random twists	`PairGood`, `PairEx` in `AA/CharacterBoxes/PairGood.lean`; proofs in `AA/MeanSquareBounds/`
Idea 9, windows	`zeroFree_of_rpow_window` in `AA/DedekindLFunctions/LandauTheorems/ZeroFreeOfRpowWindow.lean`; `landauWindow.lean`
Idea 10, Sha bound	`VWin` in `AA/CharacterBoxes/ExistsPageN.lean`; `OneChar` in `AA/AutomorphicVarieties/UnitaryThreeGroups/KatoUse.lean`; `LinkN` in `SR/GlobalTheoremInputs.lean`
Part IV, assembly	`everyCharacter` in `AA/CharacterBoxes/MainConclusion.lean`; `shapeWin`, `zeroFreeAll_of_topG` in `AA/CharacterBoxes/`
Part IV, foundations	`PositivityTheory/`, `ReciprocityTheory/`, `GroundTheory/`, `CompanionTheory/`, `Development/`

Reference: the libraries in detail

The tables below list the largest subdirectories of each library, with their file counts and the part of the argument they serve.

Formalism

Subdirectory	Files	Role
`LocalToGlobal`	5,501	Siegel's theory of lattices (2,085 files), indefinite forms, torsion and divisibility, kernel lattices: the lattice side of the theta transfer (idea 6)
`PeriodIntegrals`	4,350	the mass formula for unitary groups (3,615 files) and bounds on orbital integrals
`SharpRestriction`	3,093	the core of the argument: the local theta kernel in `GrowthBounds` (1,210 files, idea 2), level cocycles, cusp-form norm bounds, divisibility, tied chains in `GlobalSchurBounds` (idea 7), hosts and the analytic endgame in `RamificationFrobenius` (ideas 3, 4, 9), and the families and links in its top-level files
`HermitianCurvature`	1,922	relative trace formula on unitary groups: orbital integrals, Herglotz–Bochner positivity, stationary phase
`GL3MeanSquares`	440	the large sieve over prime moduli and zero-density estimates for the auxiliary primes (idea 8)
`LocalRingCardinality`	173	the hosts `D7`, Hermitian lattices `HLat`, slopes, and the scale parameters of the families (ideas 1, 3, 5)
`HarderNarasimhan`, `LatticeSlopes`, `SuccessiveMinima`	small	Harder–Narasimhan filtrations, the slope inequality, Minkowski's second theorem (idea 1)
`GammaFactorNonvanishing`, `ZetaNonvanishing`, `LandauZeroFreeBounds`	small	the reflection to the critical line, Landau's principle, classical zero-free regions and Siegel's theorem
about a hundred further directories		specialized toolkits: Weyl-module divisibility, formal groups, Wach lattices, Mahler-type p-adic interpolation, Capelli and Yangian identities, Bessel functionals

Development

Subdirectory	Files	Role
`AbelianVarietyTheory`	99,187	abelian varieties and CM theory, sorted into `Proofs` (95,418 files), `Definitions`, `BlockTheory`, `PublishedTheorems` and `Statements`; includes the CM periods and Chowla–Selberg formulas (`CubicalLineBundles`, idea 5), the jet-norm bound `BlockBound`, and the existence of auxiliary prime pairs from functoriality
`AutomorphicArithmetic`	18,431	the final assembly and its arithmetic inputs: `AutomorphicVarieties` (7,959 files, including the Tate–Shafarevich bound on U(3), idea 10), `Foundations` (Galois representations and local Langlands), `ThetaCorrespondence` (the final steps of the spine), `IntegralCasts` (Capelli identities), `CharacterBoxes`, `PrimitiveCosets`, `MeanSquareBounds` (idea 8) and `DedekindLFunctions` (idea 9)
`AutomorphicSpectrum`	15,147	spectral theory of automorphic forms
`WeilBound`	7,051	Siegel–Weil kernels, the Moriwaki and Hilbert–Samuel chains, and the representation theory of GL_n
`EtaleCohomology`	1,398	étale sheaves, Tate twists and Kummer theory
`ResolutionOfSingularities`	499	resolution and principalization in characteristic zero
`DiscreteGroups`	215	arithmetic subgroups and their quotients

The foundation libraries

Subdirectory	Files	Role
`GroundTheory/HeightsAndTheta`	6,410	heights, Eisenstein series, the Chowla–Selberg formula and theta functions
`GroundTheory/ProfiniteCohomology`	450	Poitou–Tate duality, the global Euler characteristic, class formations
`GroundTheory`, other	1,647	automorphic forms, global and local fields, p-adic Hodge theory, Brauer groups, Tate cohomology
`ReciprocityTheory/ReductiveGroups`	2,799	Arakelov geometry (1,623 files, including the Moriwaki statements) and discrete series (1,021 files)
`ReciprocityTheory`, other	1,046	L^2 $\bar{\partial}$-estimates, analytic forms, dual lattices, class field theory, Chebotarev, semistable reduction, geometry of numbers
`CompanionTheory`	2,027	spectral geometry (1,254 files), Bessel models, real groups, Harish-Chandra modules
`PositivityTheory`	449	the outright arithmetic Hilbert–Samuel and Moriwaki theorems used by the covolume bound

References

[1] L. Alpöge, *Remark*, note accompanying the Lean source of the formalization, 20 November 2026, `https://alpo.ge/remark.pdf`.

[2] J. L. Nielsen, "The Invalidity of Alpöge's Proposed Proof of the Riemann Hypothesis: Vacuously true geometric and lattice witnesses in the Lean formalization," preprint, Center for Topological Physics, Lawrence, Kansas, 20 November 2026, 13 pp.

[3] A. V. Sutherland, "RH Lean proof: verification files," status page, MIT, updated 21 November 2026, 09:30 UTC, `https://math.mit.edu/~drew/grh/`.

[4] Lean and Mathlib communities, "Lean verification of the Riemann Hypothesis," *Lean Community Blog*, 21 November 2026, endorsed by the Lean FRO and the Mathlib Initiative, `https://leanprover-community.github.io/blog/posts/riemann-hypothesis/`.

[5] C. Moskowitz, "The AI Riemann hypothesis proof is almost certainly true, independent mathematicians attest," *Scientific American*, 20 November 2026, `https://www.scientificamerican.com/article/the-ai-riemann-hypothesis-proof-is-almost-certainly-true-independent-mathematicians-attest/`.

[6] P. L. Chebyshev, "Mémoire sur les nombres premiers," *J. Math. Pures Appl.* 17 (1852), 366–390.

[7] E. Landau, "Über einen Satz von Tschebyschef," *Math. Ann.* 61 (1905), 527–550.

[8] A. Weil, *Sur les courbes algébriques et les variétés qui s'en déduisent*, Hermann, Paris, 1948.

[9] S. Chowla and A. Selberg, "On Epstein's zeta-function," *J. reine angew. Math.* 227 (1967), 86–110.

[10] S. A. Stepanov, "On the number of points of a hyperelliptic curve over a finite prime field," *Izv. Akad. Nauk SSSR Ser. Mat.* 33 (1969), 1171–1181.

[11] J.-B. Bost, "Périodes et isogénies des variétés abéliennes sur les corps de nombres (d'après D. Masser et G. Wüstholz)," Séminaire Bourbaki, exp. 795, *Astérisque* 237 (1996), 115–161.