Window methods for Dirichlet L-functions

Analytic criteria and conditional arithmetic reductions

A.U.Thor

Abstract

We develop a one-sided criterion for the nonvanishing of Dirichlet L-functions from estimates on arbitrarily large multiplicative intervals. The principal analytic object is a positive weighted combination of corrected, smoothed von Mangoldt sums. Its Mellin transform has simple poles at the zeros of the underlying L-functions, with residues that cannot cancel. On the logarithmic scale, concentration of normalized Laplace kernels turns window estimates into bounds for high derivatives. A pole separation argument then yields a zero-free half-plane. We give the transform calculations, concentration estimates, and pole argument in detail, including quantitative choices and several extensions of the window hypothesis. We also develop the finite-character, spectral, field-theoretic, and cohomological estimates that organize an arithmetic application. Detailed developments include restriction kernels on Hermitian domains, finite unitary groups and congruence quotients, interlacing dimension formulas, slope filtrations, integral monodromy, and the effect of finite-order symmetries on formal jets. The final arithmetic reduction is conditional on specified window and comparison hypotheses. No unconditional critical-line theorem is asserted without those hypotheses.

Contents

1 Introduction and the structure of the argument **3**

2 Finite characters and positive averaging **4**

3 Smoothing, summation, and Mellin kernels **7**

4 Continuation to the right half-plane and boundary nonvanishing **9**

5 Fourier analysis underlying the theta transformation **12**

6 Gamma factors, Gauss sums, and functional equations **16**

7 Growth, zero counting, and logarithmic derivatives **20**

8 Smoothed character sums and their transforms **25**

9 A one sided window criterion **28**

10 The Laplace transform argument **31**

11 The small parameter reduction 38

12 Quantitative anatomy of the Laplace kernels 41

13 Variants and stability of the window hypothesis 44

14 Imprimitive characters and the critical line 47

15 Spectral estimates and finite-dimensional comparison 49

16 Restriction operators, reproducing kernels, and orthogonality 51

17 Hermitian domains and explicit reproducing kernels 58

18 Localization kernels and weighted length bounds 63

19 Cohomological exact sequences and local conditions 66

20 Cyclotomic subfields and arithmetic scales 72

21 Finite unitary geometry and congruence-level counts 75

22 Representations, integral structures, and compatibility data 79

23 Local representations and integral monodromy 82

24 Weights, interlacing patterns, and dimension estimates 86

25 Lattice indices and local length budgets 91

26 Euclidean lattices and slope filtrations 93

27 Hermitian lattices over number fields 98

28 From a family of restriction maps to a slope bound 103

29 Formal coordinates, integral jets, and character symmetries 108

30 Prime classes and corrected character expansions 114

31 Scale calculus and simultaneous selection 119

32 Cohomological windows and conditional arithmetic comparison 123

1 Introduction and the structure of the argument

For a Dirichlet character χ, a central analytic question is whether

$$L(s,\chi) \neq 0 \qquad \text{for } \operatorname{Re} s > \tfrac{1}{2}, \quad (\chi, s) \neq (\chi_0, 1), \tag{1}$$

where χ_0 is the principal character at the given modulus. The exception refers to a meromorphic pole. The purpose of this article is to identify and prove an analytic implication that reduces this question to upper bounds for a smoothed arithmetic expression on a sparse collection of long intervals.

The sparseness is substantial. An estimate need not hold for every large x, or even on a set with a prescribed density. For a fixed $\delta > 0$ it is enough to have an interval $[Y, Y^{1+\delta}]$ beyond every prescribed lower threshold. Both the width and the location of these intervals matter: after taking logarithms, their lengths are proportional to their left endpoints. This geometry supports an exponentially concentrated family of integral kernels.

The estimate is also one-sided. A bound for a real weighted sum of errors need not bound the absolute value of that sum, or any individual summand. The reason it still detects zeros is a special residue sign. Strictly positive weights prevent cancellation of poles, and a pole on the real axis has positive residue. The proof retains a real-axis derivative term that would be lost in a purely absolute-value treatment.

The argument is organized in three stages. First we construct corrected smoothed sums and calculate their Mellin transforms. Second we prove an abstract exclusion theorem for poles of Laplace transforms under an integral window hypothesis. Third we describe quantitative conditions under which arithmetic estimates produce the required windows with exponents approaching $1/2$.

The third stage is a conditional reduction. The consequences of the cohomological and spectral estimates are proved here through exact sequences, logarithmic indices, slope filtrations, and a common-map corank theorem. The existence of arithmetic families satisfying all the stated hypotheses is not deduced from those consequences. In particular, the existence of auxiliary primes, the cohomological window bound, and compatible geometric comparison data are distinct requirements. Their separation is part of the mathematical statement, not an omission from its quantifiers.

1.1 Analytic conventions and classical inputs

Write $\Lambda(n)$ for the von Mangoldt function and γ for Euler's constant. Complex powers of a positive real number are defined by $x^s = \exp(s \log x)$ using the real logarithm. Intervals in integrals carry Lebesgue measure. A finite character family may contain characters of different conductors. A principal member of an admissible family is always the character of modulus 1; larger principal moduli are treated by their finite Euler factors.

Continuation to $\operatorname{Re} s > 0$ and nonvanishing on $\operatorname{Re} s \geq 1$ are proved below. At the boundary the proof is separate from the absolutely convergent Euler product. This matters at $s = 1$, where the correction for a nonprincipal character involves $(L'/L)(1, \chi)$. The functional equations and continuation to the full plane are derived later from Poisson summation for Schwartz functions, with a proof of the Fourier identity and its convergence statements.

An assertion of an analytic extension always includes agreement with the initial function on its original domain. A pole is an actual singularity with nonzero principal part. These conventions prevent a set of artificially marked points from entering the pole exclusion argument.

2 Finite characters and positive averaging

2.1 Characters as functions on a finite group

Let A be a finite abelian group, written multiplicatively. A complex character of A is a homomorphism $\xi : A \to \mathbb{C}^\times$. If $a \in A$ has order m, then $\xi(a)^m = 1$. Consequently $|\xi(a)| = 1$ and $\overline{\xi(a)} = \xi(a)^{-1} = \xi(a^{-1})$. The pointwise product of two characters is again a character, and the set $\widehat{A}$ of characters is an abelian group.

The unit-modulus property has two different uses. It makes averaging operators bounded, and it turns conjugation into inversion in orthogonality calculations. Neither use requires choosing arguments of roots of unity. Such arguments are rarely compatible across products, whereas the group identities are exact.

Lemma 2.1 (Character orthogonality)**.** *For a character ξ of A,*

$$\sum_{a \in A} \xi(a) = \begin{cases} |A|, & \xi = 1, \\ 0, & \xi \neq 1. \end{cases}$$

For $\xi, \eta \in \widehat{A}$,

$$\frac{1}{|A|} \sum_{a \in A} \xi(a)\overline{\eta(a)} = \begin{cases} 1, & \xi = \eta, \\ 0, & \xi \neq \eta. \end{cases}$$

Proof. The trivial character has constant value 1. If ξ is nontrivial, choose $b \in A$ with $\xi(b) \neq 1$. Multiplication by b permutes A, so with $S = \sum_a \xi(a)$ one has $S = \sum_a \xi(ba) = \xi(b)S$. Hence $(1 - \xi(b))S = 0$, and $S = 0$. Apply this result to $\xi\eta^{-1}$ for the second identity. □

For functions $f, h : A \to \mathbb{C}$, use the normalized inner product

$$\langle f, h \rangle_A = \frac{1}{|A|} \sum_{a \in A} f(a)\overline{h(a)}.$$

Distinct characters are orthonormal. The elementary decomposition of a finite abelian group into cyclic groups shows that $|\widehat{A}| = |A|$: a cyclic group of order m has exactly the m characters determined by sending a generator to an mth root of unity, and characters of a direct product are products of characters. Thus the characters form an orthonormal basis of the $|A|$-dimensional space of functions on A.

Proposition 2.2 (Fourier inversion and Parseval)**.** *Define $\widehat{f}(\xi) = \langle f, \xi \rangle_A$. Then*

$$f(a) = \sum_{\xi \in \widehat{A}} \widehat{f}(\xi)\xi(a), \qquad \frac{1}{|A|} \sum_{a \in A} |f(a)|^2 = \sum_{\xi \in \widehat{A}} |\widehat{f}(\xi)|^2.$$

Proof. Expand f in the orthonormal character basis. Taking the inner product with a basis vector identifies its coefficient as $\widehat{f}(\xi)$. Taking the inner product of the expansion with itself gives the second identity because all cross terms vanish. □

2.2 Restriction to subgroups

For a subgroup $A_0 \subset A$, its annihilator is

$$A_0^{\perp} = \{\xi \in \widehat{A} : \xi(a) = 1 \text{ for every } a \in A_0\}.$$

Applying orthogonality on A_0 gives

$$\frac{1}{|A_0|} \sum_{a \in A_0} \xi(a) = \mathbf{1}_{A_0^{\perp}}(\xi). \tag{2}$$

More generally, averaging over a coset bA_0 multiplies the right side by $\xi(b)$. In particular, subgroup averaging kills a character unless its restriction to that subgroup is trivial. This is a projection mechanism, not a cancellation estimate with an error term.

Define an averaging operator on functions by

$$(P_{A_0} f)(b) = \frac{1}{|A_0|} \sum_{a \in A_0} f(ba).$$

Then $P_{A_0}\xi = \xi$ for $\xi \in A_0^{\perp}$, and $P_{A_0}\xi = 0$ otherwise. It follows immediately from Fourier expansion that $P_{A_0}^2 = P_{A_0}$, that P_{A_0} is self-adjoint, and that $\|P_{A_0} f\|_2 \le \|f\|_2$. Its image consists of functions constant on cosets of A_0.

The exactness of this projection is useful when a comparison formula contains a large character family. Averaging can separate retained characters from a residual family. The residual family still needs an analytic estimate; orthogonality alone does not bound the values of the smoothed sums attached to the surviving characters.

2.3 Dirichlet characters and their primitive representatives

For a positive integer q, a Dirichlet character modulo q is a character of $(\mathbb{Z}/q\mathbb{Z})^{\times}$, extended to all integers by setting its value to zero when the integer is not coprime to q. This extension is completely multiplicative. If n is coprime to q, its character value has modulus 1; otherwise the value is zero. Hence $|\chi(n)| \le 1$ for every n.

The principal character modulo q is the indicator of coprimality with q. It is constant 1 only when $q = 1$. This distinction is responsible for the finite Euler factors in the passage from a general principal character to the zeta function.

A character modulo q is induced from a character modulo a divisor f if their values agree on integers coprime to q. There is a unique primitive inducing character, whose modulus is the conductor. Induction is a statement about unit groups; it does not identify the extended functions at integers divisible by primes dividing q but not f. For such integers the induced character is zero, while the primitive character may be nonzero.

For a fixed primitive character ψ of order r, the powers $1, \psi, \ldots, \psi^{r-1}$ form a finite group of characters at a common modulus. Replacing each power by its primitive inducing character produces a finite collection of primitive characters. Repetitions may occur and should be combined by adding their weights. The target ψ occurs, and inversion permutes the resulting collection. Thus this procedure provides a natural finite family containing a prescribed character and its conjugate without requiring a family of all conductors.

2.4 Euler products and logarithmic derivatives

For $\operatorname{Re} s > 1$ define

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

Absolute convergence follows from $|\chi(n)| \le 1$. The convergence is locally uniform in that half-plane, as is the differentiated series on any smaller closed half-plane $\operatorname{Re} s \ge 1 + \varepsilon$.

Proposition 2.3 (Absolutely convergent Euler product)**.** *For* $\operatorname{Re} s > 1$*,*

$$L(s,\chi) = \prod_{p} (1 - \chi(p) p^{-s})^{-1}, \qquad L(s,\chi) \neq 0.$$

Moreover,

$$-\frac{L'}{L}(s,\chi) = \sum_{n=1}^{\infty} \Lambda(n)\chi(n) n^{-s}. \tag{3}$$

Proof. Fix $\sigma > 1$. Each geometric series for a prime factor converges absolutely. The product over a finite set of primes expands, by unique factorization and complete multiplicativity, as the sum over positive integers whose prime factors lie in that set. As the set exhausts the primes, absolute convergence of $\sum n^{-\sigma}$ permits passage to the full Dirichlet series.

To prove nonvanishing, use the series

$$B(s) = \sum_{p} \sum_{m \ge 1} \frac{\chi(p)^m}{m} p^{-ms}.$$

For $\operatorname{Re} s \ge \sigma$, its absolute value is bounded by $\sum_p p^{-\sigma}/(1 - p^{-\sigma})$, which converges. Hence B is holomorphic in $\operatorname{Re} s > 1$ and $L(s,\chi) = \exp B(s)$ there. The exponential never vanishes. Differentiating the locally uniformly convergent double series gives $B'(s) = -\sum_p \sum_{m \ge 1} (\log p)\chi(p)^m p^{-ms}$. Grouping the prime powers yields (3). □

This proof constructs a logarithm in the initial half-plane. No global single-valued logarithm of the continued L-function is used later. The logarithmic derivative itself is meromorphic and is well defined without selecting a branch. At a zero, its principal part records the multiplicity, which is the local information needed by the window argument.

2.5 Positive weights and the meaning of positivity

Let I be finite and let $a_i > 0$. If nonnegative real numbers m_i are not all zero, then $\sum_i a_i m_i > 0$. This elementary fact is the exact positivity used in the residue calculation. The quantities m_i will be zero multiplicities. The smoothed errors being averaged may have either sign, and the residues themselves may be nonreal.

Allowing zero weights is harmless only after removing the corresponding characters from the family. A zero of a character with zero weight need not produce a pole of the combined transform. Negative weights are more serious: they can cancel multiplicities. Even two identical characters with weights 1 and -1 give the zero transform, which cannot detect any of their zeros. The strict positivity hypothesis therefore has mathematical content beyond making an average look natural.

3 Smoothing, summation, and Mellin kernels

3.1 The ramp as an integrated counting function

For a sequence of complex numbers $(b_n)_{n\ge 1}$, define

$$B(x)=\sum_{n\le x} b_n, \qquad R_b(x)=\sum_{n\le x} b_n\left(\frac{x}{n}-1\right), \quad x\ge 1.$$

Every sum is finite. The factor $x/n-1$ vanishes at the moving boundary $x=n$, so introducing a new term does not create a jump in R_b. Thus R_b is continuous, although B is generally a step function.

Lemma 3.1 (Two integral forms of the ramp)**.** *For $x\ge 1$,*

$$R_b(x)=\int_1^x\left(\sum_{n\le t}\frac{b_n}{n}\right)dt = x\int_1^x \frac{B(t)}{t^2}\,dt.$$

The value of an integrand at a jump point does not affect either integral.

Proof. For the first formula, integrate each of the finitely many summands on $[n,x]$. Its contribution is $b_n(x-n)/n$. For the second, interchange the finite sum with integration to obtain

$$x\sum_{n\le x} b_n\int_n^x t^{-2}\,dt = x\sum_{n\le x} b_n\left(\frac{1}{n}-\frac{1}{x}\right)=R_b(x).$$

□

At points away from the positive integers, $R_b'(x)=\sum_{n\le x} b_n/n$. Dividing the second identity by x and differentiating away from jumps gives $(R_b(x)/x)'=B(x)/x^2$. These identities show what information smoothing retains. They do not allow differentiation of an upper bound for R_b to obtain a bound for B: an inequality between functions does not in general imply the corresponding inequality between derivatives.

3.2 A precise Abel summation identity

Let f be continuously differentiable on $[1,x]$. Then

$$\sum_{n\le x} b_n f(n)=B(x)f(x)-\int_1^x B(t)f'(t)\,dt. \tag{4}$$

To verify the formula without assumptions at the integer endpoints, expand the integral as a finite sum:

$$\int_1^x B(t)f'(t)\,dt=\sum_{n\le x} b_n\int_n^x f'(t)\,dt = B(x)f(x)-\sum_{n\le x} b_n f(n).$$

This also proves the identity for complex sequences and complex f. The character application takes $b_n=\Lambda(n)\chi(n)$.

If $B(x) = O(x^{\alpha})$ and $\operatorname{Re} s > \max\{\alpha, 0\}$, the boundary term $B(x)x^{-s}$ tends to zero. Passing to the limit in (4) with $f(t) = t^{-s}$ gives

$$\sum_{n\geq 1} b_n n^{-s} = s\int_1^{\infty} B(t)t^{-s-1}\,dt,$$

provided the series is interpreted through the same partial sums. The integral converges absolutely under the displayed growth bound. This familiar passage relates counting estimates to analytic continuation, but the window argument below will not assume a global bound for B.

3.3 The Mellin transform of a ramp

For $\operatorname{Re} s > 1$ and $n \geq 1$,

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

The denominator has two factors because the ramp is one integration smoother than a sharp step. Both $s = 0$ and $s = 1$ are features of the kernel; they are present before any arithmetic sequence is chosen. Their role must be distinguished from zeros or poles of a Dirichlet series.

Suppose $\sum_n |b_n| n^{-\sigma} < \infty$ for some $\sigma > 1$. Then

$$\sum_n \int_n^{\infty} |b_n| \left(\frac{x}{n} - 1\right) x^{-\sigma-1}\,dx = \frac{1}{\sigma(\sigma-1)} \sum_n |b_n| n^{-\sigma} < \infty.$$

Consequently the Mellin transform of R_b can be evaluated termwise at every s with $\operatorname{Re} s = \sigma$. If the convergence holds locally uniformly in a half-plane, the identity is an identity of holomorphic functions there.

The point of this calculation is not only convergence. It shows that the same denominator occurs for every character. After correction at $s = 1$, the residue at a zero ρ will be a positive multiplicity times the single complex factor $-1/[\rho(\rho-1)]$. A character-dependent kernel could destroy that alignment even when all averaging weights were positive.

3.4 An elementary main-term example

The ramp can be analyzed without prime weights to see why a term of size $x \log x$ appears. Take $b_n = 1$ and write $N = \lfloor x \rfloor$. Then

$$R_b(x) = xH_N - N, \qquad H_N = \sum_{n=1}^{N} \frac{1}{n}.$$

The elementary harmonic-sum estimate $H_N = \log N + \gamma + O(1/N)$ yields

$$R_b(x) = x \log x + (\gamma - 1)x + O(1).$$

Indeed, $N = x + O(1)$ and $x \log(N/x) = O(1)$ for large x. This example is not an approximation to the von Mangoldt ramp with a claimed error term. It illustrates the effect of the kernel on a sequence with a first-order counting density. The precise linear correction for prime weights is instead determined by the Laurent expansion of the logarithmic derivative of the zeta function.

3.5 Changing scale and preserving the measure

The substitution $x = e^u$ converts a Mellin transform into a Laplace transform:

$$\int_1^\infty F(x)x^{-s-1}\,dx = \int_0^\infty F(e^u)e^{-su}\,du.$$

The factor $dx = e^u\,du$ cancels the extra x^{-1}. Omitting this Jacobian would shift every convergence and nonvanishing exponent by one. For weighted window integrals the same identity gives

$$\int_Y^{Y^{1+\delta}} R(x)x^{-\theta-1}\,dx = \int_{\log Y}^{(1+\delta)\log Y} R(e^u)e^{-\theta u}\,du.$$

The transformation is particularly useful because multiplication of the endpoint exponent becomes multiplication of the logarithmic endpoint. It is this proportional geometry, rather than mere largeness of the original interval length, that enters concentration of the derivative kernels.

3.6 Why a generous growth bound is sufficient

Suppose $F(x) = O(x(1+\log x)^2)$. Then $F(e^u) = O(e^u(1+u)^2) = O(e^{2u})$ for $u \geq 0$. The last estimate deliberately loses sharpness. It supplies one half-plane $\operatorname{Re} s > 2$ in which every differentiated Laplace integral converges, so all later derivative comparisons occur at centers in that fixed region.

More generally, a bound $|g(u)| \leq Ce^{\alpha u}$ gives convergence for $\operatorname{Re} s > \alpha$. The pole exclusion mechanism does not require α to be the true abscissa of convergence. Taking a larger α changes the size of the center used in the concentration estimate, but not the threshold θ singled out by the window majorant. This separation between a preliminary convergence line and a final zero-free line is a useful organizing principle throughout the proof.

4 Continuation to the right half-plane and boundary nonvanishing

4.1 Bounded partial sums of a nonprincipal character

Let χ be a nonprincipal Dirichlet character modulo q. Orthogonality on $(\mathbb{Z}/q\mathbb{Z})^\times$ gives $\sum_{n=1}^{q}\chi(n) = 0$. Periodicity therefore shows that

$$B_\chi(x) = \sum_{n\leq x}\chi(n)$$

is bounded in absolute value by q: every complete block of length q contributes zero, and the remaining block has at most q terms of modulus at most one.

Abel summation gives, initially for $\operatorname{Re} s > 1$,

$$L(s,\chi) = s\int_1^\infty B_\chi(x)x^{-s-1}\,dx. \tag{5}$$

The integral on the right converges absolutely for $\operatorname{Re} s > 0$. It is locally uniformly convergent there, as are its derivatives on compact subsets: powers of $\log x$ remain integrable against $x^{-1-\sigma}$ for every fixed $\sigma > 0$. Thus it defines a holomorphic continuation of $L(s,\chi)$ to the open right half-plane. In particular a nonprincipal L-function has no pole at $s = 1$.

For a principal character modulo q,

$$L(s,\chi_0) = \zeta(s)\prod_{p\mid q}(1-p^{-s})$$

in the initial half-plane. Once the zeta continuation has been constructed, this identity gives the corresponding continuation and shows that the only pole in $\operatorname{Re} s > 0$ is the simple pole at 1. Its residue is $\prod_{p|q}(1-p^{-1}) > 0$.

4.2 The zeta continuation and its Laurent coefficient

For $\operatorname{Re} s > 1$, Abel summation for the constant sequence gives

$$\zeta(s) = s\int_1^\infty \lfloor x\rfloor x^{-s-1}\,dx.$$

Writing $\lfloor x\rfloor = x - \{x\}$ yields

$$\zeta(s) = \frac{s}{s-1} - s\int_1^\infty \{x\}x^{-s-1}\,dx. \tag{6}$$

Because $0 \le \{x\} < 1$, the second term is holomorphic for $\operatorname{Re} s > 0$. Equation (6) therefore continues ζ to that region with a simple pole at 1 and residue 1.

The constant term can be evaluated directly. Define $H_N = \sum_{n=1}^N 1/n$ and $\gamma_N = H_N - \log N$. The difference

$$\gamma_N - \gamma_{N+1} = \log(1+1/N) - \frac{1}{N+1}$$

is positive, since the integral of $1/t$ over $[N, N+1]$ is greater than its minimum there. Integral comparison bounds γ_N below, so it decreases to a finite limit γ. Moreover,

$$0 \le \gamma_N - \gamma \le \sum_{k=N}^{\infty}\left(\frac{1}{k} - \frac{1}{k+1}\right) = \frac{1}{N}.$$

This proves $H_N = \log N + \gamma + O(1/N)$.

For integer $N \ge 2$, direct integration on the intervals $[n, n+1]$ gives

$$\begin{aligned}\int_1^N \frac{\{x\}}{x^2}\,dx &= \sum_{n=1}^{N-1}\left(\log\frac{n+1}{n} - \frac{1}{n+1}\right)\\ &= \log N - H_N + 1.\end{aligned}$$

Its limit is $1-\gamma$. Evaluating the holomorphic term in (6) at $s=1$ now gives

$$\zeta(1+t) = \frac{1}{t} + \gamma + O(t).$$

Thus the coefficient needed for the corrected principal ramp is obtained from an elementary integral calculation.

4.3 A positivity argument for real characters

The nonvanishing of $L(1,\chi)$ for a real nonprincipal character is not obtained by evaluating the Euler product at its boundary. We give a separate argument using positive Dirichlet coefficients.

Let χ be real and nonprincipal. In $\operatorname{Re} s > 1$,

$$\zeta(s)L(s,\chi) = \sum_{n\ge1} a_n n^{-s}, \qquad a_n = \sum_{d|n}\chi(d).$$

The coefficients are multiplicative. At a prime power,

$$a_{p^m} = \sum_{j=0}^{m} \chi(p)^j.$$

Since $\chi(p) \in \{0, 1, -1\}$, this is respectively 1, $m+1$, or 1 for even m and 0 for odd m. In every case $a_n \geq 0$, and in particular $a_{n^2} \geq 1$ for every n.

Proposition 4.1. *If χ is a real nonprincipal Dirichlet character, then $L(1, \chi) \neq 0$.*

Proof. Suppose $L(1, \chi) = 0$. Its zero cancels the sole pole of ζ in the right half-plane, so $F(s) = \zeta(s)L(s, \chi)$ is holomorphic on $\operatorname{Re} s > 0$.

At the real point $s = 2$, termwise differentiation in the absolutely convergent Dirichlet series gives

$$\frac{(-1)^k F^{(k)}(2)}{k!} = \sum_{n \geq 1} a_n n^{-2} \frac{(\log n)^k}{k!} \geq 0.$$

To justify the differentiation, note $a_n \leq d(n)$ and that the series $\sum d(n)n^{-\sigma} = \zeta(\sigma)^2$ converges for every $\sigma > 1$; logarithmic powers can be absorbed by a small decrease in σ.

The disk $|s - 2| < 2$ lies in the holomorphic domain. The Taylor series at 2 therefore converges at $s = 1/4 = 2 - 7/4$, and

$$F(1/4) = \sum_{k \geq 0} \frac{(-1)^k F^{(k)}(2)}{k!} (7/4)^k.$$

All terms after inserting the Dirichlet coefficient expression are nonnegative. Tonelli's theorem permits their interchange, even before finiteness is known, and gives

$$\begin{aligned} F(1/4) &= \sum_{n \geq 1} a_n n^{-2} \sum_{k \geq 0} \frac{((7/4) \log n)^k}{k!} \\ &= \sum_{n \geq 1} a_n n^{-1/4} \geq \sum_{m \geq 1} (m^2)^{-1/4} = \infty. \end{aligned}$$

This contradicts finiteness of the holomorphic value $F(1/4)$. □

The contradiction is an instance of a general principle: a Dirichlet series with nonnegative coefficients cannot acquire too much holomorphic continuation across its real convergence boundary. Positivity makes the Taylor expansion sum the original coefficients at a point where they should diverge. Here the square coefficients supply an explicit divergent subseries, so no abstract theorem about an abscissa of convergence is needed.

4.4 The trigonometric polynomial inequality

For real v,

$$3 + 4\cos v + \cos 2v = 2(1 + \cos v)^2 \geq 0.$$

This elementary polynomial couples three Euler products. For $\sigma > 1$, a Dirichlet character χ modulo q, and real t, absolute convergence of the logarithmic Euler expansions gives

$$\begin{aligned} &\log \left(\zeta(\sigma)^3 |L(\sigma + it, \chi)|^4 |L(\sigma + 2it, \chi^2)| \right) \\ &= \sum_{p} \sum_{m \geq 1} \frac{p^{-m\sigma}}{m} \left(3 + 4 \operatorname{Re}(\chi(p)^m p^{-imt}) + \operatorname{Re}(\chi(p)^{2m} p^{-2imt}) \right) \geq 0. \end{aligned}$$

For $p \nmid q$, the expression in parentheses has the displayed trigonometric form because $\chi(p)^m p^{-imt}$ has modulus one. For $p \mid q$, the character terms vanish and the expression is 3. Exponentiating yields

$$\zeta(\sigma)^3 |L(\sigma + it, \chi)|^4 |L(\sigma + 2it, \chi^2)| \geq 1. \tag{7}$$

Here χ^2 is taken at modulus q; when it is principal, its L-function has the finite Euler factors belonging to that modulus.

Theorem 4.2 (Nonvanishing on the line of absolute-convergence boundary)**.** *For every Dirichlet character χ, $L(s, \chi)$ is nonzero on $\operatorname{Re} s = 1$, apart from its principal pole.*

Proof. Suppose first that $t \neq 0$ and that $L(1 + it, \chi) = 0$ of order $m \geq 1$. The function $L(s, \chi)$ is holomorphic near $1 + it$ even when χ is principal, since its only right-half-plane pole is at 1. Thus $|L(\sigma + it, \chi)| = O((\sigma - 1)^m)$ as $\sigma \downarrow 1$. The factor $L(\sigma + 2it, \chi^2)$ remains bounded, while $\zeta(\sigma) = O((\sigma - 1)^{-1})$. The left side of (7) is therefore $O((\sigma - 1)^{4m-3}) \to 0$, a contradiction.

Now let $t = 0$ and let χ be nonprincipal. If χ is real, Proposition 4.1 applies. Otherwise χ^2 is nonprincipal: a character whose square is principal has only the values 1 and -1 on units. Hence $L(s, \chi^2)$ is holomorphic at 1, and the same order-of-vanishing argument in (7) rules out a zero of $L(s, \chi)$ there. The remaining case is the principal pole, which is excluded from the statement. □

Together with Euler-product nonvanishing in $\operatorname{Re} s > 1$, this theorem supplies the right boundary input used by the pole criterion. The proof also explains why real characters at the real point 1 require a separate argument: in that case χ^2 is principal, and the additional simple pole would make the crude exponent count inconclusive.

4.5 What remains of the classical analytic input

The preceding arguments establish holomorphic continuation of nonprincipal L-functions to $\operatorname{Re} s > 0$, the zeta pole and its constant term, and nonvanishing on $\operatorname{Re} s \geq 1$. These facts suffice for the construction of the corrected transforms and their pole analysis in the positive half-plane.

The functional equations and the continuation to the full plane enter only in the final reflection of zeros and the identification of the trivial zeta zeros. They will be derived from theta transformations in a later section. The one-sided window-to-right-half-plane theorem itself is independent of that additional reflection theory.

5 Fourier analysis underlying the theta transformation

The functional equation is obtained by comparing a Gaussian sum on a lattice with the corresponding sum on its dual. We prove the analytic identity used in that comparison, including its convergence statements. This also fixes the Fourier-transform convention and the powers of i that enter the root number.

5.1 Schwartz functions

A smooth function $f : \mathbb{R} \to \mathbb{C}$ is a Schwartz function if

$$\sup_{x \in \mathbb{R}} |x^m f^{(n)}(x)| < \infty \qquad (m, n \geq 0).$$

Every derivative then decays faster than every fixed negative power. In particular f and all its derivatives are integrable. We use

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(x) e^{-2\pi i x\xi}\, dx.$$

Differentiation under the integral gives

$$\widehat{f}^{(n)}(\xi) = \int_{\mathbb{R}} (-2\pi i x)^n f(x) e^{-2\pi i x\xi}\, dx,$$

because $x^n f(x)$ is integrable. Repeated integration by parts gives

$$(2\pi i\xi)^m \widehat{f}(\xi) = \widehat{f^{(m)}}(\xi).$$

Boundary terms vanish by rapid decay. Applying the same identity to $x^n f(x)$ proves that $\widehat{f}$ is again a Schwartz function.

These estimates justify all lattice sums below absolutely. They also allow termwise differentiation of periodized series uniformly on compact sets, since a compact shift of an integer does not change the decay order of its absolute value.

5.2 The Gaussian integral and its transform

The integral $I = \int_{\mathbb{R}} e^{-\pi x^2} dx$ satisfies $I = 1$. Indeed, Tonelli and polar coordinates give

$$I^2 = \int_{\mathbb{R}^2} e^{-\pi(x^2+y^2)} dx\, dy = 2\pi \int_0^\infty r e^{-\pi r^2} dr = 1,$$

and $I > 0$. For $a > 0$, scaling gives $\int e^{-\pi a x^2} dx = a^{-1/2}$.

Let $F(\xi) = \widehat{e^{-\pi a x^2}}(\xi)$. Differentiating and using $x e^{-\pi a x^2} = -(2\pi a)^{-1}(d/dx) e^{-\pi a x^2}$, integration by parts gives

$$F'(\xi) = -\frac{2\pi\xi}{a} F(\xi).$$

With $F(0) = a^{-1/2}$, the unique solution is

$$\widehat{e^{-\pi a x^2}}(\xi) = a^{-1/2} e^{-\pi\xi^2/a}. \tag{8}$$

Since $\widehat{xf}(\xi) = -(2\pi i)^{-1} \widehat{f}'(\xi)$, we also have

$$\widehat{x e^{-\pi a x^2}}(\xi) = -i\xi a^{-3/2} e^{-\pi\xi^2/a}. \tag{9}$$

The sign in the odd formula depends on the sign chosen in the Fourier exponential. It is the source of the factor i^{-1} in the odd Dirichlet root number.

5.3 Fourier inversion by Gaussian regularization

For a Schwartz function f, define

$$I_\varepsilon(x) = \int_{\mathbb{R}} \widehat{f}(\xi) e^{2\pi i x\xi} e^{-\pi\varepsilon\xi^2}\, d\xi, \qquad \varepsilon > 0.$$

The integrand is absolutely integrable. Substitute the definition of $\widehat{f}$ and use Fubini; the double absolute integral is bounded by $\|f\|_1 \int e^{-\pi\varepsilon\xi^2} d\xi < \infty$. Formula (8) gives

$$I_\varepsilon(x) = \int_{\mathbb{R}} f(y)\varepsilon^{-1/2} e^{-\pi(x-y)^2/\varepsilon} dy.$$

The Gaussian kernel has integral one and concentrates at zero. To prove $I_\varepsilon(x) \to f(x)$, split the integral into $|x-y| < \delta$ and its complement. Uniform continuity controls the first difference by the modulus of continuity of f, while boundedness of f and the Gaussian tail control the complement. First let $\varepsilon \to 0$, then $\delta \to 0$.

On the Fourier side, dominated convergence applies because $\widehat{f} \in L^1$. We conclude

$$f(x) = \int_{\mathbb{R}} \widehat{f}(\xi) e^{2\pi i x\xi}\, d\xi. \tag{10}$$

In particular $\widehat{\widehat{f}}(x) = f(-x)$. No distributional interpretation is needed for this class of functions.

5.4 Periodization

For a Schwartz function, define $P_f(x) = \sum_{n\in\mathbb{Z}} f(x+n)$. On $[0,1]$, the series and every termwise derivative converge uniformly by rapid decay. Thus P_f is a smooth function of period one. Its kth Fourier coefficient is

$$\int_0^1 P_f(x) e^{-2\pi i k x} dx = \sum_n \int_n^{n+1} f(y) e^{-2\pi i k y} dy = \widehat{f}(k).$$

The interchange is justified by absolute integrability over the line. The coefficients $\widehat{f}(k)$ decay rapidly, so their Fourier series converges absolutely and uniformly.

For completeness, a continuous periodic function is determined by its Fourier coefficients. Let

$$F_N(x) = \frac{1}{N} \left| \sum_{j=0}^{N-1} e^{2\pi i j x} \right|^2.$$

This Fejer kernel is nonnegative, has integral one, and is a finite trigonometric polynomial. Away from an integer, $F_N(x) \le 1/[N \sin^2(\pi x)]$. Hence its mass outside any fixed neighborhood of the integers tends to zero. Convolution with F_N therefore converges uniformly to every continuous periodic function, by the same near-and-far argument used for the Gaussian. If all Fourier coefficients vanish, every such convolution vanishes, so the function is zero.

Applying uniqueness to P_f and its absolutely convergent Fourier series gives

$$P_f(x) = \sum_{k\in\mathbb{Z}} \widehat{f}(k) e^{2\pi i k x}.$$

At $x = 0$ this is the Poisson summation formula

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{k\in\mathbb{Z}} \widehat{f}(k). \tag{11}$$

5.5 Scaling, translation, and modulation

Let $q > 0$ and $a \in \mathbb{R}$. Applying Poisson summation to $x \mapsto f(a + qx)$ gives

$$\sum_{n\in\mathbb{Z}} f(a+qn) = \frac{1}{q}\sum_{k\in\mathbb{Z}} e^{2\pi ika/q}\widehat{f}(k/q). \tag{12}$$

The factor q^{-1} comes from the substitution $y = a+qx$ in the Fourier integral. Similarly, multiplication by $e^{2\pi ibx}$ shifts the Fourier transform: $\widehat{e^{2\pi ibx}f(x)}(\xi) = \widehat{f}(\xi - b)$. Translation by a multiplies it by $e^{2\pi ia\xi}$ when the translated function is $f(x + a)$.

These formulas explain why sums over residue classes interact with finite Fourier transforms. The analytic transform handles the infinite lattice, while the finite sum over $a \pmod q$ handles the character. Keeping the two transformations separate is a convenient way to verify all normalization factors.

5.6 Periodic weights

Let $c : \mathbb{Z}/q\mathbb{Z} \to \mathbb{C}$ be any function, extended periodically to integers. Define its finite Fourier transform by

$$C(k) = \sum_{a \bmod q} c(a)e^{2\pi ika/q}.$$

Multiply (12) by $c(a)$ and sum over residue classes. Absolute convergence gives

$$\sum_{n\in\mathbb{Z}} c(n)f(n) = \frac{1}{q}\sum_{k\in\mathbb{Z}} C(k)\widehat{f}(k/q). \tag{13}$$

No multiplicativity assumption on c is needed. For a primitive Dirichlet character, the finite Gauss-sum identity simplifies $C(k)$ to $\overline{\chi}(k)\tau(\chi)$, including zero when k is not a unit modulo q. The proof of that finite identity is given in the functional-equation section.

If c is imprimitive, the nonunit values of $C(k)$ need not vanish in the same way. One can still use (13), but the resulting transformation involves the actual finite transform rather than a single primitive character. This is another reason to separate primitive induction before applying the standard theta functional equation.

5.7 Even and odd theta sums

Let χ be nonprincipal and primitive, and choose $a \in \{0, 1\}$ with $\chi(-1) = (-1)^a$. Set

$$\theta_\chi(t) = \sum_{n\in\mathbb{Z}} n^a\chi(n)e^{-\pi n^2t/q}, \qquad t > 0.$$

For $a = 0$, the convention at $n = 0$ is supplied by $\chi(0)$; for a nonprincipal primitive character this value is zero. For $a = 1$, the factor n makes the zero term vanish regardless. The parity choice ensures that the summands at n and $-n$ agree. Thus the sum is twice the sum over positive integers.

Apply (13) to $f(x) = x^ae^{-\pi tx^2/q}$. The Gaussian formulas give

$$\widehat{f}(k/q) = i^{-a}k^aq^{1/2}t^{-a-1/2}e^{-\pi k^2/(qt)}.$$

After multiplication by $\tau(\chi)/q$, one obtains

$$\theta_\chi(t) = \frac{\tau(\chi)}{i^a\sqrt{q}} t^{-a-1/2}\theta_{\overline{\chi}}(1/t). \tag{14}$$

This calculation exhibits the parity, square root of the conductor, and power of t directly. They are fixed by the transform convention and do not depend on a later choice of completion.

5.8 Decay and uniform differentiation

For $t \geq 1$ and fixed q, the series defining $\theta_\chi(t)$ and all its derivatives converge absolutely and uniformly on $[1, \infty)$. Each derivative introduces a power of n^2, which is dominated by Gaussian decay. More explicitly, for every $r \geq 0$ there are constants $C_{q,r}, c_q > 0$ such that

$$|\theta_\chi^{(r)}(t)| \leq C_{q,r} e^{-c_q t}$$

when the zero term is absent. To prove this, split $e^{-\pi n^2 t/q}$ into two factors, bound one by $e^{-\pi t/(2q)}$ for $n \neq 0$, and sum the other against the fixed polynomial in n at $t = 1$.

The transformation formula then gives exponential decay in $1/t$ as $t \downarrow 0$, up to the explicit power $t^{-a-1/2}$. These estimates justify the Mellin integral for every complex s after splitting at $t = 1$. On a compact set of s, powers of t and $\log t$ arising from differentiation are absorbed by the same exponential majorants.

5.9 Why the Mellin exponent is forced

For $\operatorname{Re} s > 1$, integrate the positive-index theta series term by term:

$$\int_0^\infty \theta_\chi(t) t^{(s+a)/2-1} dt = 2 \sum_{n \geq 1} n^a \chi(n) \int_0^\infty e^{-\pi n^2 t/q} t^{(s+a)/2-1} dt.$$

Substituting $u = \pi n^2 t/q$ gives a factor $(q/\pi)^{(s+a)/2} n^{-s-a} \Gamma((s+a)/2)$. The factor n^a in the theta sum cancels the extra n^{-a}, leaving $L(s, \chi)$. Thus the natural completion is

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

The integral equals $2\Lambda(s, \chi)$.

Under $t \mapsto 1/t$, the exponent and the power in (14) combine to replace s by $1 - s$. The central line is therefore $\operatorname{Re} s = 1/2$. This is a structural consequence of the theta transformation, rather than a normalization imposed afterward on the L-function.

6 Gamma factors, Gauss sums, and functional equations

6.1 The gamma factor has no zeros

For $\operatorname{Re} z > 0$, define

$$\Gamma(z) = \int_0^\infty u^{z-1} e^{-u}\, du.$$

Exponential domination at infinity and the condition $\operatorname{Re} z > 0$ at zero give absolute convergence and holomorphy. Integration by parts gives $\Gamma(z+1) = z\Gamma(z)$, with $\Gamma(1) = 1$. Nonvanishing at a complex argument does not follow from positivity of the real integrand at real arguments, so it deserves a separate proof.

Repeated integration by parts in the beta integral gives

$$\int_0^1 t^{z-1}(1-t)^n\,dt = \frac{n!}{z(z+1)\cdots(z+n)}.$$

After multiplying by n^z and substituting $u = nt$, the left side converges to $\Gamma(z)$. Dominated convergence applies because $(1-u/n)^n \le e^{-u}$ for $0 \le u < n$; on compact subsets of $\operatorname{Re} z > 0$, a common integrable bound handles both endpoints. Therefore

$$\Gamma(z) = \lim_{n\to\infty} \frac{n!n^z}{z(z+1)\cdots(z+n)}. \tag{15}$$

The reciprocals of the finite expressions can be written

$$z\exp\bigl(z(H_n - \log n)\bigr)\prod_{m=1}^{n}(1+z/m)e^{-z/m}.$$

The product converges locally uniformly, since its logarithmic tail is a normally convergent series of terms $O_K(m^{-2})$ on every compact set K. Thus

$$D(z) = ze^{\gamma z}\prod_{m\ge 1}(1+z/m)e^{-z/m} \tag{16}$$

defines an entire function. Its zeros are precisely $0, -1, -2, \ldots$, and all are simple: away from the finitely many possibly vanishing factors, the logarithmic tail has a finite sum and hence a nonzero exponential.

Taking limits in the product of the finite expressions in (15) with their reciprocals gives $\Gamma(z)D(z) = 1$ for $\operatorname{Re} z > 0$. This proves nonvanishing there and identifies $D = 1/\Gamma$. The reciprocal product extends gamma meromorphically to the plane, with poles at the nonpositive integers and no zeros. In particular, gamma factors cannot create zeros in the open critical strip.

6.2 The finite Fourier transform of a primitive character

Let χ be primitive modulo $q > 1$, and define

$$\tau(\chi) = \sum_{r \bmod q}\chi(r)e^{2\pi ir/q}, \qquad S_\chi(m) = \sum_{r \bmod q}\chi(r)e^{2\pi imr/q}.$$

If m is a unit modulo q, changing variables by multiplication with m gives $S_\chi(m) = \overline{\chi(m)}\tau(\chi)$. Primitivity makes the same formula hold when m is not a unit, with both sides zero.

To prove this last assertion, set $d = q/\gcd(m,q)$, a proper divisor of q. Reduction $(\mathbb{Z}/q\mathbb{Z})^\times \to (\mathbb{Z}/d\mathbb{Z})^\times$ is surjective: a unit modulo d can be lifted to avoid the finitely many prime divisors of q not dividing d, using the Chinese remainder theorem. If χ were trivial on the kernel of this reduction, it would factor through modulus d, contrary to primitivity. Hence there is a unit $b \equiv 1 \pmod d$ with $\chi(b) \ne 1$. Since $mb \equiv m \pmod q$, multiplication of the summation variable by b gives $S_\chi(m) = \chi(b)S_\chi(m)$, so $S_\chi(m) = 0$. We have proved

$$S_\chi(m) = \overline{\chi(m)}\tau(\chi) \qquad (m \in \mathbb{Z}). \tag{17}$$

Finite Fourier orthogonality gives

$$\sum_{m \bmod q} |S_\chi(m)|^2 = q \sum_{r \bmod q} |\chi(r)|^2 = q\varphi(q).$$

Using (17), the left side is $|\tau(\chi)|^2\varphi(q)$. Thus $|\tau(\chi)| = \sqrt{q}$. Also, conjugating the Gauss sum and replacing r by $-r$ shows

$$\overline{\tau(\chi)} = \chi(-1)\tau(\bar{\chi}), \qquad \tau(\chi)\tau(\bar{\chi}) = \chi(-1)q.$$

These identities will determine the nonzero root number.

6.3 Poisson summation and character theta series

We use Poisson summation for Schwartz functions on the real line, with Fourier transform $\widehat{f}(\xi) = \int_{\mathbb{R}} f(x)e^{-2\pi i x\xi}\,dx$. In residue-class form it states

$$\sum_{k\in\mathbb{Z}} f(r+kq) = \frac{1}{q}\sum_{m\in\mathbb{Z}} \widehat{f}(m/q)e^{2\pi i m r/q}.$$

Both sides converge absolutely for the Gaussian functions used below.

Let $a \in \{0,1\}$ satisfy $\chi(-1) = (-1)^a$, and set

$$\vartheta_\chi(t) = \sum_{n\in\mathbb{Z}} n^a \chi(n) e^{-\pi n^2 t/q}, \qquad t > 0.$$

The term at zero is zero because $q > 1$ and $\chi(0) = 0$. For $f(x) = e^{-\pi t x^2/q}$, $\widehat{f}(\xi) = \sqrt{q/t}\, e^{-\pi q\xi^2/t}$. Differentiating this transform gives

$$\widehat{xf}(\xi) = -i(q/t)\xi\sqrt{q/t}\, e^{-\pi q\xi^2/t}.$$

Apply residue-class Poisson summation and then (17). For both parities the result is

$$\vartheta_\chi(t) = \epsilon_\chi t^{-a-1/2}\vartheta_{\bar{\chi}}(1/t), \qquad \epsilon_\chi = \frac{\tau(\chi)}{i^a\sqrt{q}}. \tag{18}$$

The modulus of ϵ_χ is one. Its exact phase is irrelevant to nonvanishing, but its nonzero value is essential for reflecting zeros.

6.4 Mellin integration of the theta identity

The theta series decays exponentially as $t \to \infty$. Equation (18) shows that as $t \downarrow 0$ it is a power of t times a function decaying like $e^{-c/t}$. Consequently the integral

$$I_\chi(s) = \int_0^\infty \vartheta_\chi(t) t^{(s+a)/2-1}\,dt$$

converges for every complex s and defines an entire function. Derivatives in s merely insert powers of $\log t$, which are absorbed by the rapid decay at both endpoints.

For $\operatorname{Re} s > 1$, absolute convergence permits termwise integration. The contributions from n and $-n$ are equal, because $(-n)^a\chi(-n) = n^a\chi(n)$. Hence

$$\begin{aligned} I_\chi(s) &= 2\sum_{n\geq 1} n^a\chi(n)\Gamma((s+a)/2)(\pi n^2/q)^{-(s+a)/2} \\ &= 2(q/\pi)^{(s+a)/2}\Gamma((s+a)/2)L(s,\chi). \end{aligned}$$

Thus the completed function

$$\Lambda(s,\chi) = (q/\pi)^{(s+a)/2}\Gamma((s+a)/2)L(s,\chi)$$

extends to the entire function $I_\chi(s)/2$. Since $1/\Gamma$ is entire, this also extends $L(s,\chi)$ itself to an entire function.

Finally substitute (18) into the integral and change variables $u = 1/t$. The exponent transforms exactly as

$$t^{(s+a)/2-1}t^{-a-1/2}\,dt \longmapsto u^{(1-s+a)/2-1}\,du$$

with the reversed endpoints restoring the positive orientation. It follows that

$$\Lambda(s,\chi) = \epsilon_\chi\Lambda(1-s,\bar{\chi}). \tag{19}$$

This derivation shows where parity, conductor normalization, and the conjugate character enter. Dropping any of those features would change the completed function and its reflection law.

6.5 The zeta theta series and the principal exception

For the principal character of conductor 1, use

$$\vartheta(t) = \sum_{n\in\mathbb{Z}} e^{-\pi n^2 t}.$$

Poisson summation gives $\vartheta(t) = t^{-1/2}\vartheta(1/t)$. The constant term 1 changes the Mellin calculation. For $\operatorname{Re} s > 1$,

$$\Lambda_\zeta(s) := \pi^{-s/2}\Gamma(s/2)\zeta(s) = \frac{1}{2}\int_0^\infty (\vartheta(t)-1)t^{s/2-1}\,dt.$$

Split the integral at 1 and apply the theta transformation to the lower half. Separating the elementary term $t^{-1/2} - 1$ yields

$$\Lambda_\zeta(s) = \frac{1}{s(s-1)} + \frac{1}{2}\int_1^\infty (\vartheta(t)-1)\left(t^{s/2-1} + t^{(1-s)/2-1}\right)dt. \tag{20}$$

The integral on the right is entire by exponential decay. The expression is invariant under $s \mapsto 1-s$, so $\Lambda_\zeta(s) = \Lambda_\zeta(1-s)$ meromorphically. It has simple poles at 0 and 1 with residues -1 and 1. The function $\frac{1}{2}s(s-1)\Lambda_\zeta(s)$ is entire and symmetric about the critical line.

The principal exception is therefore a consequence of the constant term in the theta series. For primitive nonprincipal characters that constant term is absent, and the corresponding completed function is entire. Finite Euler factors then treat all imprimitive characters.

6.6 The domain in which completion preserves zeros

If $0 < \operatorname{Re} s < 1$, then $(s+a)/2$ has positive real part for either parity. The gamma factor is finite and nonzero there, and the conductor factor is an exponential and hence nonzero. Therefore $L(s,\chi)$ and its completion have precisely the same zeros in that strip, including multiplicities.

Outside the strip, gamma poles can cancel zeros of the uncompleted function. One must therefore examine the reciprocal gamma factor rather than asserting that completion preserves zeros everywhere. For the zeta function this examination produces the trivial negative even zeros, while the symmetry of the completed function controls the nontrivial zeros. The final section carries out that distinction explicitly.

7 Growth, zero counting, and logarithmic derivatives

The theta integral gives more than a functional equation. It supplies quantitative growth estimates for the completed functions and thereby controls the number of their zeros. These estimates clarify the locally finite pole sets used in the window criterion. They also distinguish local residue calculations, which need very little global information, from infinite explicit formulas, which require convergence control.

7.1 Entire completions

For a nonprincipal primitive character χ of parity a, write

$$\Lambda_\chi(s) = \left(\frac{q}{\pi}\right)^{(s+a)/2} \Gamma\left(\frac{s+a}{2}\right) L(s,\chi).$$

Splitting its theta integral at one and using the theta transformation gives

$$\Lambda_\chi(s) = \frac{1}{2}\int_1^\infty \left[\theta_\chi(t) t^{(s+a)/2-1} + \epsilon_\chi \theta_{\overline{\chi}}(t) t^{(1-s+a)/2-1}\right] dt. \tag{21}$$

Both integrals and all their derivatives in s converge uniformly on compact subsets of $\mathbb{C}$, by exponential decay of the theta sums. Thus Λ_χ is entire.

For the zeta function, define

$$\xi(s) = \tfrac{1}{2} s(s-1)\pi^{-s/2}\Gamma(s/2)\zeta(s).$$

The rational terms in the split theta integral are canceled by $s(s-1)$, making ξ entire. The functional equations are $\Lambda_\chi(s) = \epsilon_\chi \Lambda_{\overline{\chi}}(1-s)$ and $\xi(s) = \xi(1-s)$.

The values at zero are nonzero. For nonprincipal χ, the functional equation gives $\Lambda_\chi(0) = \epsilon_\chi \Lambda_{\overline{\chi}}(1)$, and the latter is nonzero by boundary nonvanishing and the nonvanishing of the gamma function at its positive argument. For zeta, the split theta formula gives $\xi(0) = 1/2$. These facts let us apply Jensen's formula about the origin without first dividing out a zero there.

7.2 An elementary factorial bound for the theta integral

Let $c > 0$ and $u \geq 0$. If $m = \lceil u \rceil$, then for $t \geq 1$ one has $t^u \leq t^m$. Hence

$$\int_1^\infty e^{-ct} t^u \, dt \leq \int_0^\infty e^{-ct} t^m \, dt = c^{-m-1} m!.$$

The last identity follows by m integrations by parts, starting from $\int_0^\infty e^{-ct}dt = 1/c$. Since $m! \le m^m$ for $m \ge 1$, the logarithm of the right side is $O_c((u+1)\log(u+2))$.

On $|s| \le R$, the powers of t in (21) are bounded by $t^{(R+2)/2}$ for $t \ge 1$. The preceding estimate therefore gives constants depending on the fixed character such that

$$\max_{|s|\le R} |\Lambda_\chi(s)| \le \exp\bigl(C_\chi(R+1)\log(R+2)\bigr). \tag{22}$$

The same estimate holds for ξ, with an absolute constant, because the additional polynomial factor contributes only $O(\log(R+2))$ to the logarithm of the maximum.

No uniformity in the conductor is claimed here. The exponential decay constant in the theta series depends on q. A statement uniform in a varying conductor requires keeping that dependence rather than absorbing it into C_χ.

7.3 The order of the completed functions

For an entire function F, its order is

$$\rho_F = \limsup_{R\to\infty} \frac{\log\log \max_{|s|\le R} |F(s)|}{\log R},$$

with harmless modifications when the maximum is small. Estimate (22) gives order at most one. The order is exactly one for the completed functions above.

Indeed, for real $\sigma \to \infty$, the absolutely convergent Dirichlet series gives $L(\sigma,\chi) \to 1$, and likewise $\zeta(\sigma) \to 1$. For real $r \ge 1$, Euler's gamma integral has the lower bound

$$\Gamma(r) \ge \int_r^{r+1} t^{r-1}e^{-t}dt \ge r^{r-1}e^{-r-1}.$$

Taking $r = (\sigma + a)/2$ shows

$$\log|\Lambda_\chi(\sigma)| \ge \tfrac{1}{2}\sigma\log\sigma - O_\chi(\sigma).$$

For ξ the extra polynomial factor does not alter this conclusion. Thus the limsup defining the order is at least one. The ratio $\log\max|F|/R$ is unbounded, so the completed functions have infinite type at order one. The coarser order statement is all that is needed for the zero-counting estimates below.

7.4 Jensen's formula

Let F be holomorphic on a neighborhood of the closed disk of radius R, with $F(0) \ne 0$ and no zero on the boundary. List its zeros inside the disk as $\rho_1, \dots, \rho_m$, repeated by multiplicity. Factoring these zeros leaves a nonvanishing holomorphic function on the disk. Its logarithmic absolute value is harmonic, so the mean-value property gives

$$\sum_{j=1}^{m} \log\frac{R}{|\rho_j|} = \frac{1}{2\pi}\int_0^{2\pi} \log|F(Re^{i\theta})|d\theta - \log|F(0)|. \tag{23}$$

For the contribution of one zero, use

$$\frac{1}{2\pi}\int_0^{2\pi} \log|Re^{i\theta} - \rho|d\theta = \log R \qquad (|\rho| < R).$$

This follows by expanding $\log(1-\rho e^{-i\theta}/R)$ into its uniformly convergent power series on a slightly smaller ratio and taking real parts; every nonconstant Fourier mode has mean zero.

If the original radius meets a zero, apply the formula at nearby larger radii avoiding zeros and pass to the required inequality by a limit. Isolated zeros ensure that such radii are available. This is enough for counting estimates even without a boundary-free exact formula at every radius.

7.5 A zero-counting bound

Let $n_F(R)$ count the zeros of F in $|s| \le R$, with multiplicity. Apply Jensen at radius $2R$ or a slightly larger radius. Every zero in the smaller disk contributes at least $\log 2$. Thus

$$n_F(R)\log 2 \le \log \max_{|s|\le 2R+1} |F(s)| - \log|F(0)|.$$

For $F = \Lambda_\chi$ or ξ, this gives

$$n_F(R) = O_\chi((R+1)\log(R+2)). \tag{24}$$

This is an upper bound, not the full zero-counting asymptotic. It requires neither an argument-principle analysis on a tall rectangle nor a precise Stirling expansion.

In particular, the zeros are locally finite, and the number with $|\operatorname{Im}\rho| \le T$ in the critical strip is $O_\chi(T\log(T+2))$. The strip lies inside the disk of radius $T+1$, so the latter follows directly from (24).

7.6 Trivial zeros and the completed zero set

The reciprocal gamma function has simple zeros at the nonpositive integers. For a primitive nonprincipal character of parity a, the identity

$$L(s,\chi) = \left(\frac{q}{\pi}\right)^{-(s+a)/2} \frac{\Lambda_\chi(s)}{\Gamma((s+a)/2)}$$

therefore gives zeros at

$$s = -a - 2m, \qquad m = 0, 1, 2, \dots.$$

These zeros are simple. At such an s, the functional equation expresses $\Lambda_\chi(s)$ through $\Lambda_{\overline{\chi}}(1-s)$, which is nonzero because $1 - s \ge 1$ and the boundary and Euler-product nonvanishing results apply. Thus only the simple reciprocal-gamma zero contributes.

At every other point with $\operatorname{Re} s \le 0$, the gamma factor is finite and nonzero and the reflected completed value is nonzero. Hence there are no other zeros in that closed half-plane. The completed function has no zeros for $\operatorname{Re} s \ge 1$ either, and its functional equation excludes $\operatorname{Re} s \le 0$. All of its zeros lie in the open critical strip and correspond to the nontrivial zeros of $L(s,\chi)$. The zeta completion has the same property; its separate polynomial factors cancel the poles at zero and one rather than adding zeros there.

7.7 Summability over the zeros

The coarse zero count implies, for every $\delta > 0$,

$$\sum_\rho \frac{1}{|\rho|^{1+\delta}} < \infty, \tag{25}$$

where zeros are counted with multiplicity. There are only finitely many zeros in a fixed disk, none at zero. For the rest, group them into dyadic annuli $2^j < |\rho| \le 2^{j+1}$. The contribution of the jth annulus is at most

$$2^{-j(1+\delta)} O_\chi(2^j(j+1)) = O_\chi((j+1)2^{-j\delta}),$$

and this series converges. In particular $\sum_\rho |\rho|^{-2}$ converges. The same calculation gives the tail estimate

$$\sum_{|\rho|>T} |\rho|^{-2} = O_\chi\left(\frac{\log(T+2)}{T+1}\right).$$

Such estimates justify grouped expressions in which each zero contributes a term of order $|\rho|^{-2}$. They do not justify an unregularized sum of $1/\rho$, which need not converge absolutely.

7.8 A convergent product with the prescribed zeros

Set $E_1(z) = (1-z)e^z$ and form

$$P(s) = \prod_\rho E_1(s/\rho).$$

For $|z| \le 1/2$,

$$\log E_1(z) = -\sum_{m\ge 2} \frac{z^m}{m}, \qquad |\log E_1(z)| \le C|z|^2.$$

On each compact set of s, all but finitely many factors fall in this range, and (25) with $\delta = 1$ gives uniform convergence of the logarithmic tail. Thus P is an entire function whose zeros are precisely the prescribed zeros with the prescribed multiplicities.

The quotient F/P is entire and nonvanishing. It therefore equals $e^{h(s)}$ for some entire function h: the logarithmic derivative $(F/P)'/(F/P)$ is entire, has a primitive on the simply connected plane, and exponentiating that primitive gives the quotient up to a nonzero constant. This elementary argument does not determine the degree of h. A stronger factorization theorem can do so using global growth, but no such conclusion is needed for the local pole argument.

7.9 A regularized logarithmic-derivative expansion

On compact sets avoiding the zeros, differentiation of the uniformly convergent logarithmic tail gives

$$\frac{F'(s)}{F(s)} = h'(s) + \sum_\rho \left(\frac{1}{s-\rho} + \frac{1}{\rho}\right). \tag{26}$$

The summand equals $s/[\rho(s-\rho)]$ and is therefore $O_K(|\rho|^{-2})$ on a fixed compact set K. This proves normal convergence of the series away from its poles.

Near a zero ρ of multiplicity m, the singular part is $m/(s-\rho)$. Neither the regularizing constants nor the entire term h' can change that residue. This is the only aspect of the global expansion needed to establish positivity of the multiplicity weights in the Mellin transform. One can alternatively obtain it directly from the local factorization $F(s) = (s-\rho)^m u(s)$ with $u(\rho) \ne 0$.

For nonprincipal primitive characters,

$$\frac{L'}{L}(s,\chi) = \frac{\Lambda_\chi'}{\Lambda_\chi}(s) - \frac{1}{2}\log\frac{q}{\pi} - \frac{1}{2}\frac{\Gamma'}{\Gamma}\left(\frac{s+a}{2}\right).$$

Thus the completed zero contribution and the gamma contribution are separated explicitly. In the open critical strip the gamma term is holomorphic, so it does not affect residues at nontrivial zeros.

7.10 An absolutely convergent inversion kernel

For $c > 1$, define

$$k_c(u) = e^{-cu}(e^u - 1)\mathbf{1}_{[0,\infty)}(u).$$

This is continuous and integrable, and a direct integral gives

$$\int_{\mathbb{R}} k_c(u)e^{-itu}du = \frac{1}{c-1+it} - \frac{1}{c+it} = \frac{1}{(c+it)(c-1+it)}.$$

The transform is integrable in t, because it decays quadratically. The Gaussian-regularization proof of Fourier inversion applies to this continuous integrable function as well: the convolution tends to k_c at every point, and dominated convergence applies on the transform side. Therefore, for every $y > 0$,

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{y^s}{s(s-1)}ds = \begin{cases} y-1, & y \geq 1, \\ 0, & 0 < y < 1. \end{cases} \tag{27}$$

At $y = 1$, both expressions give zero. The integral is absolutely convergent, so no principal-value convention is needed.

7.11 Inversion of the smoothed prime sum

For $c > 1$, the logarithmic-derivative series is absolutely convergent:

$$-\frac{L'}{L}(s,\chi) = \sum_{n\geq 1} \frac{\Lambda(n)\chi(n)}{n^s}.$$

Insert this series into the vertical integral with the ramp kernel. The sum of absolute integrals is finite, bounded by

$$x^c \sum_{n\geq 1} \frac{\Lambda(n)}{n^c} \int_{\mathbb{R}} \frac{dt}{|(c+it)(c-1+it)|}.$$

Fubini and (27) yield the exact identity

$$R_\chi(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} -\frac{L'}{L}(s,\chi)\frac{x^s}{s(s-1)}ds. \tag{28}$$

The smoothing has supplied a second denominator. An unsmoothed Perron integral usually has only one and requires more care with truncation and boundary values.

7.12 What a contour shift does and does not prove

Take a finite rectangle whose boundary avoids zeros and the kernel poles. The residue theorem applied to the integrand in (28) is an exact finite identity. A zero ρ of multiplicity m contributes

$$-\frac{mx^\rho}{\rho(\rho-1)}$$

when $\rho \neq 0, 1$. A pole of L contributes with the opposite sign. The singularities at zero and one must be computed separately because the smoothing kernel itself has poles there.

Passing to an infinite rectangle requires estimates on the remaining sides. The coarse zero count alone does not bound the logarithmic derivative uniformly near zeros. One must choose contours

separated from the zeros and prove bounds on those contours before discarding horizontal integrals. The window method avoids that particular global contour-shift requirement: it works with high derivatives at fixed points and uses local pole separation plus Cauchy estimates.

The coefficient $-m/[\rho(\rho-1)]$ nevertheless appears in both approaches. The finite residue calculation explains why. The Mellin transform and the inversion formula are two views of the same smoothing operator, and local singularities give the same coefficients regardless of which global argument is used afterward.

8 Smoothed character sums and their transforms

For $x \geq 1$, set

$$R_\chi(x) = \sum_{1 \leq n \leq x} \Lambda(n)\chi(n)\left(\frac{x}{n} - 1\right). \tag{29}$$

Define

$$E_1(x) = R_1(x) - x\log x + (1+\gamma)x, \tag{30}$$

$$E_\chi(x) = R_\chi(x) + x\frac{L'}{L}(1,\chi), \qquad \chi \neq \chi_0. \tag{31}$$

The first formula uses the character of modulus 1. For a finite admissible family $\mathcal{F} = (\chi_i, a_i)_{i\in I}$, with $a_i > 0$, put

$$G_{\mathcal{F}}(x) = \sum_{i\in I} a_i \operatorname{Re} E_{\chi_i}(x). \tag{32}$$

Proposition 8.1 (Mellin transforms)**.** *For* $\operatorname{Re} s > 1$ *and a nonprincipal character* χ*,*

$$\int_1^\infty E_\chi(x)x^{-s-1}\,dx = -\frac{(L'/L)(s,\chi) - s(L'/L)(1,\chi)}{s(s-1)}. \tag{33}$$

For the character of modulus 1 *the transform is*

$$\int_1^\infty E_1(x)x^{-s-1}\,dx = -\frac{(\zeta'/\zeta)(s)}{s(s-1)} - \frac{1}{(s-1)^2} + \frac{1+\gamma}{s-1}. \tag{34}$$

Proof. Termwise integration in the absolutely convergent half-plane gives

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

The identity $\sum_{n\geq 1}\Lambda(n)\chi(n)n^{-s} = -(L'/L)(s,\chi)$, together with $\int_1^\infty x^{-s}\,dx = (s-1)^{-1}$, gives (33). Integrating $x\log x$ gives $(s-1)^{-2}$, proving (34). Absolute convergence justifying these exchanges is established below. □

The apparent singularities of the corrected transforms at $s = 1$ are removable. For a nonprincipal character this follows because the numerator in (33) vanishes there. For ζ, extend $(s-1)\zeta(s)$ holomorphically to $s = 1$, where its value is 1 and its derivative is γ.

Let Φ_i denote the meromorphic continuation of the transform of E_{χ_i}. Set

$$\Phi_{\mathcal{F}}(s) = \frac{1}{2}\sum_i a_i\left(\Phi_i(s) + \overline{\Phi_i(\bar{s})}\right). \tag{35}$$

For $\operatorname{Re} s > 2$, this is the Laplace transform of $g_{\mathcal{F}}(u) = G_{\mathcal{F}}(e^u)$ on $u > 0$. The symmetrization in (35) is needed because (32) takes real parts.

Proposition 8.2 (No cancellation of zero poles)**.** *Let ρ be a zero with $\operatorname{Re}\rho > 0$ of a member of $\mathcal{F}$ or of its conjugate-character L-function. Then $\Phi_{\mathcal{F}}$ has a simple pole at ρ with residue*

$$c_\rho = -\frac{M_\rho}{\rho(\rho-1)}, \qquad M_\rho = \frac{1}{2}\sum_i a_i\big(m_i(\rho) + m_i(\bar\rho)\big) > 0, \tag{36}$$

where $m_i(z)$ is the zero multiplicity of $L(s,\chi_i)$ at z. For a real zero $0 < \beta < 1$, the residue c_β is positive real.

Proof. At a zero of multiplicity m, the logarithmic derivative has principal part $m/(s-\rho)$. Equations (33)–(34) therefore give residue $-m/(\rho(\rho-1))$. Symmetrization gives (36). At least one multiplicity is positive and all weights are strictly positive, so $M_\rho > 0$. Nonvanishing on $\operatorname{Re} s \ge 1$ places the zeros under discussion below that line. If $\rho = \beta$ is real, then $-1/(\beta(\beta-1)) > 0$. □

This positivity statement is specific to the chosen positive combination of logarithmic derivatives. It does not assert that an arbitrary sum of L-functions preserves zeros, nor that $G_{\mathcal{F}}$ is pointwise nonnegative.

8.1 Elementary size and regularity estimates

We supply the analytic bookkeeping behind the transform identities. These estimates depend only on elementary properties of $\Lambda(n)$ and Dirichlet characters, and do not use any zero-free region beyond what is explicitly invoked at $s = 1$.

For $x \ge 2$ the inequalities $|\chi(n)| \le 1$, $0 \le \Lambda(n) \le \log n$, and $0 \le x/n - 1 \le x/n$ give

$$|R_\chi(x)| \le x \sum_{2\le n\le x} \frac{\log n}{n} \ll x(1+\log x)^2. \tag{37}$$

For example, splitting the sum into intervals $2^j \le n < 2^{j+1}$ bounds its jth contribution by a constant times $j+1$, and there are $O(1+\log x)$ such intervals. No prime number theorem is used here. The correcting terms in (30)–(31) have at most this order of growth for each fixed character. Finiteness of the family consequently gives

$$|G_{\mathcal{F}}(x)| \le C_{\mathcal{F}} x(1+\log x)^2 \qquad (x \ge 1).$$

The constant includes the finitely many values $(L'/L)(1,\chi_i)$. It need not be uniform in the conductor, because the analytic argument fixes the family before taking its large-scale limit.

Under $x = e^u$, the preceding estimate becomes $|g_{\mathcal{F}}(u)| \le C_{\mathcal{F}} e^u (1+u)^2$. Since $(1+u)^2 e^{-u}$ is bounded on $[0,\infty)$, this implies (42). The deliberately generous exponential rate 2 is useful: it puts the initial Laplace transform and all its derivatives in a common, simple half-plane. The eventual zero-free exponent θ is independent of this preliminary growth rate.

On each interval between consecutive integers, $R_\chi(x)$ is affine in x. At an integer n, the newly introduced summand has coefficient $n/n - 1 = 0$. Thus the sum is continuous there as well. In particular it is measurable and locally integrable. Its derivative has jumps, but the Laplace-transform argument never assumes differentiability of R_χ. All high derivatives in that argument are derivatives of the transform, whose existence follows from exponential domination.

8.2 Absolute convergence and the exchange of sum and integral

Let $\sigma = \operatorname{Re} s > 1$. For every integer $n \geq 1$,

$$\int_n^\infty \left|\Lambda(n)\chi(n)\left(\frac{x}{n} - 1\right)x^{-s-1}\right| dx \leq \frac{\Lambda(n)n^{-\sigma}}{\sigma(\sigma-1)}.$$

The series on the right converges, since $\Lambda(n) \leq \log n$ and $\sum_{n\geq 2}(\log n)n^{-\sigma} < \infty$. This proves absolute integrability of the sum of the integrands. Fubini's theorem therefore applies to the complex-valued series, either directly or separately to its real and imaginary parts. This supplies a concrete justification for the first exchange in Proposition 8.1.

For completeness, the elementary integral is

$$\begin{aligned}\int_n^\infty \left(\frac{x}{n} - 1\right)x^{-s-1}\,dx &= \frac{1}{n}\frac{n^{1-s}}{s-1} - \frac{n^{-s}}{s} \\ &= n^{-s}\left(\frac{1}{s-1} - \frac{1}{s}\right) = \frac{n^{-s}}{s(s-1)}.\end{aligned}$$

The complex powers here use the real logarithm of $x > 0$. Their antiderivatives are valid in the indicated half-plane, and the boundary contributions at infinity vanish because $\sigma > 1$.

The same absolute convergence permits the familiar logarithmic-derivative identity. In $\operatorname{Re} s > 1$, the Euler product can be differentiated on compact subsets, giving

$$-\frac{L'}{L}(s,\chi) = \sum_p \sum_{r\geq 1} \chi(p)^r (\log p) p^{-rs} = \sum_{n\geq 1} \Lambda(n)\chi(n)n^{-s}.$$

The first equality is justified by locally uniform absolute convergence; the second groups prime powers. Thus smoothing introduces precisely the factor $1/[s(s-1)]$, with no additional character-dependent denominator. That common factor is what later makes the residue signs compatible.

8.3 Why the correcting terms have these coefficients

Write $H_\chi(s) = (L'/L)(s,\chi)$ for a nonprincipal character. The transform of the uncorrected ramp is $-H_\chi(s)/[s(s-1)]$. Near $s = 1$ it has an artificial simple pole caused by the integration kernel, even though $L(s,\chi)$ is regular and nonzero there. Adding $xH_\chi(1)$ changes the transform by $H_\chi(1)/(s-1)$, so that the new numerator is $-H_\chi(s) + sH_\chi(1)$. It vanishes at $s = 1$.

More precisely, put $t = s - 1$ and expand $H_\chi(1+t) = h_0 + h_1 t + O(t^2)$. Then

$$-\frac{H_\chi(1+t) - (1+t)h_0}{(1+t)t} = h_0 - h_1 + O(t).$$

The removable value of the corrected transform is therefore $H_\chi(1) - H_\chi'(1)$. The exact value is not used in the pole exclusion argument, but the cancellation is essential: an artificial pole at 1 would otherwise be indistinguishable from a genuine obstruction.

For the principal character of modulus 1, use

$$\zeta(1+t) = \frac{1}{t} + \gamma + O(t), \qquad \frac{\zeta'}{\zeta}(1+t) = -\frac{1}{t} + \gamma + O(t).$$

It follows that

$$-\frac{(\zeta'/\zeta)(1+t)}{(1+t)t} = \frac{1}{t^2} - \frac{1+\gamma}{t} + O(1).$$

The terms $-x\log x$ and $(1+\gamma)x$ cancel the double and simple poles, respectively. Both coefficients in (30) are forced by this local calculation. An error in the linear coefficient would leave a pole whose residue has no relation to zeta zeros.

8.4 Conjugation, multiplicities, and the real-valued transform

For a complex character, taking the real part of the smoothed sum is not the same operation as taking the real part of its transform at a complex point. For a real-variable function E, one has

$$\int_1^\infty \operatorname{Re} E(x)\, x^{-s-1}\, dx = \frac{1}{2}\left(\Phi(s) + \overline{\Phi(\bar{s})}\right),$$

because conjugation also changes the exponent in the second integral. This is the origin of (35). The expression $\overline{\Phi(s)}$ alone is generally antiholomorphic and would not serve as a meromorphic continuation in s.

If Φ has principal part $c/(s-\bar{\rho})$ at $\bar{\rho}$, then $\overline{\Phi(\bar{s})}$ has principal part $\bar{c}/(s-\rho)$ at ρ. Substituting $c = -m/[\bar{\rho}(\bar{\rho}-1)]$ yields $-m/[\rho(\rho-1)]$. The contributions from the two symmetrized terms therefore point in exactly the same complex direction. The positivity needed for noncancellation is positivity of their real multiplicity coefficients, not positivity of the complex residue itself.

The statement also handles a multiple zero. If $L(s,\chi) = (s-\rho)^m h(s)$ with $h(\rho) \neq 0$, then

$$\frac{L'}{L}(s,\chi) = \frac{m}{s-\rho} + \frac{h'}{h}(s).$$

Thus a zero of multiplicity m still produces a simple pole of the transform. Its multiplicity is carried in the residue. There are no higher-order poles to complicate the high-derivative asymptotics.

Finally, local finiteness of the pole set follows from isolation of zeros of a nonzero analytic function. The family is finite, so the union of its zero sets and conjugate zero sets is locally finite away from the excluded boundary. This fact is used in compact disks contained in $\operatorname{Re} s > 0$. It does not assert any global finite bound on the number of zeros.

9 A one sided window criterion

Definition 9.1 (Multiplicative windows)**.** A set $S \subset \mathbb{R}$ has windows of exponent $\lambda > 1$ if

$$\forall Y_0 \in \mathbb{R}\ \exists Y \geq Y_0 \quad [Y, Y^\lambda] \subset S. \tag{38}$$

Only arbitrarily large positive Y are relevant.

For $0 < \theta < 1$ and $\delta > 0$, a slightly weaker integral formulation asks that for every $\eta > 0$ and every Y_0 there are $Y \geq Y_0$ and a majorant R_Y on $[Y, Y^{1+\delta}]$ satisfying

$$0 \leq R_Y(x), \qquad G_{\mathcal{F}}(x) \leq R_Y(x), \tag{39}$$

$$\int_Y^{Y^{1+\delta}} R_Y(x) x^{-\theta-1}\, dx \leq Y^\eta. \tag{40}$$

The integrand is required to be integrable.

Lemma 9.2. *If $G_{\mathcal{F}}(x) \leq Cx^\theta$ throughout arbitrarily large intervals $[Y, Y^{1+\delta}]$, then* (39)–(40) *hold.*

Proof. Take $R_Y(x) = |C|x^\theta$. Its weighted integral is $|C|\delta \log Y$, which is at most Y^η for sufficiently large Y. The arbitrarily large window hypothesis permits this extra lower threshold. □

Under $x = e^u$ and $U = \log Y$, the same condition becomes the existence of a nonnegative majorant $\mathcal{R}_U$ on $[U, (1+\delta)U]$ such that

$$g_{\mathcal{F}}(u) \le \mathcal{R}_U(u), \qquad \int_U^{(1+\delta)U} \mathcal{R}_U(u) e^{-\theta u}\, du \le e^{\eta U}. \tag{41}$$

The substitution preserves integrability on each compact interval.

Theorem 9.3 (Window nonvanishing criterion)**.** *Let $\mathcal{F}$ be a finite admissible family with strictly positive weights, and let $0 < \theta < 1$ and $\delta > 0$. If for some $C \in \mathbb{R}$,*

$$\forall Y_0\ \exists Y \ge Y_0\ \forall x \in [Y, Y^{1+\delta}], \qquad G_{\mathcal{F}}(x) \le C x^\theta,$$

then $L(s, \chi_i) \neq 0$ for every i and every $\operatorname{Re} s > \theta$, with poles understood in the usual meromorphic sense.

The proof passes through the integral majorant condition and the pole exclusion theorem developed next. The principal pole at $s = 1$ is excluded from every nonvanishing assertion.

9.1 What the window hypothesis does and does not control

A window set can be very sparse. For example, choose $Y_n \to \infty$ so rapidly that $Y_{n+1} > Y_n^{(1+\delta)^2}$, and put

$$S = \bigcup_{n \ge 1} [Y_n, Y_n^{1+\delta}].$$

Then S has the required windows, although the gaps between successive windows can be arbitrarily large on the logarithmic scale. None of the arguments below needs a density estimate for S. The transformation to $u = \log x$ turns these intervals into $[\log Y_n, (1+\delta)\log Y_n]$, whose lengths increase linearly with their left endpoints.

This linear growth in logarithmic length is the useful geometric feature. The high-derivative kernel will have its mass near a point proportional to $U = \log Y$. The two endpoints of $[U, (1+\delta)U]$ leave a positive proportional margin around that point. Those margins allow both tails to be exponentially small in the derivative order. A fixed logarithmic window length would not support the same tail argument with fixed parameters.

The lower threshold can always be increased. If a property holds on every window chosen above Y_0, the next application may replace Y_0 by $\max\{Y_0, Y_1, 1\}$. This elementary operation simultaneously handles a finite collection of eventual estimates. It is the reason that the many thresholds can be combined by repeated maxima.

Lemma 9.4 (Restriction and transfer of windows)**.** *Suppose S has windows of exponent $\lambda > 1$. If $1 < \lambda' \le \lambda$, then S has windows of exponent λ'. If T contains $S \cap [X_0, \infty)$, then T has windows of exponent λ.*

Proof. Choose the window's left endpoint above 1. Then $Y^{\lambda'} \le Y^\lambda$, proving the first assertion. For the second, choose $Y \ge \max\{Y_0, X_0, 1\}$. Every point in the selected interval lies in both S and $[X_0, \infty)$, hence in T. □

A comparison estimate valid only beyond x_1 can be transferred from a family scale set to the set where the analytic inequality holds. No closure of either set under multiplication is needed.

9.2 The integral formulation and its quantifiers

The integral window condition has the order

$$\forall \eta > 0 \quad \forall Y_0 \quad \exists Y \geq Y_0 \quad \exists R_Y \quad \text{the majorant conditions hold.}$$

The selected window is allowed to depend on η. The theorem does not require a single sequence of windows on which every positive η works from the outset. In the contradiction proof, η is selected after a pole and two radii have been fixed. Only that value of η is then used.

Similarly, R_Y may depend on the selected window. Its role is local: it controls the positive part of $G_{\mathcal{F}}$ on that interval in a weighted integral. The growth hypothesis supplies the bound outside the interval. No continuation of R_Y with a global estimate is required, although it can be extended by zero when convenient for measure-theoretic notation.

The inequality is genuinely one-sided. If a real number v is negative, then $|v| + v = 0$; if v is nonnegative, then $|v| + v = 2v$. Thus

$$|v| + v = 2\max\{v, 0\}.$$

The master derivative inequality uses exactly this positive part. Large negative values of g are not directly controlled by the majorant. Their effect appears in the real-center derivative term, which the pole argument retains and treats separately.

9.3 An explicit choice of the power-bound threshold

For the power majorant in Lemma 9.2, write $B = |C|\delta$ and $v = \log Y$. The required inequality is $Bv \leq e^{\eta v}$. For $v \geq 0$, the exponential series gives $e^{\eta v} \geq \eta^2 v^2/2$. Therefore it suffices to choose

$$v \geq \frac{2B}{\eta^2}, \qquad Y \geq \exp\left(\frac{2|C|\delta}{\eta^2}\right).$$

This threshold is not optimized. Its purpose is to make explicit that a fixed multiplicative constant in the pointwise bound does not change the zero-free exponent. The threshold may depend strongly on C, δ, and η, but this is permitted by the quantifiers.

The logarithmic substitution itself is exact:

$$x = e^u, \qquad dx = e^u\,du, \qquad x^{-\theta-1}dx = e^{-\theta u}\,du,$$
$$x = Y \Longleftrightarrow u = U, \qquad x = Y^{1+\delta} \Longleftrightarrow u = (1+\delta)U.$$

For $Y > 1$, the map is a continuously differentiable bijection of the two compact intervals, and the ordinary change-of-variables theorem transfers integrability between them. Endpoint choices such as closed or open intervals do not affect these Lebesgue integrals, since the endpoints have measure zero; pointwise majorization may still be stated on the closed interval.

9.4 A useful limitation of the criterion

The pole hypotheses are not optional. Take $g(u) = -e^{\beta u}$ with $\theta < \beta < 1$. Then the majorant $\mathcal{R} = 0$ works on every window, yet the Laplace transform is $-1/(s-\beta)$ and has a pole at β. Its residue is negative. This example is excluded precisely by the positive real-residue requirement in the pole exclusion theorem. It shows why a generic one-sided Tauberian assertion for arbitrary signed functions would be false. The arithmetic character combination supplies the additional sign structure that makes the theorem applicable.

10 The Laplace transform argument

We give the mechanism behind Theorem 9.3, including why the one-sided hypothesis suffices. A real measurable function g satisfies

$$|g(u)| \le C_0 e^{2u} \quad (u \ge 0). \tag{42}$$

Its Laplace transform Φ is initially defined on $\operatorname{Re} s > 2$ and extends holomorphically to $\{\operatorname{Re} s > 0\} \setminus Z$. The set Z is locally finite there. Each $\rho \in Z$ has a nonzero simple-pole principal part $c_\rho/(s-\rho)$, and $c_\rho > 0$ whenever ρ is real. These hypotheses hold for the character-family transform by Proposition 8.2 and the growth and regularity estimates above.

For $\sigma > 2$, $\tau \in \mathbb{R}$ and $k \ge 0$, write

$$A_k(\sigma + i\tau) = \frac{(-1)^k}{k!}\Phi^{(k)}(\sigma + i\tau) = \int_0^\infty g(u)\frac{u^k}{k!}e^{-(\sigma+i\tau)u}\,du. \tag{43}$$

Differentiation under the integral is justified by (42).

10.1 Controlling the derivatives from above

Put $w_k(u) = u^k e^{-\sigma u}/k!$. On an interval where $g \le \mathcal{R}$ and $\mathcal{R} \ge 0$, the elementary inequality

$$|g| + g \le 2\mathcal{R} \tag{44}$$

is the key. The triangle inequality in (43) and a split into the window and its complement give

$$\begin{aligned} |A_k(\sigma + i\tau)| + \operatorname{Re} A_k(\sigma) \le{}& \frac{2}{(\sigma-\theta)^k}\int_U^{(1+\delta)U} \mathcal{R}(u)e^{-\theta u}\,du \\ &+ 2C_0 \int_{(0,\infty)\setminus[U,(1+\delta)U]} \frac{u^k e^{(2-\sigma)u}}{k!}\,du. \end{aligned} \tag{45}$$

Indeed, $u^k e^{-(\sigma-\theta)u}/k! \le (\sigma-\theta)^{-k}$ follows from the exponential series. Outside the window, (42) bounds $|g| + g$ by $2C_0 e^{2u}$.

The tail estimate permits a choice of $\sigma > 2$, depending only on δ, so that for every $U \ge 1$ one can choose an integer $k \ge U$ with

$$\int_{(0,\infty)\setminus[U,(1+\delta)U]} \frac{u^k e^{(2-\sigma)u}}{k!}\,du \le C_1 \sigma^{-k}. \tag{46}$$

One explicit admissible choice is

$$\ell = 2 + \frac{6}{\delta}, \qquad M = \ell(1+\delta), \qquad \sigma = M(2+\ell), \qquad k = \lceil (1+\delta/2)\sigma U \rceil .$$

The proof applies exponential bounds to the upper and lower tails of the gamma kernel. This makes its concentration on the logarithmic window quantitative, without assuming bounds for g outside that window other than (42).

Combining (41), (45), and (46), one obtains a constant C_2 such that for every $\eta > 0$ and every lower threshold there exist arbitrarily large U and integers $k \ge U$ satisfying, simultaneously for every $\tau \in \mathbb{R}$,

$$|A_k(\sigma + i\tau)| + \operatorname{Re} A_k(\sigma) \le \frac{C_2 e^{\eta U}}{(\sigma-\theta)^k}. \tag{47}$$

The real-axis term on the left is retained. Dropping it without a lower bound would invalidate the argument. The two pole cases below address exactly this point.

10.2 Separating a pole

Assume Z contains ρ_0 with $\operatorname{Re}\rho_0 > \theta$. Since the initial Laplace transform is holomorphic on $\operatorname{Re} s > 2$, every pole has real part at most 2, hence less than the chosen σ. Local finiteness yields the following dichotomy.

(i) There is $D \le \sigma - \theta$ such that $B(\sigma, D)$ contains no pole, but some pole ρ_1 satisfies $\sigma - \operatorname{Re}\rho_1 < D$.

(ii) There is a real pole $\beta > \theta$ isolated in a disk $B(\sigma, r)$ with $\sigma - \beta < r \le \sigma - \theta$.

To see the dichotomy, examine the finitely many poles in the closed disk of radius $\sigma - \theta$. If this set is empty, case (i) holds. Otherwise take a nearest pole. Unless some pole has horizontal distance strictly less than the nearest distance, equality in the Euclidean distance bound forces the nearest pole to be real and unique. Increasing the radius slightly then gives case (ii).

In case (i), perturb the ordinate of ρ_1 to choose $z = \sigma + i\tau$ having a unique closest pole ρ at distance $d = |z - \rho| < D$. Only finitely many exceptional ordinates create ties among the relevant finite set of poles. Choose $d < r \le D$ so that $B(z, r)$ contains only ρ, and then choose $d < r' < r$. Subtracting its principal part and using Cauchy's estimates gives

$$A_k(z) = \frac{c_\rho}{(z-\rho)^{k+1}} + O((r')^{-k}). \tag{48}$$

The pole-free disk at the real center also gives $A_k(\sigma) = O((r')^{-k})$. Therefore the left side of (47) is bounded below by $|c_\rho| d^{-k-1} - C_3(r')^{-k}$.

In case (ii), use the real center $z = \sigma$ and $d = \sigma - \beta$. The same principal-part estimate applies, with $c_\beta > 0$. Consequently $\operatorname{Re} A_k(\sigma) \ge c_\beta d^{-k-1} - C_3(r')^{-k}$. The other term in (47) is nonnegative. The sign of the residue is essential here.

In either case, (47) implies, for arbitrarily large U and some $k \ge U$,

$$c d^{-k} \le C_3(r')^{-k} + C_2 e^{\eta U}(\sigma - \theta)^{-k}, \qquad 0 < d < r' \le \sigma - \theta, \tag{49}$$

with $c > 0$. Choose $0 < \eta < \log(r'/d)$. After multiplication by d^k, the right side tends to zero: the second term is at most $C_2 \exp[-(\log(r'/d) - \eta)U]$, and the first decays geometrically. This contradicts $c > 0$. Thus all poles satisfy $\operatorname{Re}\rho \le \theta$. Proposition 8.2, together with the known nonvanishing on $\operatorname{Re} s \ge 1$, proves Theorem 9.3.

10.3 Differentiating the Laplace transform in full detail

Fix a compact subset K of $\operatorname{Re} s > 2$ and put $a = \inf_{s \in K} \operatorname{Re} s > 2$. For every nonnegative integer k,

$$\left| g(u)(-u)^k e^{-su} \right| \le C_0 u^k e^{-(a-2)u} \qquad (s \in K, u \ge 0).$$

The dominating function is integrable, with integral $C_0 k!/(a-2)^{k+1}$. The same estimate holds on a small neighborhood of K after decreasing a slightly. Differentiation under the integral can therefore be iterated, giving

$$\Phi^{(k)}(s) = \int_0^\infty g(u)(-u)^k e^{-su}\, du.$$

This proves (43), including the factorial normalization and its sign. It also proves that the integral defines a holomorphic function in the entire initial half-plane, independently of the continuation data.

For real $\sigma > 2$, $A_k(\sigma)$ is real because its integrand is real. We retain a real-part notation so that the same expression can be used uniformly in complex-valued inequalities. All the estimates involving A_k are made at centers in the initial half-plane. Meromorphic continuation is used to analyze their derivatives through nearby poles, not to justify an integral representation to the left of the convergence line.

The elementary integral used here is

$$\int_0^\infty \frac{u^k}{k!} e^{-au}\, du = a^{-k-1} \qquad (a > 0). \tag{50}$$

It follows by the substitution $v = au$ and repeated integration by parts. The boundary terms vanish at infinity by exponential decay and at zero by the factor v^k, with the $k = 0$ case computed directly. Formula (50) explains both the factorial normalization and the exponential-tilting estimates for the tails.

10.4 A complete tail-concentration proof

We prove the existence form of (46) directly. The explicit constants displayed earlier give an explicit choice; the proof below uses flexible auxiliary parameters to exhibit the mechanism.

Lemma 10.1 (Two gamma tails)**.** *For every $\delta > 0$ there are $\sigma > 2$ and $C_1 > 0$ such that, for each $U \geq 1$, an integer $k \geq U$ satisfies* (46)*. One may take $k = \lceil c\sigma U \rceil$ with $c = 1 + \delta/2$ and a sufficiently large fixed σ.*

Proof. Choose $a > 0$ sufficiently small that

$$c \log(1 + a) > a.$$

Such a choice exists because the difference of the two sides vanishes at $a = 0$ and has derivative $c - 1 = \delta/2 > 0$ there. Similarly, choose $0 < b < 1$ sufficiently small that

$$b(1 + \delta) + c \log(1 - b) > 0.$$

The derivative of this expression at zero is $(1 + \delta) - c = \delta/2 > 0$.

Now fix σ so large that $\sigma > 2$, $a\sigma > 2$, $1 - b - 2/\sigma > 0$, $c\sigma \geq 1$, and both strict inequalities persist in the forms

$$c \log(1 + a - 2/\sigma) > a, \tag{51}$$

$$b(1 + \delta) + c \log(1 - b - 2/\sigma) > 0. \tag{52}$$

This is possible by continuity as $\sigma \to \infty$. All these parameters depend only on δ.

Set $k = \lceil c\sigma U \rceil$, so that $c\sigma U \leq k \leq c\sigma U + 1$ and $k \geq U$. On $0 \leq u \leq U$, $1 \leq e^{a\sigma(U-u)}$. Consequently, by (50),

$$\begin{aligned} I_- &:= \int_0^U \frac{u^k}{k!} e^{-(\sigma-2)u}\, du \\ &\leq e^{a\sigma U} \int_0^\infty \frac{u^k}{k!} e^{-(\sigma-2+a\sigma)u}\, du \\ &= \frac{\sigma^{-k}}{\sigma - 2 + a\sigma} \exp\left(a\sigma U - k \log(1 + a - 2/\sigma)\right). \end{aligned}$$

The logarithm is positive and $k \geq c\sigma U$. Hence the exponential factor is at most 1 by (51), and $I_- \leq C_-\sigma^{-k}$.

On $u \geq (1+\delta)U$, use $1 \leq e^{b\sigma(u-(1+\delta)U)}$. This gives

$$\begin{aligned} I_+ &:= \int_{(1+\delta)U}^{\infty} \frac{u^k}{k!} e^{-(\sigma-2)u}\, du \\ &\leq \frac{\sigma^{-k}}{\sigma - 2 - b\sigma} \exp(-b\sigma(1+\delta)U - k\log(1-b-2/\sigma))\,. \end{aligned}$$

Here the logarithm is negative. Using $k \leq c\sigma U + 1$, the exponent is at most

$$-\sigma U\big(b(1+\delta) + c\log(1-b-2/\sigma)\big) - \log(1-b-2/\sigma).$$

The first term is negative by (52); the second is a fixed constant. Thus $I_+ \leq C_+\sigma^{-k}$. Summing the two estimates proves the lemma with $C_1 = C_- + C_+$. □

The proof exposes an important separation of choices. The parameters a, b, σ are fixed before the majorant window is selected. The integer k then depends on U. The arithmetic window hypothesis has no dependence on k, so there is no need to solve a simultaneous condition involving a prescribed derivative order and a prescribed arithmetic scale.

10.5 Deriving the master inequality without a sign error

Let $W = [U, (1+\delta)U]$ and $w_k(u) = u^k e^{-\sigma u}/k!$. Since $w_k \geq 0$, the triangle inequality gives

$$|A_k(\sigma + i\tau)| \leq \int_0^{\infty} |g(u)| w_k(u)\, du, \qquad A_k(\sigma) = \int_0^{\infty} g(u) w_k(u)\, du.$$

Adding these expressions, rather than estimating each separately, gives an upper bound by $\int (|g| + g)w_k$. On W, this is at most $2\int_W \mathcal{R} w_k$. On its complement it is at most $2C_0 \int_{W^c} e^{2u} w_k$. This proves (45) once the kernel maximum has been inserted.

For the maximum, write $v = (\sigma - \theta)u \geq 0$. The term $v^k/k!$ is one term of the series for e^v, so $v^k e^{-v}/k! \leq 1$. Therefore

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

The bound is uniform in both u and k. It is intentionally less sharp than Stirling's formula; the proof needs only exponential control in k.

Using Lemma 10.1, the right side becomes

$$2e^{\eta U}(\sigma - \theta)^{-k} + 2C_0 C_1 \sigma^{-k}.$$

Because $0 < \theta < \sigma$ and $U \geq 1$, $\sigma^{-k} \leq (\sigma - \theta)^{-k}$ and $e^{\eta U} \geq 1$. Thus one may take $C_2 = 2 + 2C_0C_1$ in (47). It is independent of η, U, k, τ. The freedom to choose η later is consequently preserved.

10.6 The finite geometry behind the pole dichotomy

We justify the separation argument more explicitly. Put $R = \sigma - \theta$. The closed disk $\overline{B}(\sigma, R)$ lies in $\operatorname{Re} s \geq \theta > 0$. Local finiteness implies that it meets Z in a finite set F. If $F = \varnothing$, the first case holds with $D = R$, because the assumed pole ρ_0 has horizontal distance $\sigma - \operatorname{Re}\rho_0 < R$.

Otherwise choose a pole in F at minimum distance d from σ. The open disk of radius d is pole-free. If any pole of Z has horizontal distance less than d, this gives the first case with $D = d$. If no pole does, then every $\rho \in Z$ satisfies $\sigma - \operatorname{Re}\rho \geq d$. For a nearest pole,

$$d^2 = (\sigma - \operatorname{Re}\rho)^2 + (\operatorname{Im}\rho)^2 \quad \text{and} \quad \sigma - \operatorname{Re}\rho \geq d.$$

Both inequalities can hold only when $\operatorname{Im}\rho = 0$ and $\rho = \sigma - d$. In particular the nearest pole is unique. The existence of ρ_0 also implies $d < R$, since $d \leq \sigma - \operatorname{Re}\rho_0 < R$ in this branch. Finiteness of F now allows one to increase d to a radius $r \leq R$ without encountering any other pole. This is the second case.

In the first case, put $z_0 = \sigma + i \operatorname{Im}\rho_1$. Its distance to ρ_1 is the horizontal distance, which is less than D. Choose a small $h > 0$ so that $D + h < \sigma$ and every center $\sigma + i\tau$ with $|\tau - \operatorname{Im}\rho_1| < h$ still lies less than D from ρ_1. Any pole within distance D of such a center belongs to the finite set $Z \cap \overline{B}(z_0, D + h)$.

For two distinct poles v, w in this finite set, equality of their squared distances to $\sigma + i\tau$ reads

$$(\sigma - \operatorname{Re} v)^2 + (\tau - \operatorname{Im} v)^2 = (\sigma - \operatorname{Re} w)^2 + (\tau - \operatorname{Im} w)^2.$$

If their imaginary parts differ, this excludes at most one value of τ. If their imaginary parts agree, equality would force their real parts to agree as well, since both real parts are less than σ; that would contradict distinctness. Avoiding finitely many ordinates gives a unique nearest pole. A slightly larger radius isolates it and remains at most D. This proves the perturbation claim used above without a genericity assumption on the zero set.

10.7 Principal parts and high derivatives

Suppose $B(z, r)$ contains exactly one pole ρ, with residue c_ρ. Then

$$H(s) = \Phi(s) - \frac{c_\rho}{s - \rho}$$

extends holomorphically across ρ and throughout $B(z, r)$. Choose $d = |z - \rho| < r' < r$. By Cauchy's formula on the circle of radius r' centered at z,

$$\frac{|H^{(k)}(z)|}{k!} \leq M_{r'}(r')^{-k}, \qquad M_{r'} = \max_{|s-z|=r'} |H(s)|.$$

Since $(-1)^k (d/ds)^k (s-\rho)^{-1}/k! = (s-\rho)^{-k-1}$, this proves (48) with an explicit constant independent of k.

For a pole-free real-center disk, the same estimate applies to Φ itself. The comparison of the two derivative sequences then depends on the strict radius inequality $d < r'$. It does not depend on the argument of c_ρ or on the phase of $(z - \rho)^{-k-1}$, because the off-axis term is inside an absolute value. This avoids having to choose a subsequence of derivative orders with a favorable oscillatory phase.

For a real pole, $z = \sigma$ and $z - \rho = d > 0$. The leading term is a positive real number because $c_\rho > 0$. Its real part therefore cannot be lost to oscillation or to cancellation with a negative leading coefficient. These two mechanisms explain the two branches of the dichotomy.

10.8 The final exponential comparison

Let $q = d/r' < 1$ and choose $0 < \eta < \log(1/q)$. Multiplication of (49) by d^k gives

$$c \le C_3 q^k + C_2 e^{\eta U} \left(\frac{d}{\sigma - \theta}\right)^k \le C_3 q^U + C_2 e^{-(\log(1/q)-\eta)U}.$$

Both terms tend to zero as $U \to \infty$. More concretely, choose U_0 so large that each term is less than $c/3$ for $U \ge U_0$. The window hypothesis supplies such a U and the tail lemma supplies its $k \ge U$, producing the contradiction $c < 2c/3$. This is a contradiction from one sufficiently large window after all constants have been fixed.

The same proof works with a preliminary growth bound $C_0 e^{au}$ in place of $C_0 e^{2u}$, provided the initial center is chosen sufficiently far to the right. The number 2 is part of the convenient analytic packaging, not an arithmetic obstruction at a distinguished line. The critical value $1/2$ enters through the arithmetic exponent and the final small-parameter argument, rather than through the Laplace-transform machinery.

10.9 Reading the pole test in a single-character example

Consider first the family consisting only of the character of modulus 1 with weight 1. Its smoothed error is E_1 and its transform is (34). If $\rho = \beta + it$ is a nonreal zeta zero of multiplicity m, then conjugation gives the same multiplicity at $\bar\rho$, so the symmetrization leaves the residue

$$c_\rho = -\frac{m}{\rho(\rho - 1)}.$$

The window criterion says that upper bounds for this one real smoothed function on sufficiently long multiplicative intervals rule out all its zeros to the right of the chosen exponent. No separate estimate for the imaginary part of a character sum is needed in this example.

For a nonreal character χ, a singleton family still uses $\operatorname{Re} E_\chi$. A zero of $L(s,\chi)$ at ρ contributes a positive multiplicity term even if $L(s,\chi)$ has no zero at $\bar\rho$. The latter absence does not cause cancellation: its multiplicity is simply zero in the formula for M_ρ. The transform also detects zeros of the conjugate character because of the reflected term in (35).

For a hypothetical real zero $\beta \in (0,1)$ of a real nonprincipal character, the residue is $m/[\beta(1-\beta)] > 0$. The real-pole branch of the proof applies directly. In particular, the argument does not have to discard a possible exceptional real zero before using the window criterion. Instead, a bound with exponent $\theta < \beta$ is incompatible with that zero. This is an analytic implication; proving the corresponding arithmetic bound remains the separate construction problem.

10.10 Why high poles do not require a height-uniform argument

Suppose a pole has a very large imaginary part. The perturbed center $\sigma + i\tau$ is chosen near that ordinate, while the real center σ remains unchanged. The integral estimate is uniform in τ because $|e^{-i\tau u}| = 1$. Thus translating the center vertically costs nothing in the upper bound.

The constants in the Cauchy estimates may depend on the chosen pole and its height. This is allowed: the contradiction excludes one assumed pole at a time. The lower threshold for U can be increased after those constants have been fixed. The proof therefore does not imply a numerical height-uniform estimate for derivatives or a quantitative zero-density theorem. It establishes nonexistence of any pole in the forbidden half-plane under the stated window hypothesis.

This point is especially relevant when comparing the argument with classical zero-free regions. A quantitative zero-free region usually controls the dependence on the conductor and the imaginary part. Here the arithmetic premise itself is strong enough to let the contradiction be carried out separately for each fixed character and each hypothetical zero. The resulting conclusion is stronger in location but is conditional on those strong window premises.

10.11 Positivity and changes of family

The requirement $a_i > 0$ guarantees that every listed character is actually detected. If one allowed $a_i = 0$, a zero belonging only to that character would make no contribution to the transform. If negative weights were allowed, even two copies of the same character with weights 1 and -1 would give the identically zero smoothed function and identically zero transform. Every pole would cancel. This illustrates failure of the detection mechanism without making any assumption about the location of the original character's zeros.

Repeated characters with positive weights pose no problem. Their weights can be added, and the residue coefficient remains positive. Likewise, passing to a larger finite family is legitimate provided the pointwise or integral window estimate is established for the enlarged weighted sum. An estimate for one family does not automatically persist after adding a new character, since the added smoothed error can have either sign.

Inversion closure of a retained level is useful for its arithmetic construction and for organizing conjugate characters. It is not needed as a separate hypothesis of the basic analytic family criterion: the symmetrized transform already incorporates conjugation. This distinction helps separate an arithmetic design condition from an analytic necessity.

10.12 Harmless and non-harmless perturbations

There is a useful stability observation at the level of the window majorant. If a real perturbation $h(x)$ satisfies $h(x) \le Cx^{\theta}$ on the same windows, then its positive contribution to the weighted integral is at most $|C|\delta \log Y$. This can be absorbed into an arbitrarily small power of Y. In this sense fixed lower-order errors at exponent at most θ are compatible with the criterion.

However, an arbitrary perturbation of $G_{\mathcal{F}}$ need not preserve the pole data. A term cx^{β} has Mellin transform $c/(s-\beta)$ in its convergence half-plane. It can introduce a pole, or cancel an existing one if its coefficient is chosen with the opposite residue. The specific corrections at $s = 1$ are justified precisely because their effect on the transform is computed explicitly. Stability of the majorant alone is not a substitute for checking the meromorphic continuation and residue conditions of a modified function.

10.13 Contrapositive: persistent positive excursions

The analytic theorem has a useful contrapositive interpretation. A zero to the right of the proposed exponent forces a positive excursion of the smoothed family sum in every sufficiently late multiplicative window. The contrapositive can be stated both pointwise and in weighted integral form.

Corollary 10.2. *Suppose a member of the admissible positive family $\mathcal{F}$ has a zero ρ with $0 < \theta < \operatorname{Re}\rho < 1$. Fix $\delta > 0$. For every $C > 0$ there is Y_0 such that every $Y \ge Y_0$ admits $x \in [Y, Y^{1+\delta}]$ with*

$$G_{\mathcal{F}}(x) > Cx^{\theta}.$$

In fact there are $\eta > 0$ and Y_1 such that, for every $Y \geq Y_1$,

$$\int_Y^{Y^{1+\delta}} \max\{G_{\mathcal{F}}(x), 0\} x^{-\theta-1}\, dx > Y^{\eta}. \tag{53}$$

Proof. If the first conclusion failed for some C, then arbitrarily large Y would satisfy $G_{\mathcal{F}}(x) \leq Cx^{\theta}$ throughout the window. Theorem 9.3 would exclude the given zero.

For the stronger integral statement, use the canonical majorant $R(x) = \max\{G_{\mathcal{F}}(x), 0\}$. It is continuous and hence integrable on each compact window. If (53) failed in the stated eventual sense for every $\eta > 0$, then for every $\eta > 0$ and every lower threshold there would be a window on which this majorant's integral is at most Y^{η}. That is precisely the integral window condition. The pole argument would again exclude ρ. Negating these quantifiers proves the assertion. □

To relate the two forms quantitatively, let

$$M(Y) = \max_{Y \leq x \leq Y^{1+\delta}} \frac{\max\{G_{\mathcal{F}}(x), 0\}}{x^{\theta}}.$$

Continuity makes the maximum finite and attained. Its weighted integral is at most $M(Y)\delta \log Y$, so the stronger conclusion gives

$$M(Y) > \frac{Y^{\eta}}{\delta \log Y} \qquad (Y \text{ sufficiently large}).$$

Thus these positive excursions cannot be avoided by moving to an arbitrarily sparse sequence of long windows. This is exactly the feature that makes sparse arithmetic control sufficient for the nonvanishing argument.

The conclusion concerns the positive part of the combined smoothed sum. It does not locate the excursion at a prescribed point, nor does it say which individual character supplies the largest contribution there. It also gives no uniform effective value of Y_1 as the zero or the character family varies. Its use here is structural: the arithmetic construction supplies windows that are incompatible with the positive excursions a forbidden pole would force.

11 The small parameter reduction

We now combine the analytic criterion with a family of power estimates whose exponents approach the critical value.

Theorem 11.1 (Conditional reduction to Dirichlet nonvanishing)**.** *For every primitive Dirichlet character ψ, suppose there are a finite admissible positively weighted family $\mathcal{F}_\psi$ containing ψ, constants $A_\psi \in \mathbb{R}$, $\varepsilon_\psi > 0$, and $\delta_\psi > 0$ such that, for every $0 < \varepsilon < \varepsilon_\psi$ and every Y_0, there is $Y \geq Y_0$ with*

$$G_{\mathcal{F}_\psi}(x) \leq x^{1/2 + A_\psi \varepsilon} \quad \text{for all } x \in [Y, Y^{1+\delta_\psi}]. \tag{54}$$

Then (1) *holds for every Dirichlet character. Every zero in $0 < \operatorname{Re} s < 1$ of every Dirichlet L-function lies on $\operatorname{Re} s = 1/2$. In particular, the Riemann hypothesis follows.*

Proof of the primitive half-plane assertion. Fix ψ and suppose $L(\rho, \psi) = 0$ with $\beta = \operatorname{Re}\rho > 1/2$. Choose

$$a = \tfrac{1}{2} \min\{\beta - \tfrac{1}{2}, \tfrac{1}{2}\}, \qquad \theta = \tfrac{1}{2} + a,$$

so that $0 < \theta < 1$ and $\theta < \beta$. Choose $0 < \varepsilon < \varepsilon_\psi$ sufficiently small that $A_\psi \varepsilon \le a$. For $x \ge 1$, (54) is then bounded by x^θ. Applying Theorem 9.3 to the fixed family $\mathcal{F}_\psi$ gives a contradiction. The extension to all characters and the critical-line conclusion are proved in Section 14. □

The order of quantifiers is substantive. Allowing A_ψ to depend arbitrarily on ε would not imply that $A_\psi \varepsilon \to 0$. A family depending on ε could also be used if each chosen family separately satisfies the analytic criterion and contains the target character. The fixed-family formulation makes the order transparent: the family and loss exponent are fixed before ε is selected from the real part of the putative zero.

11.1 An explicit admissible small parameter

The choice in Theorem 11.1 can be made quantitative. If $a > 0$, $\varepsilon_\psi > 0$, and $A_\psi \in \mathbb{R}$, put

$$\varepsilon = \min\left\{\frac{\varepsilon_\psi}{2}, \frac{a}{|A_\psi| + 1}\right\}.$$

Then $0 < \varepsilon < \varepsilon_\psi$ and $A_\psi \varepsilon \le |A_\psi| \varepsilon \le a$. This also handles a negative loss coefficient without a special case.

There is no claim that the chosen ε is effective in terms of the conductor alone. It depends on the hypothetical zero through a, and on the comparison constants already attached to its character. That is enough for a contradiction proof of nonvanishing. An explicit zero-free region with numerical constants would require additional uniform information not supplied by this reduction.

11.2 Constants, thresholds, and loss exponents

Three types of parameter should be distinguished. A multiplicative constant C in $G_{\mathcal{F}}(x) \le C x^\theta$ can be absorbed in the integral majorant by raising the scale threshold. A threshold x_1 can be absorbed by choosing a later window. An exponent coefficient A cannot be handled in either of those ways: it must be controlled before ε is made small.

For example, suppose a comparison gives

$$G_{\mathcal{F}}(x) \le B_{\varepsilon,\psi} x^{1/2 + A_\psi \varepsilon} \quad \text{on arbitrarily large windows.}$$

For fixed ε, ψ, the constant $B_{\varepsilon,\psi}$ is harmless in the window criterion. It may grow faster than any power of $1/\varepsilon$. Alternatively, for $B_{\varepsilon,\psi} > 0$, one can absorb it into an additional factor x^ε beyond $x \ge B_{\varepsilon,\psi}^{1/\varepsilon}$. The latter method consolidates the final exponent. Either method preserves the order of quantifiers.

Quantity	Permitted dependence in the reduction
Family and weights	Fixed for the chosen target character in the fixed-family formulation.
A_ψ, ε_ψ	Chosen before ε; may depend on the character or family class.
Auxiliary family	May depend on the chosen ε.
Comparison constants	May depend on that family and ε.
Starting threshold	May depend on all previously fixed data.
Window endpoint Y	Chosen arbitrarily far beyond the requested threshold.

11.3 The gluing order for existence and compatibility

An additional quantifier issue occurs before the analytic estimate. A family-existence statement and a compatibility statement can have the shapes

$$\begin{aligned} \mathsf{Ex}: &\quad \forall \varepsilon\ \forall K\ \exists \mathcal{D}, \quad P(\mathcal{D}, K) \text{ and the required windows,} \\ \mathsf{Chain}: &\quad \forall \varepsilon\ \exists K\ \forall \mathcal{D}, \quad P(\mathcal{D}, K) \Longrightarrow \text{the comparison conditions.} \end{aligned}$$

After choosing K from Chain, one uses Ex with that same K to obtain the family. This produces one family carrying both the windows and the comparison inequalities. The quantifier $\forall K$ in the existence statement is crucial for this elementary composition.

If both statements merely said $\exists K \exists \mathcal{D}$, their witnesses could differ. One family might carry the windows and another the inequalities, with no implication for a common family. In the valid composition, the permitted ε ranges are intersected by taking their minimum.

11.4 No limit of geometric objects is required here

The proof that the critical exponent can be approached does not construct a limit of the geometric families as $\varepsilon \downarrow 0$. Nor does it require compactness of a moduli space of those families. Such geometric limits may occur elsewhere in the construction, but they are not part of this final analytic deduction. For a hypothetical zero with real part $\beta > 1/2$, one selects a single ε with $1/2 + A\varepsilon < \beta$, then applies the criterion to the corresponding fixed data.

This explains why the final theorem is naturally proved pointwise in the character and the alleged zero. It avoids assertions of uniformity that would be unnecessary for the stated nonvanishing conclusion.

11.5 A version with logarithmic losses

The power bound in Theorem 11.1 is a convenient final form, but the analytic argument tolerates several familiar lower-order factors. For every fixed $B \geq 0$ and $\kappa > 0$,

$$(1 + \log x)^B \leq x^{\kappa} \quad \text{for all sufficiently large } x. \tag{55}$$

To see this, put $u = \log x$. Then $B \log(1 + u) - \kappa u \to -\infty$, since $\log(1 + u)/u \to 0$. The same argument absorbs a fixed positive multiplicative constant.

Suppose, for each small ε, one has on arbitrarily large windows the bound

$$G_{\mathcal{F}}(x) \leq C_\varepsilon x^{1/2 + A\varepsilon} (1 + \log x)^{B_\varepsilon},$$

where C_ε and B_ε are finite and $B_\varepsilon \geq 0$. Fix the chosen ε and use (55) with $\kappa = \varepsilon/2$. Raise the threshold further so that $|C_\varepsilon| \leq x^{\varepsilon/2}$. The resulting bound is $x^{1/2 + (A+1)\varepsilon}$. The small-parameter reduction still applies. There is no need for B_ε to be uniformly bounded as $\varepsilon \downarrow 0$, because this absorption occurs after one ε has been fixed.

This observation explains the frequent use of eventual logarithmic estimates in the comparison argument. It also identifies a limitation: a factor of the form x^{B_ε} is an exponent loss, not a logarithmic loss. Its behavior as $\varepsilon \downarrow 0$ must be examined rather than automatically absorbed.

11.6 A more flexible conditional reduction

Proposition 11.2 (Varying finite families)**.** *Fix a primitive character ψ. Suppose there is a sequence $\theta_n \in (0,1)$ with $\theta_n \downarrow 1/2$. For each n, suppose there are a finite admissible positively weighted family $\mathcal{F}_n$ containing ψ, a number $\delta_n > 0$, and a finite constant C_n such that*

$$G_{\mathcal{F}_n}(x) \le C_n x^{\theta_n}$$

on arbitrarily large intervals $[Y, Y^{1+\delta_n}]$. Then $L(s,\psi) \neq 0$ whenever $\operatorname{Re} s > 1/2$, away from the principal pole.

Proof. If $L(\rho,\psi) = 0$ and $\operatorname{Re}\rho > 1/2$, choose n with $\theta_n < \operatorname{Re}\rho$. Apply Theorem 9.3 to the single fixed family $\mathcal{F}_n$, its width δ_n, and its constant C_n. The conclusion contradicts the zero at ρ. □

Thus neither a uniform lower bound for the window widths nor uniform analytic constants across n is logically necessary for this variant. A fixed-family hypothesis with a fixed window exponent is stronger. The variant distinguishes the uniformity needed for the analytic reduction from the uniformity that an arithmetic construction may supply.

11.7 Why a pointwise bound for each character is not extracted

The hypothesis bounds a positive weighted sum of real smoothed errors. It need not give an individual upper bound of the same size for every summand: some real smoothed errors may be negative and compensate others at a given x. The proof therefore does not isolate the target character by discarding the other summands pointwise.

Instead it passes to the transform, where all zero contributions have the common residue direction $-1/[\rho(\rho-1)]$ with nonnegative multiplicity coefficients. This is a different positivity statement. Pointwise cancellation among smoothed errors is compatible with the argument; cancellation of the residues detecting a zero is excluded. Confusing these two forms of positivity would lead to an invalid shortcut from the family estimate to an individual character estimate.

The distinction is useful for a researcher evaluating possible substitute arithmetic inputs. A method that proves only the combined window bound can still suffice. It need not establish the corresponding individual prime-sum bounds, provided it preserves the strictly positive character weights and the exact corrected-sum normalization.

12 Quantitative anatomy of the Laplace kernels

12.1 Normalization and concentration scale

For $a > 0$ and $k \in \mathbb{N}$, let

$$p_{k,a}(u) = \frac{a^{k+1}}{k!} u^k e^{-au} \mathbf{1}_{(0,\infty)}(u).$$

The gamma integral shows that $p_{k,a}$ has integral 1. Direct integration gives its first and second moments:

$$\int_0^\infty u p_{k,a}(u)\,du = \frac{k+1}{a}, \qquad \int_0^\infty u^2 p_{k,a}(u)\,du = \frac{(k+1)(k+2)}{a^2}.$$

Its variance is therefore $(k+1)/a^2$. Thus the mean has size k/a and the standard deviation has size $\sqrt{k}/a$. When k is proportional to U and a is fixed, the window $[U,(1+\delta)U]$ has width of order k, much larger than the natural fluctuation size of the kernel.

A variance estimate alone would only give polynomial tail decay. The pole contradiction needs decay at an exponential rate in k, because the pole contribution itself grows geometrically with k. Exponential tilting provides precisely that stronger estimate.

12.2 A two-sided exponential-tilting inequality

Lemma 12.1. *Let $k \geq 0$, $0 \leq l < s$, and $0 < U < V$. Then*

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

Proof. For $u \geq V$, $1 \leq e^{l(u-V)}$, and hence $e^{-su} \leq e^{-lV}e^{-(s-l)u}$. Integrating this bound over the upper tail and enlarging the domain to $(0,\infty)$ gives the first term. For $0 < u \leq U$, use $1 \leq e^{l(U-u)}$ to obtain the second term. The denominators are positive because $s-l>0$ and $s+l>0$. Endpoints do not affect the integrals. Adding the bounds proves the result. □

The use of a single parameter l for both tails is convenient, although separate choices could sharpen the constants. For the nonvanishing theorem one needs a strictly positive exponential margin; optimizing that margin is unnecessary.

Lemma 12.2 (Comparison with a fixed geometric scale)**.** *Suppose $a, \sigma > 0$ and*

$$x + k\frac{\sigma - a}{a} \leq c.$$

Then

$$e^x a^{-k-1} \leq \frac{e^c}{a}\sigma^{-k}.$$

Proof. The inequality $1 + t \leq e^t$ holds for every real t. With $t = (\sigma - a)/a$, it gives $\sigma/a \leq e^{(\sigma-a)/a}$. Raising to the nonnegative integer power k and multiplying by e^x yields $e^x(\sigma/a)^k \leq e^c$. Divide by $a\sigma^k$. □

12.3 A fully explicit choice of center

We now verify the constants appearing in the tail bound without a limiting choice of auxiliary parameters. Fix $\delta > 0$ and put

$$l = 2 + \frac{6}{\delta}, \quad M = l(1+\delta), \quad \sigma = M(2+l), \quad c = 1 + \frac{\delta}{2}.$$

These quantities depend only on δ. They satisfy

$$c(2+l) = M - 1, \qquad c(l-2) = l + 1, \qquad \sigma - 2 - l = (M-1)(l+2). \tag{56}$$

For example, the first identity follows from $\delta l = 2\delta + 6$ after expanding both sides. In particular $M > l > 2$, $\sigma > 2 + l$, and $M - 1 > 1$.

Proposition 12.3 (Explicit concentration bound)**.** *For every $U \geq 1$, set $k = \lceil c\sigma U \rceil$. Then $k \geq U$ and*

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

where

$$C_\delta = \frac{e}{\sigma - 2 - l} + \frac{1}{\sigma - 2 + l}.$$

Proof. We have $c\sigma U \leq k \leq c\sigma U + 1$. Since $c\sigma > 1$, this implies $k \geq U$. Apply Lemma 12.1 with $s = \sigma - 2$ and $V = (1+\delta)U$.

For the upper-tail term take $a_- = \sigma - 2 - l$. By (56),

$$\frac{\sigma - a_-}{a_-} = \frac{1}{M-1}, \qquad c\sigma = M(M-1).$$

Consequently

$$-l(1+\delta)U + \frac{k}{M-1} \leq -MU + MU + \frac{1}{M-1} \leq 1.$$

Lemma 12.2 gives the bound $e\, a_-^{-1}\sigma^{-k}$.

For the lower-tail term take $a_+ = \sigma - 2 + l$. Then $(\sigma - a_+)/a_+ = -(l-2)/a_+$. We claim $k(l-2) \geq lUa_+$. Indeed,

$$k(l-2) \geq c\sigma U(l-2) = (l+1)\sigma U,$$

and $(l+1)\sigma - l(\sigma - 2 + l) = \sigma + 2l - l^2 > 0$, since $\sigma > l(l+2)$. Thus $lU - k(l-2)/a_+ \leq 0$, and Lemma 12.2 gives $a_+^{-1}\sigma^{-k}$. Adding the two estimates proves the proposition. □

As δ decreases, this choice of σ grows on the order of δ^{-2}. This does not obstruct the theorem: δ is fixed before the pole argument begins. It does explain why the same proof does not automatically provide estimates uniform in arbitrarily shrinking window widths. A separate uniform argument would need to track how the pole-separation radii and derivative orders change with σ.

12.4 The factorial normalization

The coefficient $1/k!$ is important in two places. First, it turns Cauchy's derivative estimate into a geometric bound:

$$\left| \frac{h^{(k)}(z)}{k!} \right| \leq Mr^{-k}.$$

Second, it makes the exponential-series bound $v^k e^{-v}/k! \leq 1$ available for all $v \geq 0$. These two uses match the complex-analytic and real-variable sides of the argument.

Using an unnormalized derivative would introduce a common factor $k!$ into both sides of the final comparison. It would eventually cancel, but would obscure the exponential competition between d^{-k} and r^{-k}. The normalized notation displays the only growth rates that matter.

12.5 Pole derivatives and analytic remainders

For $z \neq \rho$,

$$\frac{d^k}{dz^k} \frac{c}{z-\rho} = (-1)^k k! \frac{c}{(z-\rho)^{k+1}}.$$

Thus the normalization preserves the residue and raises the inverse distance to the $(k+1)$st power. If $\Phi - c/(s-\rho)$ is holomorphic on $B(z,r)$, then for every $r' < r$ the compact circle $|s-z| = r'$ has a finite maximum M. Cauchy's estimate gives

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

The strict inequality $|z-\rho| < r'$ gives an exponential advantage to the pole term. Merely having a weak inequality would leave both terms at the same exponential rate and would not yield the desired contradiction.

This also clarifies why one selects a uniquely nearest pole. If two poles occur at the same distance, their normalized derivative contributions can have phases that depend on k and can cancel for infinitely many values of k. The small perturbation of the center removes that ambiguity before any window or derivative order is selected.

12.6 A quantitative terminal contradiction

Suppose $c > 0$, $0 < d < r \le R$, and nonnegative M, C satisfy

$$cd^{-k} \le Mr^{-k} + Ce^{\eta U} R^{-k}, \qquad k \ge U.$$

Set $a = \log(r/d) > 0$ and choose $\eta = a/2$. Multiplying by d^k gives

$$c \le Me^{-aU} + Ce^{-aU/2}.$$

For example, if

$$U > \max\left\{0, \frac{1}{a}\log\frac{2M+1}{c}, \frac{2}{a}\log\frac{2C+1}{c}\right\},$$

both terms are strictly less than $c/2$. The contradiction is therefore obtained at a finite, although not necessarily effective in arithmetic terms, threshold depending on the selected pole and analytic disks.

The window hypothesis supplies arbitrarily large U after a has been fixed. This is why no lower bound for a uniform over all zeros is needed. The proof excludes each possible zero separately, with its own finite configuration of constants.

13 Variants and stability of the window hypothesis

13.1 The canonical positive majorant

Let $F : [1,\infty) \to \mathbb{R}$ be continuous and write $F_+(x) = \max\{F(x), 0\}$. For fixed θ and $\delta > 0$, define

$$I_{\theta,\delta}(Y) = \int_Y^{Y^{1+\delta}} F_+(x) x^{-\theta-1}\, dx, \qquad Y > 1.$$

The integral exists on every compact window. The integral majorant condition is equivalent to

$$\forall \eta > 0 \quad \forall Y_0 > 1 \quad \exists Y \ge Y_0 : \; I_{\theta,\delta}(Y) \le Y^{\eta}. \tag{57}$$

Indeed, any nonnegative majorant of F also majorizes F_+, and conversely F_+ itself is an admissible majorant. Thus the freedom to choose a majorant does not enlarge the condition for continuous functions; it is a convenient way to verify it using an explicit bound.

Proposition 13.1 (A logarithmic liminf formulation)**.** *Condition* (57) *is equivalent to*

$$\liminf_{Y\to\infty} \frac{\log \max\{1, I_{\theta,\delta}(Y)\}}{\log Y} = 0.$$

Proof. The quotient is nonnegative. If (57) holds, then for every $\eta > 0$ the quotient is at most η at arbitrarily large Y, which forces the liminf to be zero. Conversely, if the liminf is zero, then for every $\eta > 0$ and every threshold there is a larger Y with quotient less than η. Exponentiation gives the required bound. The use of $\max\{1, I\}$ also covers windows on which F_+ vanishes identically. □

This reformulation uses only a liminf. A limsup bound would be stronger and would require control of all sufficiently late windows. The analytic proof needs only the ability to choose one favorable window after the pole geometry has fixed an exponent η.

13.2 Monotonicity in the exponent

If $\theta' \geq \theta$, then for $x \geq 1$, $x^{-\theta'-1} \leq x^{-\theta-1}$. Thus the canonical window condition at exponent θ implies it at exponent θ'. This monotonicity is consistent with the conclusion: excluding poles to the right of θ is stronger than excluding them to the right of θ'.

The reverse implication need not hold. For the simple function $F(x) = x^\beta$ with $\theta < \beta < \theta'$, the integral at exponent θ' is bounded as Y grows, while the integral at exponent θ grows like a fixed positive power of Y. The threshold in the majorant is therefore not just a bookkeeping parameter.

13.3 Weighted L^p estimates

For $1 < p < \infty$, define

$$J_p(Y) = \int_Y^{Y^{1+\delta}} (F_+(x)x^{-\theta})^p \frac{dx}{x}.$$

Proposition 13.2. *Suppose that for every $\varepsilon > 0$ there are arbitrarily large Y with $J_p(Y) \leq Y^\varepsilon$. Then the integral window condition at exponent θ holds.*

Proof. Hölder's inequality for the measure dx/x gives

$$I_{\theta,\delta}(Y) \leq J_p(Y)^{1/p} (\delta \log Y)^{1-1/p}.$$

Given $\eta > 0$, use the hypothesis with $\varepsilon = p\eta/2$. On the resulting windows the first factor is at most $Y^{\eta/2}$. For all sufficiently large Y, the logarithmic factor is also at most $Y^{\eta/2}$. Raising the threshold before selecting the window proves (57). □

The measure dx/x is natural here because it becomes ordinary Lebesgue measure on the logarithmic interval. A bound measured with dx has a different scaling and cannot be substituted without accounting for the extra power of x.

For $p = 1$ the statement is the definition. A uniform pointwise bound $F_+(x)x^{-\theta} \leq C$ is the corresponding elementary L^∞ case, giving $I_{\theta,\delta}(Y) \leq C\delta \log Y$. These statements place the power-bound criterion inside a family of integral sufficient conditions.

13.4 Exceptional subsets of a window

One can permit a bad subset of each selected window if its weighted contribution is controlled. Let $E_Y \subset [Y, Y^{1+\delta}]$ be measurable. Suppose

$$F(x) \le Cx^\theta \quad \text{outside } E_Y, \qquad |F(x)| \le C_0 x^a \quad \text{throughout the window,}$$

where $a \ge \theta$ and the constants are fixed. Then

$$I_{\theta,\delta}(Y) \le |C|\delta \log Y + C_0 Y^{(1+\delta)(a-\theta)} \int_{E_Y} \frac{dx}{x}. \tag{58}$$

The proof is a direct split of the integral. On the bad set, $F_+(x)x^{-\theta} \le C_0 x^{a-\theta}$, whose maximum occurs at the upper endpoint because $a - \theta \ge 0$.

Thus a small logarithmic measure can compensate for a large pointwise bound on the exceptional set. For example, if

$$\int_{E_Y} \frac{dx}{x} \le Y^{-(1+\delta)(a-\theta)}$$

on arbitrarily large selected windows, the bad-set contribution in (58) is bounded. This is a sufficient condition, not a claim that the arithmetic construction supplies such exceptional sets. It explains exactly what must be checked if a pointwise estimate is available only almost everywhere with a quantitative exceptional set.

13.5 Perturbations of smaller order

Let F satisfy the integral window condition at exponent θ. Suppose H is measurable and $|H(x)| \le Cx^\alpha$ for all sufficiently large x, where $\alpha < \theta$. Then $F + H$ satisfies the same condition. Indeed,

$$(F + H)_+ \le F_+ + |H|,$$

and

$$\int_Y^{Y^{1+\delta}} |H(x)| x^{-\theta-1}\, dx \le \frac{C}{\theta - \alpha} Y^{\alpha-\theta}.$$

Given a target $\eta > 0$, first apply the window condition for F with $\eta/2$. The resulting contribution is at most $Y^{\eta/2}$, and the perturbation contribution tends to zero. Their sum is at most Y^η beyond a further threshold.

The strict inequality $\alpha < \theta$ is more than necessary. At $\alpha = \theta$ the extra contribution is $C\delta \log Y$, which can still be absorbed. A perturbation of size $x^{\theta+\varepsilon}$ for a fixed positive ε is different: its weighted integral has a fixed power growth and generally changes the exponent that can be excluded. The distinction mirrors the difference between multiplicative constants and exponent losses in the small-parameter reduction.

13.6 Sums require common windows

If F and H each satisfy the window condition, it does not follow solely from their separate quantifiers that $F + H$ does. Their favorable windows may lie in disjoint regions. A valid sufficient condition is that for each η there are arbitrarily large *common* windows on which both canonical integrals are at most $Y^{\eta/2}$. Then $(F + H)_+ \le F_+ + H_+$ gives the conclusion after absorbing the factor 2.

This issue has an arithmetic counterpart. Bounds for different ranks, local lengths, or character averages cannot be combined merely because each holds on some long interval. The estimates must be realized on the same interval and for the same selected objects. The analytic argument has enough freedom to move its starting threshold; it does not have the power to intersect arbitrary sparse families of intervals.

13.7 Continuous versus measurable formulations

Continuity makes the canonical integral finite on compact windows, but is not needed in the abstract Laplace theorem. A measurable function with an exponential growth bound is locally integrable on $(0,\infty)$, so its positive part can be used as a majorant almost everywhere. Altering such a function on a null set does not alter its Laplace transform or any integral window estimate.

A pointwise upper bound on a closed interval is stronger than an almost-everywhere bound there. Both are sufficient for the integral argument. This distinction can be useful when an arithmetic expression is first defined as a measurable representative and only later shown to be continuous. The corrected ramp sums considered here are continuous, so no representative ambiguity remains in their final pointwise statements.

14 Imprimitive characters and the critical line

Proposition 14.1 (Change of conductor)**.** *If every primitive Dirichlet L-function is nonzero in* $\operatorname{Re} s > 1/2$*, away from the principal pole, then* (1) *holds for every Dirichlet character.*

Proof. Let χ modulo N be induced by the primitive character χ^* of conductor $f \mid N$. The change-of-level identity is

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

For $\operatorname{Re} s > 0$, $|\chi^*(p)p^{-s}| \le p^{-\operatorname{Re} s} < 1$, so each extra factor is nonzero. The pole condition transfers correctly: if the inducing character is principal, then so is its change of level. The primitive nonvanishing therefore implies the assertion. □

Proposition 14.2 (Reflection in the critical strip)**.** *Assume* (1)*. If* $0 < \operatorname{Re}\rho < 1$ *and* $L(\rho,\chi) = 0$*, then* $\operatorname{Re}\rho = 1/2$*.*

Proof. The case $\operatorname{Re}\rho > 1/2$ is excluded directly. For $\operatorname{Re}\rho < 1/2$, first pass to the primitive inducing character, using the nonvanishing of the extra Euler factors. For a primitive χ^*, write

$$\Lambda(s,\chi^*) = (f/\pi)^{(s+a)/2}\Gamma((s+a)/2)L(s,\chi^*), \qquad a \in \{0,1\}, \quad \chi^*(-1) = (-1)^a.$$

In the open critical strip the gamma and conductor factors are finite and nonzero. The functional equation $\Lambda(s,\chi^*) = \epsilon_{\chi^*}\Lambda(1-s,\overline{\chi^*})$ therefore sends the zero at ρ to a zero of the inverse character at $1-\rho$. Its real part exceeds $1/2$, contradicting (1). □

For the zeta function, the same argument gives the critical-strip statement. The usual functional equation and the nonvanishing on $\operatorname{Re} s \ge 1$ show that the zeros outside the strip are the negative even integers, with $\zeta(0) = -1/2$. Accounting for these boundary and trivial-zero cases gives

$$\zeta(s) = 0, \quad s \notin \{-2,-4,-6,\dots\}, \quad s \ne 1 \quad \Longrightarrow \quad \operatorname{Re} s = \tfrac{1}{2}.$$

This completes the remaining deductions in Theorem 11.1.

14.1 Why missing Euler factors do not affect the open strip

Let χ modulo N be induced by a primitive character χ^* of conductor f. In the absolutely convergent half-plane, the Euler products differ only at primes dividing N but not f. This gives the identity in Proposition 14.1 there; meromorphic continuation extends it to the rest of the plane away from the usual exceptional point for the principal character.

For a missing Euler factor, write $\chi^*(p) = e^{i\alpha}$ when its value is nonzero. Its zeros satisfy

$$e^{i\alpha}e^{-s\log p} = 1.$$

Taking absolute values forces $\operatorname{Re} s = 0$, and the remaining condition is $\operatorname{Im} s = (\alpha + 2\pi k)/\log p$ for some integer k. If $\chi^*(p) = 0$, the factor is identically 1 and has no zeros. Thus every extra Euler-factor zero lies on the boundary line $\operatorname{Re} s = 0$.

This explains the scope of the critical-strip statement for imprimitive characters. The assertion concerns zeros in $0 < \operatorname{Re} s < 1$; it does not claim that every zero of an imprimitive L-function in the whole complex plane is inherited from its primitive inducing character. The additional boundary zeros are a genuine feature of the change of modulus.

At $s = 1$, a principal character has a meromorphic pole. The condition $(\chi, s) \neq (\chi_0, 1)$ is preserved when passing between a principal character and its primitive inducing character of modulus 1. For a nonprincipal character, the inducing character cannot be principal, since changing the level of the principal character again gives a principal character. Thus the exceptional pole condition is respected by the change-of-conductor identity.

14.2 The nonvanishing factors in the functional equation

For primitive conductor f and parity $a \in \{0, 1\}$, the completed function is

$$\Lambda(s, \chi) = (f/\pi)^{(s+a)/2}\Gamma((s + a)/2)L(s, \chi).$$

When $0 < \operatorname{Re} s < 1$, the gamma argument has positive real part. The gamma function is holomorphic and nonzero there. The conductor factor is a complex exponential with positive real base, hence is also nonzero. Therefore a zero of L in this strip is exactly a zero of Λ.

The functional equation multiplies the reflected completed function by a nonzero root number. The argument needs only its nonvanishing; its exact phase is irrelevant. A zero at ρ with real part below $1/2$ produces a zero at $1 - \rho$ for the conjugate character, whose real part is above $1/2$. The already established right-half-plane assertion for all primitive characters then gives the contradiction.

There are two equivalent ways to express the symmetry. The functional equation relates χ at s to $\bar{\chi}$ at $1 - s$. Conjugating the Dirichlet series and then continuing analytically gives $L(\bar{s}, \bar{\chi}) = \overline{L(s, \chi)}$. Combining the two relates zeros of a fixed character to their reflections $1 - \bar{\rho}$. The all-character hypothesis lets us apply the first form directly.

14.3 Boundary lines and the trivial zeta zeros

The Riemann hypothesis is a statement about the nontrivial zeta zeros, so the boundary and negative-real-axis cases must be accounted for. The nonvanishing on $\operatorname{Re} s \geq 1$, away from the pole,

was proved before the transform discussion. It does not follow simply by substituting $\operatorname{Re} s = 1$ into the absolutely convergent Euler product; nonvanishing on that boundary is a separate analytic result.

For $\operatorname{Re} s \le 0$, use the completed functional equation in the form

$$\zeta(s) = \pi^{s-1/2}\Gamma((1-s)/2)D(s/2)\zeta(1-s), \qquad D = 1/\Gamma. \tag{59}$$

For $s \neq 0$ in this region, the exponential factor is nonzero, the gamma argument has positive real part, so its value is finite and nonzero, and $\zeta(1-s)$ is finite and nonzero by the boundary theorem and Euler-product nonvanishing. The zeros therefore occur precisely at the zeros of $D(s/2)$, namely the negative even integers after $s = 0$ is excluded. A nonzero point on the imaginary axis contributes no additional zero.

At $s = 0$, the zero of $D(s/2)$ cancels the reflected zeta pole. The reciprocal product gives $D(s/2) = s/2 + O(s^2)$, while $\zeta(1-s) = -1/s + O(1)$. The Gaussian integral gives $\Gamma(1/2) = \sqrt{\pi}$, so taking the limit in (59) yields $\zeta(0) = -1/2$. Thus every zero other than a negative even integer lies in the open critical strip, where Proposition 14.2 applies under its stated nonvanishing hypothesis.

All statements here concern meromorphic functions in the usual sense. The pole at $s = 1$ is excluded from assertions about zeros or finite values.

15 Spectral estimates and finite-dimensional comparison

15.1 Gram operators and the dimension of an image

Let V and H be finite-dimensional complex Hilbert spaces and $T : V \to H$ a linear map. Its Gram operator is $G = T^*T$. It is self-adjoint and positive semidefinite, since $\langle Gv, v\rangle = \|Tv\|^2 \ge 0$. Moreover, $\ker G = \ker T$, so $\operatorname{rank} G = \operatorname{rank} T$.

Let $\lambda_1, \dots, \lambda_r$ be its positive eigenvalues, counted with multiplicity. Cauchy–Schwarz gives

$$(\operatorname{Tr} G)^2 = \left(\sum_{j=1}^{r} \lambda_j\right)^2 \le r \sum_{j=1}^{r} \lambda_j^2 = r \operatorname{Tr}(G^2).$$

If $T \neq 0$, this yields

$$\operatorname{rank} T \ge \frac{(\operatorname{Tr} G)^2}{\operatorname{Tr}(G^2)}. \tag{60}$$

The right side is sometimes called an effective rank. Its importance here is that traces can be estimated block by block, without constructing a basis of the image.

15.2 Block estimates

Suppose $V = \bigoplus_{u \in I} V_u$, where I is finite, and let $T_u : V_u \to H$ be the restrictions of T. Put

$$d_u = \dim V_u, \qquad N = \sum_{u \in I} d_u, \qquad K_{uv} = T_u^* T_v.$$

The matrix (K_{uv}) represents G. For a linear map between finite Hilbert spaces the Hilbert–Schmidt norm is $\|A\|_{\mathrm{HS}}^2 = \operatorname{Tr}(A^*A)$. In orthonormal bases it is the sum of the squared absolute values of the matrix entries. Thus

$$\operatorname{Tr}(G^2) = \|G\|_{\mathrm{HS}}^2 = \sum_{u,v} \|K_{uv}\|_{\mathrm{HS}}^2.$$

Proposition 15.1 (Rank from approximate orthogonality)**.** *Suppose $N > 0$, $0 \le e < 1$, $\eta \ge 0$, and $D > 0$. Assume*

$$\|K_{uu} - I\|_{\mathrm{HS}} \le e\sqrt{d_u}, \qquad \|K_{uv}\|_{\mathrm{HS}}^2 \le (1+\eta)\frac{d_u d_v}{D} \quad (u \ne v).$$

Then

$$\operatorname{rank} T \ge \frac{(1-e)^2 N}{(1+e)^2 + (1+\eta)N/D}.$$

Proof. The Hilbert–Schmidt inner product with the identity gives $|\operatorname{Tr}(K_{uu}-I)| \le \sqrt{d_u}\|K_{uu}-I\|_{\mathrm{HS}} \le e d_u$. Hence $\operatorname{Tr} G \ge (1-e)N > 0$. The triangle inequality gives $\|K_{uu}\|_{\mathrm{HS}} \le (1+e)\sqrt{d_u}$, and therefore

$$\begin{aligned}\operatorname{Tr}(G^2) &\le (1+e)^2 N + \frac{1+\eta}{D}\sum_{u \ne v} d_u d_v \\ &\le (1+e)^2 N + \frac{1+\eta}{D} N^2.\end{aligned}$$

Insert these inequalities into (60) and cancel a factor of N. □

This is a model spectral estimate. Applying it to a geometric family requires actual maps T_u whose adjoint products are the kernels under consideration. Bounds on an arbitrary array of matrices do not by themselves produce a Gram realization. Positive semidefiniteness of the whole block matrix, or a direct construction of the maps, is an additional requirement.

15.3 When the ambient dimension is recovered

The denominator in Proposition 15.1 has two contributions. The term $(1+e)^2$ records diagonal errors. The term $(1+\eta)N/D$ records the collective off-diagonal contribution. If N/D is large, the bound approaches

$$\operatorname{rank} T \gtrsim \frac{(1-e)^2}{1+\eta} D.$$

If N/D is small, the bound instead approaches a positive proportion of N. Thus the estimate interpolates between the size of the domain and a reference ambient dimension.

A rank estimate close to D need not give surjectivity onto a D-dimensional space unless the error is small enough to exploit integrality. If $\dim H = D \in \mathbb{N}$ and one proves $\operatorname{rank} T > D - 1$, then surjectivity follows because rank is an integer. A relative lower bound $(1 - o(1))D$ is not enough when the product of the relative error and D still grows. This distinction matters when dimensions grow rapidly with the arithmetic scale.

15.4 Perturbation of an invertible map

A different route to surjectivity is a norm perturbation argument. Suppose $T_0 : V \to H$ is an isomorphism and $\|T_0^{-1}(T - T_0)\|_{\mathrm{op}} < 1$. Then

$$T = T_0(I + A), \qquad A = T_0^{-1}(T - T_0),$$

and $(I+A)^{-1} = \sum_{n \ge 0}(-A)^n$ converges in operator norm. Thus T is an isomorphism and

$$\|T^{-1}\|_{\mathrm{op}} \le \frac{\|T_0^{-1}\|_{\mathrm{op}}}{1 - \|A\|_{\mathrm{op}}}.$$

This conclusion depends on a reference isomorphism, which is stronger geometric information than a collection of trace estimates.

Hilbert–Schmidt control implies operator-norm control because $\|A\|_{\mathrm{op}} \le \|A\|_{\mathrm{HS}}$. The converse involves a factor of the square root of rank. Consequently one must keep track of which norm an estimate supplies before applying a perturbation theorem in a family whose dimension varies.

15.5 Shared objects in a comparison argument

Suppose a geometric calculation produces a rank r_1, an arithmetic calculation produces a rank r_2, and each is bounded by an expression that resembles the other. One cannot combine the bounds until the two ranks have been identified, or a comparison map between the corresponding spaces has been established. The same issue arises for metrics, lattices, normalizations of volume, and the choice of a trace pairing.

A useful mathematical way to formulate a comparison hypothesis is to choose a single collection of spaces, maps, lattices, and measures first, and impose all equalities and inequalities on that collection. This prevents a statement of the form “there exists a space with the rank bound” from being combined with an unrelated statement of the form “there exists a lattice with the divisibility bound.” Existence in the two statements is not automatically simultaneous.

The analytic nonvanishing criterion is insensitive to how these objects are constructed. It needs their final scalar estimate on common windows. The construction problem, however, is sensitive to every identification that makes that scalar estimate meaningful.

16 Restriction operators, reproducing kernels, and orthogonality

The rank estimate used in the arithmetic comparison is most useful when its Gram matrix comes from a restriction operator. We develop this construction before discussing the estimates imposed on it. The distinction between an evaluation kernel and a projection kernel is particularly important: both involve sums of matrix coefficients, but their normalization is determined by different measures.

16.1 Evaluation in a finite-dimensional space of functions

Let (X, μ) be a measure space and let H be a d-dimensional subspace of $L^2(X, \mu)$. Suppose that each element of H has a specified representative as a function on a set $X_0 \subset X$. This assumption is needed because evaluation at a point is not defined for an arbitrary L^2 equivalence class. In applications the specified representatives are continuous or smooth. Fix an orthonormal basis $e_1, \dots, e_d$ of H and put

$$K(x, y) = \sum_{j=1}^{d} e_j(x)\overline{e_j(y)}, \qquad x, y \in X_0.$$

We use inner products linear in the first variable throughout this section. For each $y \in X_0$, the function $K_y = K(\,\cdot\,, y)$ belongs to H. If $f = \sum_j a_j e_j$, then

$$\langle f, K_y \rangle = \sum_j a_j e_j(y) = f(y).$$

Thus K_y represents the evaluation functional at y.

Proposition 16.1 (Intrinsic kernel)**.** *The kernel K is independent of the chosen orthonormal basis. It satisfies*

$$\begin{aligned} K(y,x) &= \overline{K(x,y)}, \\ |K(x,y)|^2 &\le K(x,x)K(y,y), \\ \int_X K(x,z)K(z,y)\,d\mu(z) &= K(x,y), \\ \int_X K(x,x)\,d\mu(x) &= d. \end{aligned}$$

The integral identities are interpreted using the specified representatives where they are defined almost everywhere.

Proof. The vector representing a linear functional on a finite-dimensional Hilbert space is unique, so a change of orthonormal basis does not change K_y. Alternatively, the change-of-basis matrix is unitary, and expanding the double sum cancels its two matrix factors. Hermitian symmetry follows from the definition. Since $K(x,y) = \langle K_y, K_x\rangle$, the second identity is Cauchy–Schwarz. For the third, expansion gives

$$\sum_{i,j} e_i(x)\overline{e_j(y)} \int_X \overline{e_i(z)}e_j(z)\,d\mu(z) = \sum_i e_i(x)\overline{e_i(y)}.$$

All sums are finite, so no interchange theorem beyond linearity is needed. The last identity follows from $\|e_j\|_2 = 1$. □

The diagonal $K(x,x)$ is nonnegative and equals the squared norm of the evaluation functional. It is consequently a local measure of the size of H. Its integral is the dimension, but it need not be constant. A homogeneous space often supplies enough symmetry to make it constant; that conclusion requires invariance of both the space of functions and the measure.

16.2 Restriction to a finite set

Choose points $x_1, \dots, x_N \in X_0$ and positive weights $w_1, \dots, w_N$. Define

$$T : H \longrightarrow \mathbb{C}^N, \qquad (Tf)_i = \sqrt{w_i} f(x_i),$$

where $\mathbb{C}^N$ carries its ordinary Euclidean inner product. The adjoint is

$$T^* a = \sum_{i=1}^{N} \sqrt{w_i} a_i K_{x_i}.$$

Indeed, substituting the reproducing identity into $\langle Tf, a\rangle$ gives $\langle f, T^*a\rangle$. Thus the two Gram operators are

$$T^*Tf = \sum_i w_i f(x_i) K_{x_i}, \qquad (TT^*)_{ij} = \sqrt{w_i w_j} K(x_i, x_j).$$

They have the same positive eigenvalues, with multiplicities. To see this directly, if $T^*Tv = \lambda v$ and $\lambda > 0$, then $Tv \neq 0$ and $TT^*(Tv) = \lambda Tv$. Conversely use T^*; these maps identify the two eigenspaces up to the nonzero factor λ.

Proposition 16.2 (Sampling rank formula)**.** *If at least one evaluation functional is nonzero, then*

$$\operatorname{rank} T \geq \frac{\left(\sum_i w_i K(x_i, x_i)\right)^2}{\sum_{i,j} w_i w_j |K(x_i, x_j)|^2}.$$

Proof. The trace of TT^* is the sum of its diagonal entries. Its squared Hilbert–Schmidt norm is the sum of the squared absolute values of all entries. Apply (60) to TT^*. □

This formula identifies precisely which off-diagonal quantity must be small. A bound on $\sum_{i \neq j} K(x_i, x_j)$ is insufficient because its terms may cancel. The rank estimate requires the nonnegative sum of $|K(x_i, x_j)|^2$. Conversely, the diagonal must contribute a substantial positive trace. A small total Hilbert–Schmidt norm alone says nothing useful about rank if the entire operator is also small.

16.3 Exact quadrature and tight frames

Suppose the weighted points integrate all products of functions from H:

$$\sum_i w_i f(x_i)\overline{g(x_i)} = \int_X f(x)\overline{g(x)}\, d\mu(x) \qquad (f, g \in H). \tag{61}$$

Then $T^*T = I_H$. This follows by pairing T^*Tf with every $g \in H$. In particular, T is injective, $N \geq d$, and the vectors $\sqrt{w_i} K_{x_i}$ form a tight frame:

$$\sum_i |\langle f, \sqrt{w_i} K_{x_i} \rangle|^2 = \|f\|^2.$$

The matrix TT^* is an orthogonal projection of rank d. Consequently

$$\sum_i w_i K(x_i, x_i) = d, \qquad \sum_{i,j} w_i w_j |K(x_i, x_j)|^2 = d.$$

Proposition 16.2 is then an equality.

A tight frame can contain more vectors than the dimension. Its Gram matrix on $\mathbb{C}^N$ then has $N - d$ zero eigenvalues. It would be incorrect to infer that TT^* is close to I_N merely because $T^*T = I_H$. The correct identity matrix depends on which side of the restriction operator is being considered.

Example 16.3 (Roots of unity)**.** Let H consist of the trigonometric polynomials $f(z) = \sum_{j=0}^{d-1} a_j z^j$ on the unit circle, with normalized Haar measure. Choose the Nth roots of unity and weights $1/N$, where $N \geq d$. Then

$$\frac{1}{N} \sum_{z^N = 1} z^{j-k} = \begin{cases} 1, & j = k, \\ 0, & j \neq k, \end{cases} \qquad 0 \leq j, k < d.$$

Thus (61) holds. The kernel is $K(z, w) = \sum_{j=0}^{d-1} (z\overline{w})^j$. For $z \neq w$, the geometric-series expression $(1 - (z\overline{w})^d)/(1 - z\overline{w})$ describes its off-diagonal size. When $N = d$, all off-diagonal entries vanish. When $N > d$, they usually do not vanish, although the frame identity remains exact.

16.4 Approximate quadrature and eigenvalue control

Suppose that, for some $0 \le \eta < 1$,

$$\left| \sum_i w_i |f(x_i)|^2 - \|f\|^2 \right| \le \eta \|f\|^2 \qquad (f \in H).$$

Since $T^*T - I$ is self-adjoint, the inequality is equivalent to $\|T^*T - I\|_{\mathrm{op}} \le \eta$. Every eigenvalue of T^*T lies in $[1-\eta, 1+\eta]$, so T is injective and

$$\sqrt{1-\eta}\,\|f\| \le \|Tf\| \le \sqrt{1+\eta}\,\|f\|.$$

The inverse from $T(H)$ to H has norm at most $(1-\eta)^{-1/2}$. This is stronger than a trace rank bound: it controls every singular value rather than their first two moments.

There is a useful intermediate conclusion when only a Hilbert–Schmidt estimate is available. If $\|T^*T - I\|_{\mathrm{HS}}^2 \le E$, then at most E/η^2 eigenvalues lie outside $[1-\eta, 1+\eta]$. Indeed, each such eigenvalue contributes more than η^2 to the sum $\sum_j (\lambda_j - 1)^2$. In particular,

$$\dim \ker T \le E,$$

because every zero eigenvalue contributes exactly one. An absolute bound $E < 1$ gives injectivity. A relative bound $E = o(d)$ gives a small relative kernel, but permits a kernel whose dimension tends to infinity.

16.5 Compact-group averaging

Let K now be a compact group with Haar measure of total mass one, and let $\pi : K \to U(V)$ be a continuous unitary representation on a finite-dimensional complex vector space. For $A \in \mathrm{End}(V)$ define

$$\mathcal{A}(A) = \int_K \pi(k) A \pi(k)^{-1}\, dk.$$

The integral can be taken entry by entry in a basis. Left invariance of Haar measure shows that $\mathcal{A}(A)$ commutes with every $\pi(h)$. If π is irreducible, it follows that

$$\mathcal{A}(A) = \frac{\operatorname{Tr} A}{\dim V} I_V. \tag{62}$$

Here is a proof of the needed algebraic fact. An endomorphism B of a nonzero finite-dimensional complex space has an eigenvalue λ. If B commutes with $\pi(K)$, its nonzero eigenspace $\ker(B - \lambda I)$ is invariant. Irreducibility makes it all of V, so $B = \lambda I$. Taking traces in the average then determines λ and gives (62).

For an orthonormal basis, apply the formula to the matrix unit E_{jl} and read the (i,k) entry. One obtains

$$\int_K \pi(g)_{ij} \overline{\pi(g)_{kl}}\, dg = \frac{\delta_{ik}\delta_{jl}}{\dim V}. \tag{63}$$

If π and σ are inequivalent irreducible representations, the corresponding cross integral is zero. To prove this, average a map $A : W \to V$ as $\int_K \pi(g) A \sigma(g)^{-1} dg$. The average intertwines σ and π. A nonzero intertwiner between irreducibles has invariant kernel and image, hence is an isomorphism, contrary to inequivalence.

16.6 Multiplicity spaces

A finite-dimensional unitary representation splits orthogonally as

$$H = \bigoplus_{\pi} V_\pi \otimes M_\pi,$$

where V_π carries an irreducible representation and M_π is a multiplicity space on which the group acts trivially. To obtain the splitting, choose an invariant irreducible subspace of minimal positive dimension. Its orthogonal complement is invariant because the action is unitary. Repeat until the space is exhausted, and collect equivalent summands.

An equivariant endomorphism has the form $\bigoplus_\pi I_{V_\pi} \otimes B_\pi$. Indeed, the cross blocks between inequivalent irreducibles vanish by the preceding argument, and a block between equivalent copies is scalar on the irreducible factor. Therefore

$$\operatorname{Tr} B = \sum_{\pi} (\dim V_\pi) \operatorname{Tr} B_\pi, \qquad \operatorname{rank} B = \sum_{\pi} (\dim V_\pi) \operatorname{rank} B_\pi.$$

This explains how representation dimensions enter a restriction estimate. A scalar estimate on the multiplicity space can represent a much larger rank contribution after multiplication by $\dim V_\pi$.

The decomposition is an orthogonal decomposition of a Hilbert space. It does not automatically preserve a specified integral lattice. A rational or integral comparison therefore needs further information about its projectors. The denominator of an averaging projector may contribute to a lattice index even when its complex operator norm equals one.

16.7 Schur's lemma on a Hilbert space

The finite-dimensional eigenvalue proof does not apply to an arbitrary bounded operator on an infinite-dimensional Hilbert space: its spectrum may contain no eigenvalues. The appropriate replacement uses the continuous functional calculus for bounded self-adjoint operators.

Proposition 16.4 (Bounded commutant)**.** *Let π be an irreducible unitary representation on a nonzero complex Hilbert space. A bounded operator commuting with all $\pi(g)$ is scalar.*

Proof. First suppose $A = A^*$. If its spectrum contained distinct points a, b, choose continuous real functions f, g on the spectrum with disjoint supports, such that $f(a) = g(b) = 1$. The operators $f(A)$ and $g(A)$ are nonzero by the spectral mapping theorem, and $f(A)g(A) = 0$. Each commutes with the representation. The closure of the range of $g(A)$ is thus a nonzero invariant closed subspace, while it is contained in the proper closed subspace $\ker f(A)$. This contradicts irreducibility. Hence the spectrum is a singleton $\{a\}$ and the functional calculus gives $A = aI$.

For general A, commutation with the unitary operators $\pi(g)$ also implies commutation of A^* with them, by taking adjoints and replacing g by g^{-1}. Both $(A + A^*)/2$ and $(A - A^*)/(2i)$ are self-adjoint members of the commutant. They are real scalars by the first part, and therefore A is a complex scalar. □

The closedness of invariant subspaces is essential in this statement. The closure of a range is used because the range of a bounded operator need not be closed. The proof also makes explicit why unitarity is useful: it ensures that the commutant is closed under taking adjoints.

16.8 Formal degrees and measure normalization

For a unitary representation π of a locally compact group, write $c_{v,w}(g) = \langle \pi(g)v, w\rangle$. Suppose a positive number d_π satisfies the square-integrable orthogonality relation

$$\int_G c_{v,w}(g)\overline{c_{v',w'}(g)}\,dg = \frac{\langle v, v'\rangle\langle w', w\rangle}{d_\pi}. \tag{64}$$

We call d_π the formal degree for the chosen Haar measure. In the compact case with probability measure, (63) gives $d_\pi = \dim \pi$. For a noncompact group, existence of a relation such as (64) is an additional representation-theoretic statement; it is not obtained merely by using the compact calculation.

If Haar measure is replaced by $c\,dg$ with $c > 0$, the formal degree becomes d_π/c. If a quotient covolume is computed with the same measure, that covolume is multiplied by c. Thus the product $d_\pi \operatorname{vol}(\Gamma\backslash G)$ is unchanged. This is the natural normalization for a dimension prediction coming from a projection kernel on a quotient.

For a unit vector w, the relation implies

$$\int_G |\langle \pi(g)w, w\rangle|^2\,dg = d_\pi^{-1}.$$

More generally, for nonzero w, $d_\pi = \|w\|^4/\int_G |c_{w,w}(g)|^2 dg$. The expression is independent of w precisely because the orthogonality relation holds for all vectors. A formula using a single vector does not by itself prove that it gives a formal degree for the representation.

16.9 Convolution and projection

Let G be unimodular and let $f \in L^1(G)$. For a unitary representation R on a Hilbert space, define

$$R(f)v = \int_G f(g)R(g)v\,dg.$$

The norm estimate $\|R(f)\| \le \|f\|_1$ follows by integrating $\|f(g)R(g)v\|$. With convolution $(f{*}h)(x) = \int_G f(y)h(y^{-1}x)dy$, Fubini gives $R(f{*}h) = R(f)R(h)$. The involution is $f^*(g) = \overline{f(g^{-1})}$, and unitarity gives $R(f^*) = R(f)^*$. Therefore, if $f * f = f = f^*$, the operator $R(f)$ is an orthogonal projection.

Each assertion has an analytic hypothesis. Absolute integrability gives the bounded operator and justifies Fubini. Unimodularity removes the modular-function factor from the involution. If a matrix coefficient is only in $L^2(G)$, these L^1 arguments cannot be used without an additional construction or approximation. Square integrability and integrability play different roles in the passage from coefficients to projection operators.

16.10 Periodization on a quotient

Let Γ be a discrete subgroup of a unimodular group G. Suppose f is sufficiently integrable and its periodization converges absolutely so that the following manipulations are justified. For a function F on $\Gamma\backslash G$, right convolution can be written

$$R(f)F(x) = \int_{\Gamma\backslash G} K_f(x,y)F(y)\,dy, \qquad K_f(x,y) = \sum_{\gamma\in\Gamma} f(x^{-1}\gamma y).$$

The formula follows by decomposing the integral over G into translates of a measurable fundamental domain. The kernel is independent of the chosen representatives of x and y because changing either one reindexes the sum over Γ.

On the diagonal,

$$K_f(x,x) = f(1) + \sum_{\gamma \neq 1} f(x^{-1}\gamma x).$$

If the operator is a finite-rank projection with a continuous kernel and the diagonal trace formula is justified, then

$$\dim \operatorname{im} R(f) = f(1)\operatorname{vol}(\Gamma\backslash G) + \int_{\Gamma\backslash G} \sum_{\gamma \neq 1} f(x^{-1}\gamma x)\, dx. \tag{65}$$

This is the elementary shape of the main-term and remainder decomposition behind many dimension estimates. Formula (65) does not claim that its hypotheses hold for every coefficient or every quotient. In a noncompact quotient the remainder and the continuous spectrum require separate analysis.

16.11 A uniform decay criterion for the geometric remainder

The convergence issue can be isolated in a simple counting lemma. Let $\ell : \Gamma \to [0,\infty)$ be a proper length function, and suppose

$$\#\{\gamma : m \le \ell(\gamma) < m+1\} \le Ae^{bm} \qquad (m \in \mathbb{N}).$$

If $|a(\gamma)| \le Ce^{-c\ell(\gamma)}$ with $c > b$, then

$$\sum_{\gamma \in \Gamma} |a(\gamma)| \le AC \sum_{m \ge 0} e^{-(c-b)m} = \frac{AC}{1 - e^{-(c-b)}}.$$

The same proof gives a uniform bound for a family a_x when A, C, b, c are uniform in x. If only the terms with $\ell(\gamma) \ge R$ occur, the bound gains a factor of order $e^{-(c-b)R}$.

The strict inequality $c > b$ is the decisive point. Exponential decay of an individual coefficient is not enough when the number of group elements grows exponentially at a faster rate. At equality the displayed majorant is not summable. Quantitative geometry enters by producing both a counting exponent and a coefficient decay exponent, with a positive gap that remains uniform over the chosen family.

16.12 Tensor products and accumulated errors

Suppose local Hilbert spaces H_v and operators A_v are given at a finite set of places. On $\bigotimes_v H_v$ one has

$$\operatorname{Tr}\left(\bigotimes_v A_v\right) = \prod_v \operatorname{Tr} A_v, \qquad \left\|\bigotimes_v A_v\right\|_{\mathrm{HS}}^2 = \prod_v \|A_v\|_{\mathrm{HS}}^2.$$

These identities follow by choosing product orthonormal bases. If $A_v = I + E_v$ and $\|E_v\|_{\mathrm{op}} \le \eta_v$, expansion of the tensor product gives

$$\left\|\bigotimes_v A_v - I\right\|_{\mathrm{op}} \le \prod_v (1+\eta_v) - 1 \le \exp\left(\sum_v \eta_v\right) - 1.$$

Thus local errors must be summable over the set of places. A fixed small error at each place need not give a small global error when the number of places grows. If $\sum_v \eta_v \le 1$, the last expression is at most $(e-1)\sum_v \eta_v$ by convexity of the exponential on $[0,1]$.

The same issue appears when normalizing formal degrees, volume factors, and restriction maps. A product of apparently harmless constants can be large. Taking logarithms converts such products into sums and makes the accumulated dependence visible. This is one reason why the later arithmetic comparison is naturally expressed through logarithmic indices and normalized slopes.

17 Hermitian domains and explicit reproducing kernels

The restriction kernels at infinite places arise from unitary representations. A concrete model is provided by holomorphic functions on a matrix ball. In this model the invariant measure, the reproducing kernel, and its decay can all be computed directly. The calculation illustrates how an archimedean weight creates the decay needed to control off-diagonal terms.

17.1 Hermitian forms of mixed signature

Let $p, q \ge 1$. A nondegenerate Hermitian form on $\mathbb{C}^{p+q}$ has a basis in which its matrix is

$$J = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}.$$

To obtain such a basis, diagonalize its Hermitian matrix unitarily and rescale the eigenvectors by the square roots of the absolute values of the nonzero eigenvalues. The numbers p and q are intrinsic: a positive-definite subspace has dimension at most p, and the first p coordinate directions attain that bound. The corresponding unitary group is

$$U(p,q) = \{g \in \mathrm{GL}_{p+q}(\mathbb{C}) : g^* J g = J\}.$$

Taking determinants gives $|\det g| = 1$.

The block diagonal subgroup $U(p) \times U(q)$ preserves the positive and negative coordinate spaces. The quotient by this compact subgroup can be realized as the set of negative q-planes. A negative plane projects isomorphically onto the last q coordinates: a nonzero vector with vanishing last coordinates would have positive norm. It is therefore the graph of a matrix $Z \in M_{p,q}(\mathbb{C})$, and negativity is the condition $Z^*Z - I_q < 0$.

17.2 The matrix ball

Define

$$\mathcal{D}_{p,q} = \{Z \in M_{p,q}(\mathbb{C}) : I_q - Z^*Z > 0\}.$$

Equivalently $\|Z\|_{\mathrm{op}} < 1$. This set is open, bounded, and convex, since the operator norm is a norm. Its complex dimension is pq. The identity

$$\det(I_p - ZZ^*) = \det(I_q - Z^*Z)$$

follows because the nonzero eigenvalues of ZZ^* and Z^*Z agree with multiplicity. Both determinants are positive on the domain.

Write an element of $U(p,q)$ in blocks as $g = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. Its action on graphs is

$$g \cdot Z = (AZ + B)(CZ + D)^{-1}. \tag{66}$$

The denominator is invertible. If $(CZ + D)v = 0$ for nonzero v, the negative vector (Zv, v) would be sent to a vector with zero negative coordinates, contradicting preservation of the Hermitian form. The image graph is again negative, so the action preserves the domain.

17.3 The denominator identity

Let $Z' = g \cdot Z$ and $J_g(Z) = CZ + D$. The graph identity is

$$g \begin{pmatrix} Z \\ I_q \end{pmatrix} = \begin{pmatrix} Z' \\ I_q \end{pmatrix} J_g(Z).$$

Pairing both sides using J gives

$$I_q - Z'^* Z' = J_g(Z)^{-*}(I_q - Z^* Z) J_g(Z)^{-1}. \tag{67}$$

Taking determinants and writing $j(g, Z) = \det J_g(Z)$,

$$\det(I - Z'^* Z') = \frac{\det(I - Z^* Z)}{|j(g, Z)|^2}.$$

A two-point version follows by pairing the graph at Z with the graph at W:

$$I - (gW)^*(gZ) = J_g(W)^{-*}(I - W^* Z) J_g(Z)^{-1}. \tag{68}$$

This identity will determine the reproducing kernel transformation.

The denominators also satisfy a cocycle identity: $J_{gh}(Z) = J_g(hZ) J_h(Z)$, and hence $j(gh, Z) = j(g, hZ) j(h, Z)$. This is obtained by composing the two graph identities. It is the algebraic reason that the automorphy factors below define a group action.

17.4 Transitivity

For $W \in \mathcal{D}_{p,q}$, put

$$A_W = (I_p - WW^*)^{-1/2}, \qquad D_W = (I_q - W^* W)^{-1/2}.$$

The relation $A_W W = W D_W$ follows first for polynomials in WW^* and W^*W, and then for the inverse-square-root function by uniform polynomial approximation on their compact positive spectra. Define

$$g_W = \begin{pmatrix} A_W & W D_W \\ W^* A_W & D_W \end{pmatrix}.$$

A block multiplication using $A_W W = W D_W$ shows $g_W^* J g_W = J$. Moreover $g_W \cdot 0 = W D_W D_W^{-1} = W$. Thus $U(p,q)$ acts transitively, with stabilizer of zero equal to $U(p) \times U(q)$. The domain is therefore a concrete realization of that quotient.

This formula does not require a choice of singular vectors. It is continuous in W and is well suited to transporting the evaluation functional at zero to an arbitrary point.

17.5 The Jacobian

Differentiate (66). For a tangent matrix H,

$$d(g\cdot)_Z(H) = (A - Z'C)H(CZ + D)^{-1}.$$

Left multiplication by a $p \times p$ matrix L on $M_{p,q}$ has complex determinant $(\det L)^q$, while right multiplication by a $q \times q$ matrix R has determinant $(\det R)^p$. Also

$$\det g = \det(CZ + D)\det(A - Z'C),$$

which follows by multiplying g by the determinant-one block matrix $\begin{pmatrix} I_p & Z \\ 0 & I_q \end{pmatrix}$ and taking a Schur complement. Therefore the squared absolute value of the complex Jacobian is

$$|\operatorname{Jac}_{\mathbb{C}}(g\cdot)(Z)|^2 = |j(g, Z)|^{-2(p+q)}. \tag{69}$$

The real volume Jacobian is this square, because the map is complex linear on each tangent space.

Combining (67) and (69) shows that

$$d\mu(Z) = \det(I - Z^*Z)^{-(p+q)}dZ \tag{70}$$

is invariant. Here dZ is ordinary Lebesgue measure on the $2pq$ real coordinates. Any other invariant measure differs by a positive constant, which must be reflected in a formal-degree calculation.

17.6 A matrix beta integral

For $\alpha > -1$, let

$$I_{p,q}(\alpha) = \int_{\mathcal{D}_{p,q}} \det(I - Z^*Z)^{\alpha}dZ.$$

We claim

$$I_{p,q}(\alpha) = \pi^{pq}\prod_{j=1}^{q}\frac{\Gamma(\alpha + j)}{\Gamma(\alpha + p + j)}. \tag{71}$$

Split $Z = (Z_0, z)$ into its first $q - 1$ columns and its last column. The Schur-complement identity gives

$$\det(I - Z^*Z) = \det(I - Z_0^*Z_0)\left(1 - z^*(I - Z_0Z_0^*)^{-1}z\right).$$

For fixed Z_0, substitute $z = (I - Z_0Z_0^*)^{1/2}u$. The allowed u form the unit ball in $\mathbb{C}^p$, and the real Jacobian is $\det(I - Z_0Z_0^*) = \det(I - Z_0^*Z_0)$. Hence

$$I_{p,q}(\alpha) = I_{p,q-1}(\alpha + 1)\int_{\|u\|<1}(1 - \|u\|^2)^{\alpha}du.$$

The ball integral is $\pi^p\Gamma(\alpha + 1)/\Gamma(\alpha + p + 1)$, by polar coordinates and the beta integral. Iterating the recurrence proves (71).

The convergence condition $\alpha > -1$ is visible already in the last radial integral near the boundary. The recurrence proves convergence in every matrix dimension under that same condition. It also shows that increasing the number of columns shifts the exponent in the remaining integral; ignoring the Jacobian would miss that shift.

17.7 The holomorphic Hilbert space

Fix an integer $\lambda \geq p+q$ and put $\alpha = \lambda - p - q \geq 0$. Let $\mathcal{H}_\lambda$ consist of holomorphic functions on $\mathcal{D}_{p,q}$ with finite norm

$$\|f\|_\lambda^2 = \int_{\mathcal{D}_{p,q}} |f(Z)|^2 \det(I - Z^*Z)^\alpha dZ.$$

Point evaluation is continuous. On a small Euclidean ball whose closure lies in the domain, the weight is bounded below by a positive constant. The mean-value inequality for $|f|^2$ on that ball bounds the value at its center by a constant times the displayed norm. The same argument uniformly on compact subsets shows that an L^2-Cauchy sequence of holomorphic functions converges locally uniformly to a holomorphic limit. Thus $\mathcal{H}_\lambda$ is a Hilbert space.

The integer restriction on λ lets us use powers of the holomorphic nonzero function $j(g,Z)$ without choosing branches. More general parameters can be treated on covering groups, but that additional representation-theoretic choice is unnecessary for the present calculation.

17.8 Evaluation at the origin

The measure is invariant under $Z \mapsto e^{i\theta}Z$. For any $f \in \mathcal{H}_\lambda$, Fubini and the mean-value property in the scalar variable give

$$\int_{\mathcal{D}_{p,q}} f(Z) \det(I - Z^*Z)^\alpha dZ = I_{p,q}(\alpha) f(0).$$

The integral is absolutely convergent by Cauchy–Schwarz and the finiteness of $I_{p,q}(\alpha)$. For each fixed Z in the domain, the circle average of $f(e^{i\theta}Z)$ is $f(0)$, because the scalar holomorphic function extends to a disk of radius greater than one.

Thus the vector representing evaluation at zero is the constant function $C_\lambda = I_{p,q}(\alpha)^{-1}$. By (71),

$$C_\lambda = \pi^{-pq} \prod_{j=1}^{q} \frac{\Gamma(\lambda - q + j)}{\Gamma(\lambda - p - q + j)}. \tag{72}$$

The constant function one has squared norm $1/C_\lambda$.

17.9 The unitary action and the kernel

The operators

$$(T_g f)(Z) = j(g,Z)^{-\lambda} f(gZ)$$

are unitary by the measure transformation formulas. They form an anti-representation, and $\pi(g) = T_{g^{-1}}$ is a representation. Transporting evaluation vectors under these operators gives the kernel transformation

$$K_\lambda(gZ, gW) = j(g,Z)^\lambda \overline{j(g,W)}^\lambda K_\lambda(Z,W).$$

The two-point determinant identity (68), together with the value at zero and transitivity, therefore yields

$$K_\lambda(Z,W) = C_\lambda \det(I - W^*Z)^{-\lambda}. \tag{73}$$

The determinant is nonzero because $\|W^*Z\| < 1$. It is holomorphic in Z and antiholomorphic in W, as required.

The calculation proves the reproducing identity through unitary transport, not merely by proposing a function with the right diagonal. A positive diagonal alone would not determine a reproducing kernel. The full covariance identity and evaluation at every point are what identify (73).

17.10 Irreducibility in this model

The span of the evaluation vectors is dense in $\mathcal{H}_\lambda$: a function orthogonal to every evaluation vector vanishes at every point and is therefore zero. Transitivity shows that these vectors belong, up to nonzero scalars, to the orbit of the constant evaluation vector at zero.

Let P be the orthogonal projection onto a closed invariant subspace. Because the representation is unitary, the orthogonal complement is also invariant, so P commutes with the action. The subgroup $\mathrm{diag}(e^{i\theta}I_p, I_q)$ acts by scalar rotation of Z and fixes the constant function. Its fixed holomorphic functions are precisely the constants: expansion at zero into homogeneous polynomials kills every positive degree. Hence P sends the constant vector to a scalar multiple of itself. As a projection, that scalar is zero or one. Commutation and density of the orbit then give $P = 0$ or $P = I$. The representation is irreducible.

This proof uses a special one-dimensional compact-subgroup type and a cyclic vector. In more general unitary representations the lowest compact type may have larger dimension, and the corresponding argument requires its full representation structure.

17.11 Normalized overlaps

Let k_Z be the unit vector obtained by normalizing the evaluation vector at Z. Formula (73) gives

$$|\langle k_Z, k_W\rangle| = \left(\frac{\det(I - Z^*Z)^{1/2}\det(I - W^*W)^{1/2}}{|\det(I - W^*Z)|}\right)^{\lambda}. \tag{74}$$

The factor in parentheses is at most one, as also follows abstractly from Cauchy–Schwarz. At $W = 0$ it is $\det(I - Z^*Z)^{1/2}$.

Let $r_1, \dots, r_m$ be the nonzero singular values of Z, where $m \le \min(p, q)$, and write $r_j = \tanh t_j$ with $t_j \ge 0$. Then

$$|\langle k_Z, k_0\rangle| = \prod_{j=1}^{m}(1 - r_j^2)^{\lambda/2} = \prod_{j=1}^{m}(\cosh t_j)^{-\lambda}.$$

Since $\cosh t \ge e^t/2$, this is at most $2^{m\lambda}e^{-\lambda\sum_j t_j}$. Near the origin, $\log\cosh t = \frac{1}{2}t^2 + O(t^4)$, so the overlap has approximately Gaussian decay in the small radial parameters, with concentration scale $\lambda^{-1/2}$.

17.12 The formal-degree normalization

Choose Haar measure on $U(p, q)$ so that its quotient by the compact stabilizer, equipped with probability Haar measure, is (70). The squared absolute value of the matrix coefficient of the unit constant vector is $\det(I - Z^*Z)^{\lambda}$. Its integral over the group is therefore

$$\int_{\mathcal{D}_{p,q}} \det(I - Z^*Z)^{\lambda - p - q} dZ = I_{p,q}(\lambda - p - q) = C_\lambda^{-1}.$$

In a Schur orthogonality normalization as in (64), the formal degree is consequently C_λ. Rescaling Haar measure rescales this value inversely, exactly as discussed earlier.

The matrix beta integral makes the dependence on the weight explicit. For integer λ, the gamma ratios are finite products, and C_λ is a polynomial in λ of degree pq, times π^{-pq}. More precisely,

$$C_\lambda = \pi^{-pq}\prod_{j=1}^{q}\prod_{r=0}^{p-1}(\lambda - p - q + j + r).$$

Every factor is positive for $\lambda \geq p+q$. Thus the local formal degree grows like $\pi^{-pq}\lambda^{pq}$ at fixed p, q.

17.13 From local kernels to quotient estimates

Suppose a discrete group acts on the domain and a periodized kernel is formed from (73). The identity element gives the local diagonal term C_λ times the chosen invariant volume. Every other term contains an overlap between a point and one of its translates. Formula (74) supplies an explicit decay factor for that overlap.

To sum these factors, one still needs a uniform count of group elements with bounded radial parameters. If the number in a shell of radius R is at most Ae^{bR} and the coefficient is at most Ce^{-cR} with $c > b$, the geometric remainder converges by the shell-counting lemma. The weight λ helps make c large, while the geometry of the arithmetic quotient controls A and b. This identifies the two quantitative ingredients required for an off-diagonal estimate.

The calculation does not construct an arithmetic lattice in the holomorphic representation, a global automorphic realization, or a comparison with local cohomology. Those are additional mathematical constructions. It does provide the explicit archimedean kernel and its normalization, so that such constructions can be checked against a concrete analytic model rather than an unspecified norm estimate.

18 Localization kernels and weighted length bounds

18.1 Continuous degree-one cohomology

Let G be a profinite group and M a discrete G-module. A continuous one-cocycle is a continuous function $c : G \to M$ satisfying

$$c(\sigma\tau) = c(\sigma) + \sigma c(\tau).$$

A one-coboundary has the form $c_m(\sigma) = \sigma m - m$ for some $m \in M$. The abelian group $H^1(G, M)$ is the quotient of continuous one-cocycles by one-coboundaries. If M is a module over a ring commuting with the action of G, then the cocycles, coboundaries, and cohomology inherit that module structure.

A continuous homomorphism $G' \to G$ gives a restriction map by precomposing cocycles. A G-equivariant homomorphism $M \to M'$ gives a map on cohomology by postcomposition. Both operations preserve coboundaries, which is the elementary check needed for them to descend to the quotient.

For a trivial action, the cocycle condition reduces to the homomorphism condition and every coboundary is zero. Thus $H^1(G, M) = \mathrm{Hom}_{\mathrm{cont}}(G, M)$ in that case. For a nontrivial Tate twist this simplification is not valid, so the cocycle description is useful for keeping the action visible.

18.2 The negative Tate twist

Fix a prime ℓ and a number field E. The ℓ-adic cyclotomic character is $\kappa_\ell : G_E \to \mathbb{Z}_\ell^\times$, describing the action on all ℓ-power roots of unity. For an integer j, let

$$M_\ell(j) = (\mathbb{Q}_\ell/\mathbb{Z}_\ell)(-j)$$

denote the discrete module whose action is $\sigma \cdot m = \kappa_\ell(\sigma)^{-j} m$. In particular, positive j means a negative twist. The underlying abelian group is unchanged when j varies; its Galois action is what changes.

Let S_ℓ consist of the places above ℓ together with the archimedean places, and let G_{E,S_ℓ} be the Galois group of the maximal extension unramified outside S_ℓ. The cyclotomic action factors through this group. Choose compatible embeddings of decomposition groups at the finite places above ℓ. These give a localization map

$$H^1(G_{E,S_\ell}, M_\ell(j)) \longrightarrow \prod_{v \mid \ell} H^1(G_{E_v}, M_\ell(j)).$$

Define its kernel by

$$\operatorname*{Sha}_\ell(E, j) = \ker \left(H^1(G_{E,S_\ell}, M_\ell(j)) \longrightarrow \prod_{v \mid \ell} H^1(G_{E_v}, M_\ell(j)) \right). \tag{75}$$

This notation refers to the stated localization problem. No abelian variety is implicit in the definition.

18.3 Local triviality and coboundaries

A global class belongs to (75) precisely when, for each $v \mid \ell$, its restricted cocycle is a coboundary. Thus a representative c satisfies

$$c(\sigma) = \sigma m_v - m_v \qquad (\sigma \in G_{E_v})$$

for some $m_v \in M_\ell(j)$ that may depend on v. There is no requirement that the same m_v work at every place, or that one global m trivialize the cocycle. The possible failure of that global trivialization is exactly what the kernel measures.

Changing a decomposition-group embedding by conjugation does not change the localization condition under its natural transport. To see the cocycle identity behind this, define $c_g(\sigma) = g\, c(g^{-1}\sigma g)$. Expanding the cocycle relation gives

$$c_g(\sigma) = c(\sigma) + (\sigma - 1)c(g).$$

Thus c_g differs from c by a coboundary. This calculation explains why compatible changes of embeddings preserve the relevant cohomology class. When quantitative hypotheses are required uniformly over all allowed local choices, that uniformity should nevertheless be stated explicitly.

18.4 Length over the discrete valuation ring

For a $\mathbb{Z}_\ell$-module M, its length is the number of simple factors in a composition series when such a finite series exists. A finite-length $\mathbb{Z}_\ell$-module has simple factors isomorphic to $\mathbb{F}_\ell$, since $\mathbb{Z}_\ell$ is a local principal ideal domain with unique maximal ideal $\ell\mathbb{Z}_\ell$.

Lemma 18.1. *If M has finite $\mathbb{Z}_\ell$-length s, then M is finite and $|M| = \ell^s$. For a short exact sequence of finite-length modules $0 \to M' \to M \to M'' \to 0$,*

$$\operatorname{length} M = \operatorname{length} M' + \operatorname{length} M''.$$

Proof. A composition series has s quotients of cardinality ℓ. Cardinality multiplies in a finite short exact sequence, so induction on s gives $|M| = \ell^s$. For the second assertion, concatenate a composition series of M' with the inverse images in M of a composition series of M''. The resulting series has the asserted number of factors. □

Finiteness of a length is an input to this argument, not a conclusion from assigning it a natural-number notation. A localization kernel may have infinite length unless a separate theorem or hypothesis excludes that possibility.

18.5 The uniform weighted bound

For an integer $t \geq 0$ and a real number $B \geq 0$, write $\mathcal{V}(E, t, B)$ for the following condition. All the groups $\mathrm{Sha}_\ell(E, j)$ have finite length for $1 \leq j < t$, and for every finite set P of rational primes,

$$\sum_{\ell \in P} (\log \ell) \sum_{1 \leq j < t} \operatorname{length}_{\mathbb{Z}_\ell} \operatorname{Sha}_\ell(E, j) \leq B. \tag{76}$$

When local auxiliary data enter the construction of the localization map, the condition is imposed for every compatible choice of those data. The arguments below can first be applied with one such choice fixed.

Proposition 18.2 (Consequences of the weighted bound)**.** *Suppose* (76) *holds, and put* $s_\ell = \sum_{1 \leq j < t} \operatorname{length}_{\mathbb{Z}_\ell} \mathrm{Sha}_\ell(E, j)$. *Then*

(i) $s_\ell \leq B / \log \ell$ *for every prime* ℓ*;*

(ii) at most $\lfloor B / \log 2 \rfloor$ *primes have* $s_\ell > 0$*;*

(iii) if $s_\ell > 0$*, then* $\ell \leq e^B$*;*

(iv) the finite product $\prod_\ell \ell^{s_\ell}$ *is at most* e^B*.*

Proof. Take $P = \{\ell\}$ to prove (i). If $s_\ell > 0$, it is an integer at least 1, so the contribution from that prime is at least $\log 2$. Any finite collection of such primes therefore has size at most $B / \log 2$. If there were more primes in the support, a finite subcollection would contradict this bound. This proves (ii), including finiteness of the support.

For a supported prime, $\log \ell \leq s_\ell \log \ell \leq B$, which proves (iii). Apply (76) to the full finite support and exponentiate to obtain (iv). □

By Lemma 18.1, the last product is the product of the cardinalities of the localization kernels in the indicated twist range. Thus the logarithmic weights are natural: they convert additive length information at different primes into a single multiplicative size bound. The proposition does not identify those groups or construct their elements; it gives structural consequences of the assumed estimate.

18.6 Monotonicity and budget allocation

If $t' \leq t$ and $B \leq B'$, then $\mathcal{V}(E, t, B)$ implies $\mathcal{V}(E, t', B')$. The twist range shrinks, all summands are nonnegative, and the budget increases. This simple observation allows one broad cohomological estimate to be used by constructions requiring fewer twists.

For any disjoint finite sets of primes $P_1, \dots, P_m$, the sum of their weighted contributions is bounded by B, because their union is finite. It is therefore generally incorrect to spend the full budget B independently on many disjoint subsets and then claim that no factor of m is lost. The original uniform finite-set inequality is stronger than separate singleton bounds and should be retained until the final sum.

If M_ℓ is replaced by a subquotient of a finite-length module N_ℓ, then $\operatorname{length} M_\ell \le \operatorname{length} N_\ell$ by length additivity. Consequently any collection of genuine subquotient comparisons transfers a weighted length bound. The existence of the comparison maps and the subquotient property must be proved before this numerical monotonicity can be used.

18.7 A localization diagram and its kernel estimate

Consider a commutative diagram of modules with exact rows

$$\begin{array}{ccccccccc} 0 & \to & A' & \to & A & \to & A'' & \to & 0 \\ & & \downarrow f' & & \downarrow f & & \downarrow f'' & & \\ 0 & \to & B' & \to & B & \to & B'' & \to & 0. \end{array}$$

There is an exact sequence $0 \to \ker f' \to \ker f \to \ker f''$. Indeed, an element of $\ker f$ maps to $\ker f''$ by commutativity; if its image in A'' is zero, it comes from A', and injectivity of $B' \to B$ shows that the corresponding element lies in $\ker f'$. If the two outer kernels have finite length, it follows that

$$\operatorname{length} \ker f \le \operatorname{length} \ker f' + \operatorname{length} \ker f''.$$

This is a common elementary mechanism for bounding a localization kernel after filtering its coefficient module. A decomposition in the Grothendieck group alone would not suffice; the exact sequences are what justify the inequality for kernels.

19 Cohomological exact sequences and local conditions

A cohomological bound becomes useful only after the group being bounded has been related to a concrete obstruction. In degree one, this relation can often be described without higher derived functors. We develop the cocycle calculations explicitly, including their compatibility with restriction to local subgroups. These calculations explain both the appearance of local kernels and the care needed when passing between finite coefficients and divisible coefficients.

19.1 Continuous cocycles and extensions

Let G be a profinite group and M a discrete abelian group with a continuous G-action. A continuous map $c : G \to M$ is a cocycle when

$$c(gh) = c(g) + gc(h) \qquad (g, h \in G).$$

Putting $g = h = 1$ gives $c(1) = 0$, and substituting $h = g^{-1}$ gives $c(g^{-1}) = -g^{-1}c(g)$. For $m \in M$, the function $c_m(g) = gm - m$ is a cocycle. Quotienting the additive group of cocycles by these coboundaries defines $H^1(G, M)$.

The same formula describes affine actions. Define $g * x = gx + c(g)$ on the underlying set of M. Associativity of this action is exactly the cocycle identity. Replacing the origin by m changes c by

the coboundary $gm - m$. Consequently $H^1(G, M)$ classifies such affine actions up to translation of the origin. The class is zero precisely when the affine action has a fixed point: the equation $gx + c(g) = x$ is equivalent to $c(g) = g(-x) - (-x)$.

This interpretation is helpful for local conditions. A cohomology class can be globally nonzero while every chosen local subgroup has a fixed point for the corresponding affine action. The local fixed points need not be the restrictions of a single global fixed point. The kernel of global-to-local restriction measures this failure of compatibility.

19.2 The connecting homomorphism

Consider an exact sequence of discrete continuous G-modules

$$0 \longrightarrow A \xrightarrow{i} B \xrightarrow{p} C \longrightarrow 0. \tag{77}$$

If $c \in C^G$, choose $b \in B$ with $p(b) = c$. Since c is invariant, $gb - b$ lies in $i(A)$. Identify A with its image in B and define

$$\delta(c)(g) = gb - b.$$

This is a continuous cocycle. A different lift $b + a$ changes it by the coboundary $ga - a$, so its cohomology class is independent of the lift. Addition of lifts shows that $\delta : C^G \to H^1(G, A)$ is a homomorphism.

Proposition 19.1 (Exact sequence through degree one)**.** *The sequence*

$$0 \longrightarrow A^G \longrightarrow B^G \longrightarrow C^G \xrightarrow{\delta} H^1(G, A) \longrightarrow H^1(G, B) \longrightarrow H^1(G, C)$$

is exact at every displayed term except the final one, at which no further map has been specified.

Proof. Exactness at A^G and B^G follows directly from (77). An invariant lift of $c \in C^G$ gives the zero cocycle. Conversely, if $\delta(c)$ is a coboundary $ga - a$, then $b - a$ is an invariant lift of c. This proves exactness at C^G.

The image of δ becomes a coboundary in B. Conversely, let an A-valued cocycle u become the coboundary $gb - b$ in B. Then $p(b)$ is invariant and u represents $\delta(p(b))$. This proves exactness at $H^1(G, A)$.

Finally, let a B-valued cocycle v have trivial image in $H^1(G, C)$. Then $p(v(g)) = gc - c$ for some $c \in C$. Choose a lift b of c. The cocycle $v(g) - (gb - b)$ takes values in A and represents the same class as v in $H^1(G, B)$. The reverse inclusion is immediate. □

No choice of lifts is claimed to be canonical. The cohomology class is canonical because changes of lift are precisely the coboundaries that are factored out. This is the recurring mechanism by which a division problem produces a cohomological obstruction.

19.3 Division and torsion coefficients

Let M be a discrete continuous G-module for which multiplication by ℓ^n is surjective. The exact sequence

$$0 \longrightarrow M[\ell^n] \longrightarrow M \xrightarrow{\ell^n} M \longrightarrow 0$$

gives

$$0 \longrightarrow M^G/\ell^n M^G \longrightarrow H^1(G, M[\ell^n]) \longrightarrow H^1(G, M)[\ell^n] \longrightarrow 0. \tag{78}$$

To verify the right-hand term, use exactness: the image in $H^1(G, M)$ is the kernel of multiplication by ℓ^n. The left-hand kernel is the image of the connecting map, whose kernel on M^G is $\ell^n M^G$.

For the trivial action on $M = \mathbb{Q}_\ell/\mathbb{Z}_\ell$, the first term vanishes because $M^G = M$ is divisible. For a nontrivial twist, M remains divisible as an abstract group, but M^G need not be divisible. Hence the first term need not vanish. Omitting it can change the size of a finite-coefficient obstruction group.

The connecting cocycle has a direct meaning. Given $m \in M^G$, choose b with $\ell^n b = m$. Then $gb - b$ is killed by ℓ^n and records the failure of the chosen division point b to be fixed by G. An invariant division point exists exactly when this cohomology class is zero. This is the elementary prototype of the obstruction measured by a division kernel in an arithmetic construction.

19.4 Changing the set of local conditions

Let $\{G_v\}_{v \in S}$ be a finite or infinite collection of closed subgroups of G. Define

$$H^1_S(G, M) = \ker\left(H^1(G, M) \longrightarrow \prod_{v \in S} H^1(G_v, M) \right).$$

If $S \subset S'$, then $H^1_{S'}(G, M) \subset H^1_S(G, M)$. Moreover restriction induces an injection

$$H^1_S(G, M)/H^1_{S'}(G, M) \hookrightarrow \prod_{v \in S' \setminus S} H^1(G_v, M). \tag{79}$$

The kernel of this restriction map consists exactly of the classes already satisfying all conditions indexed by S'.

If the quotient and the local groups have finite $\mathbb{Z}_\ell$-length, then

$$\operatorname{length} H^1_S(G, M) - \operatorname{length} H^1_{S'}(G, M) \leq \sum_{v \in S' \setminus S} \operatorname{length} H^1(G_v, M)$$

when the right-hand set is finite. For an infinite set, one must either know that only finitely many factors contribute or use a finite subset at each stage. The product of infinitely many finite groups may have infinite length. A finite-support assertion is therefore substantive, not a consequence of the finiteness of each individual local group.

More general local conditions choose subgroups $L_v \subset H^1(G_v, M)$ and take the inverse image of $\prod_v L_v$. The condition used above is $L_v = 0$. If $L_v \subset L'_v$, enlarging the local conditions enlarges the global group, and its quotient injects into $\prod_v L'_v/L_v$ by the same kernel calculation.

19.5 A diagram lemma for local kernels

The following elementary result gives the length inequality needed in many comparisons. Consider a commutative diagram of module homomorphisms

$$\begin{array}{ccc} A & \xrightarrow{r} & B \\ \downarrow f & & \downarrow g \\ A' & \xrightarrow{r'} & B'. \end{array}$$

The induced map $\ker r \to \ker r'$ has kernel contained in $\ker f$. Its image is a submodule of $\ker r'$. Therefore, if the relevant modules have finite length,

$$\operatorname{length} \ker r \leq \operatorname{length} \ker f + \operatorname{length} \ker r'. \tag{80}$$

This estimate does not use surjectivity of either vertical arrow. It is often enough when $\ker f$ is a controlled coefficient-change term and $\ker r'$ is a cohomological group for which a global bound is available.

If f is injective, the estimate improves to an actual injection $\ker r \hookrightarrow \ker r'$. If f is surjective, there may still be a failure of surjectivity on kernels. Indeed, a lift a of $a' \in \ker r'$ satisfies only $r(a) \in \ker g$. One must adjust a by an element of $\ker f$ to kill $r(a)$. The obstruction lies in the quotient

$$\ker g / r(\ker f).$$

This description is a concrete portion of the snake-lemma mechanism. It identifies the extra term that would be missed by assuming that surjective coefficient maps automatically give surjective maps between local kernels.

19.6 Inflation and restriction

Let N be a closed normal subgroup of G, and put $Q = G/N$. The invariant subgroup M^N is a Q-module. A cocycle on Q with values in M^N can be pulled back to G; this is inflation. Restriction to N gives a map $H^1(G, M) \to H^1(N, M)$. The quotient group acts on $H^1(N, M)$ by conjugation and transport of coefficients.

Proposition 19.2 (Exactness of inflation and restriction)**.** *There is an exact sequence*

$$0 \longrightarrow H^1(Q, M^N) \xrightarrow{\inf} H^1(G, M) \xrightarrow{\mathrm{res}} H^1(N, M)^Q.$$

No surjectivity onto the final term is asserted.

Proof. Suppose an inflated cocycle is the coboundary $gm - m$. On N the inflated cocycle vanishes, so $m \in M^N$. It is therefore already a coboundary on Q. Inflation is injective.

Now suppose the restriction of a cocycle c to N is a coboundary. Subtract that coboundary, so that $c(n) = 0$ for $n \in N$. Then $c(gn) = c(g)$, so c is constant on right cosets of N. Also

$$c(ng) = nc(g), \qquad c(ng) = c(g(g^{-1}ng)) = c(g).$$

Thus $c(g) \in M^N$. The resulting map on Q is continuous by the quotient topology and is a cocycle. This proves exactness at $H^1(G, M)$.

For the last assertion, conjugating a restricted cocycle by an element of G changes it by a coboundary. This follows from the identity

$$gc(g^{-1}ng) = c(n) + (n-1)c(g), \qquad n \in N,$$

which is obtained by expanding the cocycle identity twice. Hence its class is fixed by Q. □

The normality assumption is used twice: it gives the quotient group and makes conjugation preserve N. For a local decomposition subgroup, conjugation usually changes the chosen subgroup rather than preserving it. One then compares the corresponding cohomology groups by a transport isomorphism. The vanishing of a restricted class is independent of the choice of conjugate subgroup.

19.7 Finite groups of order prime to the coefficient prime

Let G be finite of order n, and suppose multiplication by n is an automorphism of M. Then

$$H^1(G, M) = 0. \tag{81}$$

Here is an explicit proof. For a cocycle c, let $S = \sum_{g \in G} c(g)$. For $h \in G$, reindex the sum by $g \mapsto hg$:

$$S = \sum_{g} c(hg) = nc(h) + hS.$$

Thus $c(h) = h(-n^{-1}S) - (-n^{-1}S)$, a coboundary.

For an ℓ-primary torsion module, multiplication by n is an automorphism whenever $\ell \nmid n$. Therefore a finite group of order prime to ℓ contributes no degree-one cohomology with such coefficients. The assertion fails when $\ell \mid n$. For example, with trivial action of a cyclic group of order ℓ on $\mathbb{Z}/\ell\mathbb{Z}$, every homomorphism is a cocycle and every coboundary is zero. The cohomology group is then $\mathbb{Z}/\ell\mathbb{Z}$.

The proof also shows why rational averaging and integral averaging behave differently. In characteristic zero one may divide by n in a vector space. In a lattice or a torsion module that division is possible only when n acts invertibly. The primes dividing the group order are exactly where an averaging projector may introduce denominators.

19.8 Cyclic cohomology by a single equation

Suppose $G = \langle \sigma \rangle$ is cyclic of order n. A cocycle is determined by $a = c(\sigma)$, since

$$c(\sigma^r) = a + \sigma a + \cdots + \sigma^{r-1}a.$$

The relation $\sigma^n = 1$ requires $Na = 0$, where $N = 1 + \sigma + \cdots + \sigma^{n-1}$. Conversely every $a \in \ker N$ defines a cocycle by this formula. A coboundary has $a = (\sigma - 1)m$. Hence

$$H^1(G, M) = \ker N / (\sigma - 1)M. \tag{82}$$

The inclusion $(\sigma - 1)M \subset \ker N$ follows from $N(\sigma - 1) = \sigma^n - 1 = 0$.

For trivial action, $N = n$ and the denominator is zero, so $H^1(G, M) = M[n]$. For an action of order two with $\sigma = -1$, $N = 0$ and $\sigma - 1 = -2$, giving $H^1(G, M) = M/2M$. These two examples have different answers even though the underlying group has the same order. Thus the coefficient action, and in particular a Tate twist, cannot be suppressed in a cohomological estimate.

19.9 Unramified classes

Let G_v be a local Galois group with inertia subgroup I_v. The unramified subgroup is

$$H^1_{\mathrm{ur}}(G_v, M) = \ker\bigl(H^1(G_v, M) \to H^1(I_v, M)\bigr).$$

Inflation–restriction identifies it with $H^1(G_v/I_v, M^{I_v})$. When a finite quotient of the procyclic group G_v/I_v suffices to describe both the action and a particular cocycle, formula (82) computes that class in terms of Frobenius and the corresponding norm operator.

A familiar quotient $M^{I_v}/(F - 1)M^{I_v}$ occurs in cohomology of an infinite cyclic group. For a procyclic group and discrete coefficients, continuity imposes an additional finite-level condition. One

should not replace the continuous cohomology by the infinite-cyclic calculation without checking that condition. If the module is finite, a continuous cocycle factors through a finite quotient after shrinking an open subgroup that acts trivially and lies in its zero set. The norm relation at that quotient is then part of the calculation.

For a finite ℓ-primary module with trivial Frobenius action, continuous homomorphisms from $\widehat{\mathbb{Z}}$ are determined by the image of 1, and every element occurs. Thus in this case the unramified cohomology is the module itself. This example illustrates that unramified does not mean zero: it means that the class is already visible in the Frobenius quotient.

19.10 Finite images of continuous cocycles

Because a profinite group is compact and a discrete space has no infinite compact subsets, every continuous cocycle $c : G \to M$ has finite image. The subgroup generated by that image need not be finite for an arbitrary discrete module. For an ℓ-primary torsion module, however, a subgroup generated by finitely many torsion elements is finite as an abelian group. Its G-stable closure is also finite when the action on those finitely many elements has finite orbits, as continuity guarantees.

Consequently each continuous cocycle into an ℓ-primary discrete module takes values in a finite G-stable submodule. Likewise, a coboundary witnessing the equality of two classes uses a single element and its finite orbit. It follows that

$$H^1(G, M) = \varinjlim_{M_0} H^1(G, M_0),$$

where M_0 ranges over the finite G-stable submodules of M. For clarity, surjectivity follows by putting the finite image of a cocycle inside such an M_0; injectivity of the colimit map follows by also adjoining the element that witnesses a coboundary.

This is a statement about the direct limit of cohomology groups, not about equality with one finite stage. A group can be a union of finite subgroups and still have infinite length. The module $\mathbb{Q}_\ell/\mathbb{Z}_\ell$ itself is the basic example: its subgroup of elements killed by ℓ^n has length n, and these lengths are unbounded. Uniform length bounds are exactly what prevent this phenomenon.

19.11 Finite length from bounded finite stages

Let A be an ℓ-primary abelian group that is the union of an increasing family of finite subgroups A_i. If $\operatorname{length}_{\mathbb{Z}_\ell} A_i \le B$ for all i, then A is finite and has length at most $\lfloor B \rfloor$. Indeed, the cardinalities $|A_i| = \ell^{\operatorname{length} A_i}$ are positive integers bounded by $\ell^{\lfloor B \rfloor}$. Choose a subgroup of maximal cardinality. Directedness implies that every other A_i is contained in it, because a common upper stage cannot have larger cardinality. Thus the union equals that stage.

The same argument works for any directed union, not only a sequence. It provides a concrete justification for promoting bounded finite-level obstructions to a finite divisible-coefficient obstruction group. The bound must apply to all stages of the same directed system. Bounds for unrelated finite groups do not establish a statement about their union.

19.12 Twists and the range of uniformity

Let $\kappa_\ell : G \to \mathbb{Z}_\ell^\times$ be a continuous character. For an integer j, write $M(j) = \mathbb{Q}_\ell/\mathbb{Z}_\ell$ with action $g \cdot m = \kappa_\ell(g)^j m$. As an abstract abelian group $M(j)$ is independent of j. Its invariant subgroup,

local cohomology, and restriction kernel depend on j through the action. The sign convention for j therefore matters.

A bound of the form

$$\sum_{\ell} \log \ell \sum_{1 \le j < t} \operatorname{length} A_{\ell,j} \le B$$

contains three separate uniformities: it controls the coefficient prime, the twist, and any varying field implicit in $A_{\ell,j}$. Fixing two of these parameters does not settle the third. In particular, statements that each $A_{\ell,j}$ is finite provide no numerical bound on the sum. The arithmetic window hypothesis supplies a bound for the entire sum at the chosen scale. The algebra developed here explains how such a bound can be transported to obstruction modules once suitable comparison maps have been constructed.

20 Cyclotomic subfields and arithmetic scales

20.1 A common-kernel construction

Fix a primitive Dirichlet character ψ of conductor f and order r. Let $a \ge 0$ be an integer and q, d distinct odd primes. Choose N divisible by $2^{a+2}, f, q, d$. The Galois group of $K = \mathbb{Q}(\zeta_N)$ is identified with $(\mathbb{Z}/N\mathbb{Z})^\times$ through $\sigma_b(\zeta_N) = \zeta_N^b$.

Consider the homomorphism

$$\varphi : (\mathbb{Z}/N\mathbb{Z})^\times \longrightarrow ((\mathbb{Z}/2^{a+2}\mathbb{Z})^\times/\{\pm 1\}) \times \operatorname{im}\psi \times \{\pm 1\} \times \{\pm 1\} \tag{83}$$

whose components are reduction modulo 2^{a+2} up to sign, the value of ψ, and the two Legendre symbols modulo q and d. Every component is well defined because its modulus divides N. Let $H = \ker \varphi$ and $E = K^H$. Then H consists exactly of those residue classes satisfying

$$b \equiv \pm 1 \pmod{2^{a+2}}, \qquad \psi(b) = 1, \qquad \left(\frac{b}{q}\right) = \left(\frac{b}{d}\right) = 1. \tag{84}$$

The sign in the first congruence is a single alternative, not an independent choice at each prime-power factor.

Proposition 20.1 (Degree bound)**.** *The extension $E/\mathbb{Q}$ is abelian Galois and*

$$[E : \mathbb{Q}] = |\operatorname{im}\varphi| \le 4r\,2^a.$$

Proof. Every subgroup of the abelian group $\mathrm{Gal}(K/\mathbb{Q})$ is normal, so the fixed field is Galois with group $\mathrm{Gal}(K/\mathbb{Q})/H$. The first isomorphism theorem identifies this quotient with $\operatorname{im}\varphi$. The unit group modulo 2^{a+2} has order 2^{a+1}, and its subgroup $\{\pm 1\}$ has order 2, also when $a = 0$. Thus the first target factor has order 2^a. The remaining target factors have orders $r, 2, 2$, giving the upper bound. □

Equality would require the components of (83) to vary independently, that is, surjectivity onto the displayed product. It is not a consequence of the common-kernel definition. For example, a quadratic component of ψ may coincide with one of the other quadratic characters. The bound deliberately avoids any independence assumption.

20.2 The compositum interpretation

Let H_i be the kernel of the ith component of φ and $E_i = K^{H_i}$. Then

$$H = \bigcap_i H_i, \qquad E = E_1E_2E_3E_4$$

inside K. Indeed, the subgroup fixing the compositum is precisely the intersection of the subgroups fixing its individual fields. This gives an equivalent description in terms of fields cut out by four separate character conditions.

The compositum description helps explain possible degree losses. Overlaps among the fields E_i create relations among the characters and reduce the degree of their compositum. The product of their degrees is always an upper bound, while an equality requires additional linear-disjointness information. The common-kernel construction keeps those overlaps automatically, without choosing a decomposition of the Galois group.

20.3 Complex conjugation and the CM property

Suppose now that $q \equiv 3 \pmod 8$ and $d \equiv 7 \pmod{16}$. Both primes are 3 modulo 4, so $(-1/q) = (-1/d) = -1$. Complex conjugation in K corresponds to the residue class -1. Its image under (83) has last two coordinates $(-1,-1)$ and is therefore nontrivial.

Proposition 20.2. *Under these congruence conditions, E is a CM field. Its maximal real subfield E^+ has index 2 in E.*

Proof. The restriction c of complex conjugation to E is a nontrivial involution by the preceding calculation. Its fixed field E^+ is therefore of index 2. Since $E/\mathbb{Q}$ is abelian Galois, c commutes with every embedding of E into $\mathbb{C}$. If $x \in E^+$ and τ is an embedding, then $\overline{\tau(x)} = \tau(c(x)) = \tau(x)$, so every embedding of E^+ is real.

There is no real embedding of E: such an embedding would satisfy $\tau(c(x)) = \overline{\tau(x)} = \tau(x)$ for every x, and its injectivity would force c to be trivial. Thus E is a totally imaginary quadratic extension of a totally real field, as required. □

The proof uses the abelian Galois nature of $E/\mathbb{Q}$ to transport complex conjugation uniformly among embeddings. A general nonreal number field need not be CM. Here the character construction supplies the necessary symmetry.

20.4 Neighboring powers of two

For $\varepsilon > 0$ and $x \ge 1$, one may choose an integer $a(x)$ with

$$x^{2+8\varepsilon} \le 2^{a(x)} \le 2x^{2+8\varepsilon}. \tag{85}$$

For instance, take $a(x) = \lceil (2+8\varepsilon)\log x/\log 2 \rceil$. The ceiling inequalities give the lower bound and the strict upper bound unless the real number is an integer, in which case the upper bound still holds.

Combining (85) with Proposition 20.1 gives

$$[E:\mathbb{Q}] \le 8r\, x^{2+8\varepsilon}.$$

This is a degree upper bound. It does not assert that the degree is asymptotic to a constant times that power of x. The selected integer $a(x)$ can remain constant on intervals and jump at discrete scales; the associated field can change when it jumps.

Taking logarithms in (85) yields

$$(2+8\varepsilon)\log x \le a(x)\log 2 \le (2+8\varepsilon)\log x + \log 2.$$

Thus the parameter controlling the power of two grows only logarithmically with x, even though the degree upper bound grows polynomially.

20.5 The auxiliary prime conditions

Fix $0 < \varepsilon < 1/52$ and $T > 1$. The prime pair condition asks for primes q, d with

$$\begin{aligned} q &\equiv 3 \pmod 8, & d &\equiv 7 \pmod{16}, \\ \tfrac{1}{2}T^{\varepsilon/5} &< \log q \le T^{\varepsilon/5}, \\ \tfrac{1}{2}T^{\varepsilon/5}(\log T)^2 &< \log d \le T^{\varepsilon/5}(\log T)^2, \end{aligned} \tag{86}$$

and with

$$q \mid d+1, \qquad p \mid d+1 \quad \text{for every odd prime } p \le T^{\varepsilon/5}.$$

Call such a pair admissible at (ε, T). The bounds are on logarithms. Thus q and d themselves lie on exponential scales, with d substantially larger in the logarithmic comparison because of the factor $(\log T)^2$.

For a fixed admissible choice of q, the congruence requirements on d are algebraically compatible. Let m be the squarefree product of q and all odd primes up to $T^{\varepsilon/5}$, taking a repeated prime only once. The conditions are

$$d \equiv -1 \pmod m, \qquad d \equiv 7 \pmod{16}.$$

The Chinese remainder theorem gives a unique class modulo $16m$, and that class is coprime to $16m$. This proves compatibility of the congruences, but does not place a prime in the required interval. The modulus and interval both vary with T. Their quantitative interaction is an additional prime-existence problem.

20.6 Why common auxiliary primes matter

Suppose a construction must work for every $x \in [Y, Y^2]$ using a single pair (q, d). Proving for each x that there exists some pair (q_x, d_x) is weaker. A cohomological bound attached to the first pair cannot simply be transferred to a field attached to the second pair. The field is defined by its character kernel, so changing an auxiliary prime changes that kernel and usually changes the field.

The admissible prime condition is consequently imposed at a common starting scale, while $a(x)$ is allowed to vary across the window. This preserves enough arithmetic structure to compare the family along the interval. The universal quantifier over associated fields can then handle the variation caused by the power-of-two parameter without changing the chosen prime pair.

21 Finite unitary geometry and congruence-level counts

At finite places, dimension and volume comparisons involve the sizes of matrix groups and their level subgroups. The relevant estimates can be made uniform in the rank by retaining their exact product formulas. We develop general linear and unitary examples and then pass to congruence quotients over a discrete valuation ring.

21.1 General linear groups

An invertible $n \times n$ matrix over $\mathbb{F}_q$ is an ordered basis of $\mathbb{F}_q^n$. Its first column has $q^n - 1$ choices. After j independent columns have been chosen, their span has q^j elements, so the next column has $q^n - q^j$ choices. Therefore

$$|\operatorname{GL}_n(\mathbb{F}_q)| = \prod_{j=0}^{n-1} (q^n - q^j) = q^{n^2} \prod_{j=1}^{n} (1 - q^{-j}). \tag{87}$$

The same argument counts ordered independent r-tuples in an n-dimensional space as $\prod_{j=0}^{r-1}(q^n - q^j)$. Dividing by the number of ordered bases of an r-space gives the Gaussian binomial coefficient

$$\begin{bmatrix} n \\ r \end{bmatrix}_q = \prod_{j=0}^{r-1} \frac{q^n - q^j}{q^r - q^j},$$

the number of r-dimensional subspaces.

The dimension of the Grassmannian suggested by this formula is $r(n-r)$: extracting the powers of q gives

$$\begin{bmatrix} n \\ r \end{bmatrix}_q = q^{r(n-r)} \prod_{j=0}^{r-1} \frac{1 - q^{j-n}}{1 - q^{j-r}}.$$

The remaining product is bounded by constants independent of n, r when $q \geq 2$. A proof of the required product bounds is given below. Thus the exact finite count has the expected leading power with a uniform multiplicative correction.

21.2 Matrices of prescribed rank

A rank-r linear map $\mathbb{F}_q^n \to \mathbb{F}_q^m$ can be factored as a surjection onto an abstract r-space followed by an injection into $\mathbb{F}_q^m$. Changing the basis of the intermediate space changes both factors without changing the composite, and this is the only ambiguity. Hence the number of rank-r matrices is

$$\frac{\prod_{j=0}^{r-1}(q^m - q^j)(q^n - q^j)}{\prod_{j=0}^{r-1}(q^r - q^j)}. \tag{88}$$

For $r = 0$, the empty product gives the single zero matrix. For $r = \min(m, n)$, division by q^{mn} gives the probability of full possible rank for a uniformly chosen matrix.

Such a probability is a statement about a uniform finite sample. It does not prove that a geometrically defined family of matrices behaves randomly. To use the count in a construction, one needs a map to the finite matrix space and a proof that its image meets the desired rank locus, or an estimate showing that it is sufficiently large.

21.3 Norms in a quadratic finite field

Let $k = \mathbb{F}_q$ and $k_2 = \mathbb{F}_{q^2}$, with involution $\overline{z} = z^q$. The norm is $N(z) = z^{q+1}$. The multiplicative group of a finite field is cyclic. One elementary proof is to take the least common multiple m of all element orders. For each prime power dividing m, an element realizes that prime-power part; multiplying the corresponding commuting elements gives an element of order m. Every group element is a root of $X^m - 1$, so the field root bound gives $|k_2^\times| \le m$. The reverse inequality is Lagrange's theorem, hence equality and cyclicity.

It follows that $N : k_2^\times \to k^\times$ is surjective and its kernel has order $q+1$. Every nonzero norm value has exactly $q+1$ preimages, while zero has the single preimage zero. The trace $z \mapsto z+\overline{z}$ is a nonzero k-linear map from a two-dimensional space to a one-dimensional space, hence is also surjective. To see it is nonzero in characteristic two, choose $z \notin k$; then $z + z^q \neq 0$. In odd characteristic the trace of one is already nonzero.

21.4 Hermitian forms over finite fields

A Hermitian form is linear in the first variable and satisfies $h(y,x) = \overline{h(x,y)}$. A nondegenerate Hermitian space over k_2 has an orthonormal basis. First there is a vector v with $h(v,v) \neq 0$. Otherwise, expanding $h(x+ay, x+ay)$ for all $a \in k_2$ and using the nondegeneracy of the trace pairing would force $h(x,y) = 0$ for all x, y. This contradicts nondegeneracy. Since $h(v,v) \in k^\times$ and the norm map is surjective, rescale v to make its norm one. Its orthogonal complement is nondegenerate, and induction finishes the construction.

Consequently all nondegenerate Hermitian forms of a given dimension are isometric over k_2. We may use the standard form $h(x,y) = \sum_j x_j \overline{y_j}$ and write $U(n,q)$ for its group of isometries. This finite-field uniqueness differs from the signature classification over $\mathbb{C}$: over finite fields there is no ordering that distinguishes positive from negative values.

21.5 Counting vectors of a given norm

Choose a nontrivial additive character $\psi : k \to \mathbb{C}^\times$. For $u \neq 0$,

$$\sum_{z \in k_2} \psi(uN(z)) = 1 + (q+1) \sum_{a \in k^\times} \psi(ua) = -q.$$

For $u = 0$, the sum is q^2. Let $N_n(a)$ count vectors in k_2^n of norm a. Additive-character orthogonality gives

$$N_n(a) = \frac{1}{q} \sum_{u \in k} \psi(-ua) \left(\sum_{z \in k_2} \psi(uN(z)) \right)^n.$$

If $a \neq 0$, the sum of $\psi(-ua)$ over $u \neq 0$ is -1; if $a = 0$, it is $q-1$. Therefore

$$N_n(a) = q^{n-1}(q^n - (-1)^n), \qquad a \neq 0, \tag{89}$$

$$N_n(0) = q^{2n-1} + (-1)^n (q-1) q^{n-1}. \tag{90}$$

The calculation is valid in every residue characteristic, including two. The sign $(-1)^n$ is an ordinary integer sign in the counting formula, not a field element whose distinction disappears in characteristic two.

21.6 The order of the finite unitary group

An element of $U(n,q)$ is an ordered orthonormal basis. There are $N_n(1)$ choices for the first vector, and its orthogonal complement is a nondegenerate Hermitian space of dimension $n-1$. Thus

$$|U(n,q)| = N_n(1)|U(n-1,q)|.$$

Starting with the zero-dimensional group of order one yields

$$|U(n,q)| = q^{n(n-1)/2}\prod_{j=1}^{n}(q^j-(-1)^j) = q^{n^2}\prod_{j=1}^{n}(1-(-1)^j q^{-j}). \tag{91}$$

For $n=1$ this is $q+1$, the norm-one subgroup of $k_2^\times$. For $n=2$ it is $q(q+1)(q^2-1)$.

The exponent n^2 in the normalized formula agrees with the dimension of the unitary algebraic group over k. Although the matrices have entries in a quadratic extension and therefore $2n^2$ scalar coordinates over k, the Hermitian preservation equation imposes n^2 independent conditions at a smooth point.

21.7 Isotropic lines and hyperbolic planes

A nonzero isotropic vector spans a line with q^2-1 nonzero vectors. Thus the number of isotropic lines is

$$I_n(q) = \frac{(q^n-(-1)^n)(q^{n-1}-(-1)^{n-1})}{q^2-1}. \tag{92}$$

This follows by subtracting the zero vector in (90) and factoring the numerator. For $n \ge 2$ the count is positive, so isotropic lines exist.

Given an isotropic vector v, choose w with $h(v,w)=1$. Replacing w by $w-av$ changes $h(w,w)$ by $-a-\overline{a}$. Surjectivity of trace lets us choose a so that the new w is also isotropic. The span of v,w then has Gram matrix $\begin{pmatrix}0&1\\1&0\end{pmatrix}$ and is nondegenerate. It splits off orthogonally as a hyperbolic plane. Repeating proves that the maximal dimension of a totally isotropic subspace is $\lfloor n/2 \rfloor$.

For a totally isotropic subspace U, the quotient $U^\perp/U$ is nondegenerate of dimension $n-2\dim U$. Nondegeneracy follows because the radical of $U^\perp$ is exactly U, using the nondegeneracy of the ambient form and the dimension formula for orthogonal complements.

21.8 Counting isotropic flags

An isotropic flag of length r is a chain $0 \subset U_1 \subset \cdots \subset U_r$ with $\dim U_j = j$. Choose U_1, then pass to $U_1^\perp/U_1$ and continue. The number of such flags is

$$\prod_{j=0}^{r-1} I_{n-2j}(q), \qquad 0 \le r \le \lfloor n/2 \rfloor.$$

Every r-dimensional k_2-space has $\prod_{j=1}^{r}(q^{2j}-1)/(q^2-1)$ complete flags, by the elementary subspace count. Dividing yields the number of totally isotropic r-subspaces:

$$\frac{\displaystyle\prod_{k=n-2r+1}^{n}(q^k-(-1)^k)}{\displaystyle\prod_{j=1}^{r}(q^{2j}-1)}. \tag{93}$$

The formula is a count of subspaces, not of ordered isotropic frames. The distinction changes the denominator and can create large powers of q if omitted.

These counts also give indices of stabilizers. The unitary group acts transitively on isotropic subspaces of a given dimension: extend a basis of each to hyperbolic pairs and then to an orthonormal basis of the remaining complement. The resulting isometry carries one subspace to the other. Orbit–stabilizer therefore identifies (93) with the index of the corresponding parabolic stabilizer.

21.9 Uniform bounds for the product factors

For $q \geq 2$,

$$\prod_{j \geq 1} (1 + q^{-j}) \leq \exp\left(\sum_{j \geq 1} q^{-j} \right) = e^{1/(q-1)} \leq e.$$

Also $\log(1 - x) \geq -x/(1 - x)$ for $0 \leq x < 1$, as follows by integrating $1/(1 - t) \leq 1/(1 - x)$ on $[0, x]$. Therefore

$$\prod_{j \geq 1} (1 - q^{-j}) \geq \exp\left(-\frac{q}{(q-1)^2} \right) \geq e^{-2}.$$

Every partial product is bounded between these constants. Equations (87) and (91) imply

$$e^{-2} q^{n^2} \leq |\operatorname{GL}_n(\mathbb{F}_q)| \leq q^{n^2}, \qquad e^{-2} q^{n^2} \leq |U(n, q)| \leq e q^{n^2}.$$

The constants are uniform in both rank and field size. Consequently

$$\left| \log |U(n, q)| - n^2 \log q \right| \leq 2.$$

Keeping this bounded correction is substantially sharper than bounding each of the n factors separately by a fixed constant and obtaining an artificial error proportional to n.

21.10 General linear congruence quotients

Let R be a complete discrete valuation ring with residue field $\mathbb{F}_q$ and uniformizer ϖ. Reduction gives a surjection

$$\operatorname{GL}_n(R/\varpi^{m+1}) \longrightarrow \operatorname{GL}_n(R/\varpi^m).$$

Any lift of an invertible matrix remains invertible because its determinant is a unit modulo ϖ. The kernel consists of $I + \varpi^m X$ with $X \in M_n(\mathbb{F}_q)$ and therefore has q^{n^2} elements. Induction gives

$$|\operatorname{GL}_n(R/\varpi^m)| = q^{n^2(m-1)} |\operatorname{GL}_n(\mathbb{F}_q)|. \tag{94}$$

At consecutive levels the group law in the kernel is additive, because the product term $\varpi^{2m} XY$ vanishes modulo ϖ^{m+1} for $m \geq 1$.

21.11 Unitary congruence quotients

Let R_2/R be an unramified quadratic extension of complete discrete valuation rings, with involution reducing to $z \mapsto z^q$ on the residue field. Choose a uniformizer $\varpi \in R$ and a unimodular Hermitian matrix H. Suppose a matrix g preserves H modulo ϖ^m. A lift to the next level has an error

$$g^* H g - H = \varpi^m E \pmod{\varpi^{m+1}},$$

where E is Hermitian over the residue field. Replacing g by $g(I+\varpi^m X)$ changes E to $E+X^*H+HX$.

The map $X \mapsto X^*H + HX$ onto Hermitian matrices is surjective over the residue field. After a change of basis take $H = I$. For off-diagonal entries, choose one of each pair of entries freely and solve for the other. On the diagonal, solve a trace equation, which is possible even in residue characteristic two. The domain has dimension $2n^2$ over $\mathbb{F}_q$ and the Hermitian target has dimension n^2, so the kernel has dimension n^2. Thus reduction on unitary groups is surjective and every fiber has q^{n^2} elements. Consequently

$$|U(H, R_2/\varpi^m)| = q^{n^2(m-1)}|U(n, q)|. \tag{95}$$

The unramified assumption is used in the residue-field involution and the trace argument. A ramified quadratic extension requires a separate analysis and need not have this same simple lifting formula.

21.12 Level indices and logarithmic budgets

For a compact integral unitary group, let $K(m)$ be the kernel of reduction modulo ϖ^m. Completeness and the successive lifting argument identify the finite quotient with the group counted in (95). Hence

$$[K(0) : K(m)] = q^{n^2 m} \prod_{j=1}^{n} (1 - (-1)^j q^{-j}),$$

where $K(0)$ denotes the full integral group. Taking logarithms,

$$\log[K(0) : K(m)] = n^2 m \log q + O(1),$$

with the absolute error bounded by two under the stated hypotheses. Across a finite set of places these logarithmic indices add.

This is the finite-place analogue of the determinant calculations for Hermitian lattices. A global comparison can combine the exact leading term with local corrections, but the number of contributing places must still be controlled. A bounded error at each of an increasing number of places becomes an error proportional to that number. The weighted cohomological and level budgets are designed to keep such accumulated local contributions visible.

22 Representations, integral structures, and compatibility data

22.1 Continuous representations and stable lattices

Let K be a finite extension of $\mathbb{Q}_\ell$, with ring of integers $\mathcal{O}$, uniformizer ϖ, and residue field k. A continuous K-linear representation of a profinite group G on a finite-dimensional space V is a continuous homomorphism $\rho : G \to \mathrm{GL}(V)$. A lattice is a finitely generated $\mathcal{O}$-submodule of V that spans V over K and is free of rank $\dim_K V$.

Proposition 22.1 (Existence of a stable lattice)**.** *Every such representation admits a G-stable lattice.*

Proof. Choose a basis and identify V with K^n, with standard lattice $L_0 = \mathcal{O}^n$. The image $\rho(G)$ is compact because G is compact and ρ is continuous. Each matrix-entry function is bounded on

that image. Hence there is an integer $m \geq 0$ with $\rho(g)L_0 \subset \varpi^{-m}L_0$ for every $g \in G$. Let L be the $\mathcal{O}$-span of all $\rho(g)L_0$. Then

$$L_0 \subset L \subset \varpi^{-m}L_0.$$

Since $\mathcal{O}$ is noetherian, the submodule L of the finitely generated module $\varpi^{-m}L_0$ is finitely generated. It is torsion-free, hence free over the discrete valuation ring, and it has rank n because it contains L_0. Finally, multiplication by $\rho(h)$ permutes the generating submodules $\rho(g)L_0$, giving $\rho(h)L = L$ for every $h \in G$. □

The proof uses continuity and compactness, but not semisimplicity. It produces an integral realization of a representation that already exists. It does not construct a Galois representation from automorphic data, nor does it prove local compatibility of such a representation. Those are separate arithmetic questions.

22.2 Reduction modulo the maximal ideal

A stable lattice gives a representation on $L/\varpi L$, a finite-dimensional vector space over the finite field k. Its image is finite. The corresponding kernel is open because the action is continuous and the target is discrete and finite.

Two stable lattices need not be equal. Any two lattices L, L' in the same K-space are commensurable: there is m with $\varpi^m L \subset L' \subset \varpi^{-m}L$. This follows by writing a basis of each in terms of a basis of the other and clearing finitely many denominators. The resulting finite quotients measure the integral difference, even when the rational representation is the same.

Changing a lattice can change the extension structure of its reduction. One should therefore distinguish a rational representation, an integral lattice in it, and the residual representation obtained from that lattice. A statement about traces of the rational representation does not by itself select the integral structure needed for a divisibility argument.

22.3 Semisimplicity in a finite-image case

Suppose G acts through a finite group F on a vector space over a field of characteristic zero. If $W \subset V$ is an invariant subspace, choose any linear projection $p : V \to W$ that is the identity on W. Then

$$\overline{p} = \frac{1}{|F|} \sum_{g \in F} \rho(g) p \rho(g)^{-1}$$

is equivariant, still maps to W, and is still the identity on W. Its kernel is an invariant complement. Inducting on dimension proves semisimplicity.

Over the ring of integers, the division by $|F|$ need not preserve a lattice when $\ell \mid |F|$. Over a field of characteristic ℓ, the formula may not even make sense. This elementary distinction explains why a rational semisimplicity assertion cannot be substituted for an integral splitting theorem.

22.4 Characteristic polynomials and traces

Let A be an endomorphism of an n-dimensional vector space over a field of characteristic zero. Write

$$\det(1 - TA) = \sum_{j=0}^{n} (-1)^j e_j T^j, \qquad p_m = \operatorname{Tr}(A^m).$$

The coefficients satisfy Newton’s identities

$$j e_j = \sum_{m=1}^{j} (-1)^{m-1} e_{j-m} p_m, \qquad e_0 = 1. \tag{96}$$

One way to verify this is to extend scalars to an algebraic closure, write the determinant as $\prod_i (1 - \alpha_i T)$, and differentiate its logarithm as a formal power series:

$$-T \frac{d}{dT} \log \det(1 - TA) = \sum_{m \ge 1} p_m T^m.$$

Multiplying by the determinant and comparing coefficients gives (96). Thus traces of the first n powers determine the characteristic polynomial recursively.

For two representations of a group, equality of characteristic polynomials on a specified set of group elements is a concrete compatibility statement. Extending it to an isomorphism of representations requires hypotheses about that set and about semisimplicity. Equality at one element, or on a small subgroup, is not a substitute for those hypotheses.

22.5 A distinction between dimension and local compatibility

Dimension equality says only that two spaces have the same number of basis vectors. Equality of characteristic polynomials at selected operators is stronger. Equality of the associated representations is stronger still. When a local object also carries a nilpotent operator, its compatibility data must account for that operator separately.

To make the last point explicit, consider a two-dimensional space with basis e_1, e_2 and an invertible operator $F = \operatorname{diag}(q^{-1}, 1)$, where $q > 1$. Both $N = 0$ and the operator defined by $N(e_2) = e_1$, $N(e_1) = 0$ satisfy

$$F N F^{-1} = q^{-1} N.$$

The operator F and all its characteristic polynomials are identical in the two examples, but the nilpotent operators differ. The kernels of N have dimensions two and one, respectively.

For an unramified Weil–Deligne datum with these operators, the local factor is formed from F on $\ker N$. In the first example its reciprocal polynomial is $(1 - q^{-1}T)(1 - T)$; in the second it is $1 - q^{-1}T$. Hence the local factor distinguishes the two data even though the semisimple operator F is unchanged. A theorem comparing only characteristic polynomials must not silently be read as a theorem comparing all monodromy data.

22.6 Weight multisets and conjugation

For a CM field E, let c denote complex conjugation. Suppose a representation is equipped with an integer multiset H_τ of size n for each embedding $\tau : E \to \overline{\mathbb{Q}}_\ell$, and suppose there is an integer w such that

$$H_{\tau \circ c} = \{ w - h : h \in H_\tau \} \tag{97}$$

with multiplicities. This is a numerical symmetry of the weight data. If H_τ has distinct elements $h_1 < \cdots < h_n$, the partner multiset in increasing order is $w - h_n < \cdots < w - h_1$. In particular, regularity is preserved.

Summing (97) gives

$$\sum H_{\tau \circ c} = n w - \sum H_\tau.$$

The sums are the weights naturally associated to determinants when the weight theory is additive under tensor products and exterior powers. The multiset identity is more precise than equality of those sums: many different multisets have the same total weight.

Fix the convention that twisting by a cyclotomic power m adds m to each weight. Then the new weights are $H_\tau + m$, and their conjugation symmetry has central parameter $w + 2m$:

$$w + 2m - (h + m) = w - h + m.$$

Thus a normalization change affects both the individual weights and the polarization parameter. Checking only one side of this change would create a sign or shift error in the compatibility statement.

22.7 What a representation-theoretic construction must supply

A construction used in the arithmetic comparison must provide more than a vector space with a continuous Galois action. Depending on the application, the required data can include semisimplicity, characteristic polynomial identities away from the residual prime, crystalline or potentially semistable conditions at selected places, weight multisets, and compatibility of those weights with conjugation.

These conditions have different logical strengths and different local ranges. For example, a theorem may impose a particular characteristic polynomial identity only at places where the local parameter is tame, squarefree, or has zero monodromy. Such a restricted statement must remain restricted when used in a later proof. The elementary tools in this section explain some operations that preserve or compare the data; they do not establish the existence theorem attaching all of them to a prescribed automorphic representation.

23 Local representations and integral monodromy

The arithmetic comparison uses more local information than the trace of Frobenius. It must distinguish the semisimple action, the inertia action, and the nilpotent monodromy operator. This section develops the elementary linear algebra and integral formulas that keep those pieces separate. Existence of a global representation with the required local properties remains a distinct construction problem.

23.1 A local representation and its invariants

Let F be a nonarchimedean local field with residue field of cardinality q, and let ℓ be a prime different from its residue characteristic. Let $E/\mathbb{Q}_\ell$ be a finite extension and let $r : G_F \to \mathrm{GL}(V)$ be a continuous representation on a finite-dimensional E-vector space. A stable lattice $L \subset V$ allows reduction modulo the maximal ideal of $\mathcal{O}_E$ and finite congruence quotients of the image.

The inertia subgroup I_F is the kernel of the action on the residue field. If inertia acts trivially, the representation is unramified and is described by a Frobenius operator. If inertia acts nontrivially, the Frobenius polynomial on all of V does not encode its action. Even when a Weil–Deligne description is available, a nilpotent operator must be retained in addition to the Weil action.

We use geometric Frobenius in the formulas below. The compatibility relation for its operator F_0 and monodromy N is

$$F_0 N F_0^{-1} = q^{-1} N. \tag{98}$$

With arithmetic Frobenius the power of q is inverted. A choice of convention must be made before comparing local Euler factors or weights.

23.2 Nilpotent logarithms

Let A be an endomorphism with $A^n = 0$ over a characteristic-zero field. Define

$$\log(I + A) = \sum_{j=1}^{n-1} \frac{(-1)^{j+1}}{j} A^j.$$

For a nilpotent N with $N^n = 0$, define

$$\exp N = \sum_{j=0}^{n-1} \frac{N^j}{j!}.$$

These are finite expressions. The formal identities $\exp(\log(1 + X)) = 1 + X$ and $\log(\exp X) = X$ descend to the quotient ring $E[X]/(X^n)$, giving inverse correspondences between unipotent operators $I + A$ and nilpotent operators N.

If A preserves an integral lattice and $\ell \geq n$, every integer $1, \dots, n-1$ and every factorial up to $(n-1)!$ is a unit in $\mathcal{O}_E$. Hence both logarithm and exponential preserve integral endomorphisms. This elementary denominator bound explains one use of a hypothesis that the coefficient prime is large compared with the representation dimension. It does not by itself establish any local-global compatibility statement.

23.3 Powers and continuous tame parameters

For a nilpotent N, the polynomial $\exp(tN)$ satisfies

$$\exp(tN) \exp(uN) = \exp((t + u)N).$$

This follows by multiplying the finite series and using the binomial identity for the coefficient of each N^k. Thus it defines an additive-parameter one-parameter subgroup. For integral N and unit factorials as above, $t \in \mathbb{Z}_\ell$ gives an integral matrix depending continuously on t.

If $U = \exp N$, then $U^m = \exp(mN)$ for every integer m. Consequently a conjugacy relation $F_0 U F_0^{-1} = U^{q^{-1}}$, interpreted through the continuous $\mathbb{Z}_\ell$-parameter because q is an ℓ-adic unit, gives (98) after taking the nilpotent logarithm. The converse follows by exponentiating.

This explains the form of monodromy for a unipotent tame action when such an action has been identified. It does not assert that every continuous inertia action is already unipotent. Passing to a subgroup where that occurs is an additional local theorem in the general case.

23.4 Generalized Frobenius eigenspaces

Extend coefficients so that the characteristic polynomial of F_0 splits. If V_α is its generalized eigenspace of eigenvalue α, then

$$N(V_\alpha) \subset V_{\alpha/q}. \tag{99}$$

Indeed,

$$(F_0 - \alpha/q)N = q^{-1} N (F_0 - \alpha),$$

and iterating this identity shows that a sufficiently high power of $F_0 - \alpha/q$ kills Nv whenever a sufficiently high power of $F_0 - \alpha$ kills v.

Since F_0 is invertible, all eigenvalues are nonzero. Multiplication by q^{-1} has infinite order in the characteristic-zero coefficient field, so a finite set of eigenvalues contains only finite strings along each orbit. Formula (99) therefore organizes monodromy into arrows between consecutive members of such strings.

A useful sufficient condition for $N = 0$ is that the spectrum contain no pair $\alpha, \alpha/q$. This is stronger than requiring the characteristic polynomial to be square free. Distinct eigenvalues can still differ by a factor of q, so square freeness alone does not eliminate monodromy.

23.5 A single monodromy string

For $m \geq 1$ and $\alpha \neq 0$, let V have basis $e_0, \dots, e_{m-1}$ and define

$$F_0 e_j = \alpha q^{-j} e_j, \qquad N e_j = \begin{cases} e_{j+1}, & j < m-1, \\ 0, & j = m-1. \end{cases}$$

Then (98) holds. The characteristic polynomial in Euler-factor form is

$$\det(1 - T F_0 \mid V) = \prod_{j=0}^{m-1} (1 - \alpha q^{-j} T),$$

while $\ker N$ is the one-dimensional span of e_{m-1}. Thus the determinant on the monodromy kernel is only $1 - \alpha q^{-(m-1)} T$.

A direct sum of such strings gives a transparent collection of examples with the same total dimension but different monodromy ranks. The number of strings is $\dim \ker N$, while their lengths determine the nilpotent Jordan type. The full Frobenius spectrum records all points on the strings; it does not by itself specify which arrows between those points are present.

23.6 The local Euler factor

For Weil–Deligne data (r, N), with a specified finite inertia action compatible with N, the standard local factor is

$$L(s, r, N) = \det\left(1 - q^{-s} F_0 \mid (\ker N)^{I_F}\right)^{-1},$$

under the geometric-Frobenius convention in use. The precise convention for the Weil operator must be aligned with this formula. The essential point for the comparison is the space on which the determinant is taken: both the monodromy kernel and the inertia invariants occur.

If inertia is trivial and $N = 0$, the determinant is on all of V. If $N \neq 0$, it is generally on a smaller subspace. Thus equality of $\det(1 - T F_0 \mid V)$ for two representations does not imply equality of their local Euler factors. The two-dimensional example given earlier is the shortest instance of this phenomenon.

23.7 Jordan type from ranks of powers

For a nilpotent operator N, put $r_k = \operatorname{rank} N^k$, with $r_0 = \dim V$ and $r_k = 0$ for large k. A Jordan block of length m contributes $\max(m - k, 0)$ to r_k. Therefore

$$r_{k-1} - r_k = \#\{\text{blocks of length at least } k\},$$

and

$$r_{k-1} - 2r_k + r_{k+1} = \#\{\text{blocks of length exactly } k\}. \tag{100}$$

The rank sequence determines the complete Jordan type. Its successive drops are nonincreasing, and its second differences are nonnegative. These conditions are necessary for a proposed list of monodromy ranks.

When comparing two local representations, matching these ranks in addition to the Frobenius action can provide information absent from the characteristic polynomial alone. However, identifying Jordan type still does not identify the full Weil representation unless the placement of the strings among the Frobenius eigenspaces is also controlled.

23.8 Reduction can lose monodromy

Let $L = \mathcal{O}_E e_1 \oplus \mathcal{O}_E e_2$ and take $Ne_1 = \varpi e_2$, $Ne_2 = 0$. Over E the operator has rank one, but its reduction modulo ϖ is zero. Thus monodromy rank is not preserved automatically by reduction of a stable lattice. The defect is measured by the failure of the image to be saturated:

$$(E \cdot N(L)) \cap L = \mathcal{O}_E e_2, \qquad \operatorname{length}\big(((E \cdot N(L)) \cap L)/N(L)\big) = 1.$$

More generally, Smith normal form of $N : L \to L$ records the nonzero elementary divisors. The number of unit elementary divisors is the rank after reduction, while the number of nonzero elementary divisors is the rank over E.

This connects local representation theory with the lattice comparison. A rational rank equality can coexist with a nontrivial integral cokernel. The latter is exactly the sort of finite-length term that must be bounded before a rational comparison becomes an integral one.

23.9 Extension classes and local splitting

Let $\chi_1, \chi_2 : G \to E^\times$ be characters. An upper-triangular representation with these diagonal characters has matrices

$$r(g) = \begin{pmatrix} \chi_1(g) & b(g) \\ 0 & \chi_2(g) \end{pmatrix}.$$

The homomorphism condition is $b(gh) = \chi_1(g)b(h) + b(g)\chi_2(h)$. Set $c(g) = b(g)/\chi_2(g)$ and $\chi = \chi_1\chi_2^{-1}$. Then

$$c(gh) = c(g) + \chi(g)c(h),$$

so c is a cocycle with coefficients in the one-dimensional module $E(\chi)$. Changing the lift of the second basis vector by a multiple of the first changes c by a coboundary. Hence the extension splits precisely when its cohomology class vanishes.

After choosing integral lattices, the same formulas apply with integral or torsion coefficients when the entries lie in the required modules. Local splitting means that the class vanishes after restriction to a local subgroup. Global splitting requires one global coboundary. The local-kernel groups developed in Section 19 therefore measure exactly a possible obstruction to making all local splittings compatible.

23.10 Traces, semisimplification, and what remains to compare

For an upper-triangular extension as above,

$$\operatorname{Tr} r(g) = \chi_1(g) + \chi_2(g), \qquad \det r(g) = \chi_1(g)\chi_2(g),$$

independently of $b(g)$. Thus traces and characteristic polynomials cannot detect whether the extension splits. Passing to semisimplification forgets the extension class while preserving these polynomials. This is a second reason, distinct from the monodromy example, why polynomial compatibility can be weaker than compatibility of the actual representations.

A complete local comparison must specify which level of structure is asserted: the semisimplified Weil action, a Frobenius polynomial, the nilpotent operator, an inertia condition, a filtered module at the coefficient prime, or an integral lattice. The arithmetic argument may need only some of these in a restricted range, but that range must be stated. The elementary identities in this section make the distinctions visible and explain why stronger comparison hypotheses cannot be replaced by a single equality of characteristic polynomials.

24 Weights, interlacing patterns, and dimension estimates

The ranks of the spaces entering a restriction map depend on algebraic weights. A dominant weight contains more information than its largest coordinate: differences between its coordinates determine the size of the associated type. Interlacing patterns give a convenient exact combinatorial model for these dimensions and for successive restrictions. We develop the counting identities and the estimates that follow from them.

24.1 Dominant weights and shifts

A dominant integral weight of length n is a tuple $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathbb{Z}^n$ with $\lambda_1 \ge \cdots \ge \lambda_n$. Its total weight is $|\lambda| = \sum_i \lambda_i$. Adding the same integer c to every coordinate changes the total weight by nc but leaves all differences $\lambda_i - \lambda_j$ unchanged. In a general linear representation, this operation corresponds to a determinant twist. It changes a central character while preserving dimension.

The dual dominant weight is $\lambda^\vee = (-\lambda_n, \dots, -\lambda_1)$. Reversal is needed to restore the decreasing order after changing signs. A polarization condition of the form

$$\lambda_i + \lambda_{n+1-i} = w \qquad (1 \le i \le n)$$

is equivalent to $\lambda^\vee = \lambda - w(1, \dots, 1)$. Thus duality agrees with a central twist. The corresponding symmetry of Hodge weights is the reflection $h \mapsto w - h$.

24.2 Shifted coordinates

Define strictly decreasing integers

$$a_i = \lambda_i + n - i.$$

Then $a_i - a_j = \lambda_i - \lambda_j + j - i > 0$ for $i < j$. The dimension expression is

$$D_n(\lambda) = \prod_{1 \le i < j \le n} \frac{\lambda_i - \lambda_j + j - i}{j - i} = \frac{\prod_{i<j}(a_i - a_j)}{\prod_{k=1}^{n-1} k!}. \tag{101}$$

The equality of denominators follows because the difference k occurs $n-k$ times in $\prod_{i<j}(j-i)$, and the same multiplicity occurs in $\prod_{k=1}^{n-1} k!$. Although (101) is displayed as a rational product, it is a positive integer. The interlacing count below proves this integrality without relying on cancellation in the product.

24.3 Interlacing

A weight μ of length $n-1$ interlaces λ if

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \cdots \geq \mu_{n-1} \geq \lambda_n.$$

The inequalities automatically make μ dominant. Set $b_i = \mu_i + n - 1 - i$. In shifted coordinates, interlacing becomes

$$a_{i+1} \leq b_i < a_i \qquad (1 \leq i < n). \tag{102}$$

These intervals are independent, and their ordering forces $b_i > b_{i+1}$. The unweighted number of possible next rows is $\prod_{i=1}^{n-1}(\lambda_i - \lambda_{i+1} + 1)$.

A full interlacing pattern has a top row λ of length n, a row of length $n-1$ interlacing it, and so on down to length one. All entries are integers. The number of full patterns is not merely the number of choices for the next row; each next row carries its own further multiplicity.

24.4 A determinant for the Vandermonde quotient

For integers $b_1 > \cdots > b_r$, one has

$$\det\left[\binom{b_i}{r-j}\right]_{1\leq i,j\leq r} = \frac{\prod_{i<j}(b_i - b_j)}{\prod_{k=1}^{r-1} k!}. \tag{103}$$

Here $\binom{x}{k} = x(x-1)\cdots(x-k+1)/k!$ is interpreted as a polynomial, so negative integers cause no difficulty. The determinant is alternating in the b_i and has total degree at most $r(r-1)/2$. It is therefore a constant multiple of the Vandermonde product, which has that same degree. Comparing leading terms in the columns gives the constant $1/\prod_{k=0}^{r-1} k!$.

This proof uses only polynomial divisibility. An alternating polynomial vanishes when $b_i = b_j$, so it is divisible by $b_i - b_j$; the distinct linear factors are relatively prime, so their product divides it. Degree then determines the quotient to be constant.

24.5 The interlacing summation identity

Proposition 24.1. *For every dominant integral weight of length $n \geq 2$,*

$$D_n(\lambda) = \sum_{\mu \ minterlaces \ \lambda} D_{n-1}(\mu).$$

Proof. Use the shifted intervals (102) and express $D_{n-1}(\mu)$ by (103). Since the choices of the b_i are independent, multilinearity of the determinant lets us sum each row separately. The polynomial identity

$$\sum_{b=v}^{u-1} \binom{b}{k} = \binom{u}{k+1} - \binom{v}{k+1}$$

follows by telescoping Pascal's identity and holds for all integer endpoints $u \geq v$. The sum is therefore the determinant with entries

$$\binom{a_i}{n-j} - \binom{a_{i+1}}{n-j}, \qquad 1 \leq i, j \leq n-1.$$

Start instead with the $n \times n$ matrix $[\binom{a_i}{n-j}]$. Replace its first $n-1$ rows simultaneously by each row minus the following row, keeping the last row. This row transformation has determinant one. The last column becomes zero in the first $n-1$ rows and remains one in the last. Expansion along it gives exactly the preceding determinant. Formula (103) at size n identifies it with $D_n(\lambda)$. □

Since $D_1(\lambda_1) = 1$, induction shows that $D_n(\lambda)$ is the number of full interlacing patterns with top row λ. This gives positivity and integrality of the dimension expression. It also provides an exact recursive decomposition of that dimension under a sequence of restrictions.

24.6 Small-rank examples

For $n = 2$, the lower entry can be any integer between λ_2 and λ_1, so $D_2(\lambda) = \lambda_1 - \lambda_2 + 1$. For $n = 3$,

$$D_3(\lambda) = \frac{(\lambda_1 - \lambda_2 + 1)(\lambda_2 - \lambda_3 + 1)(\lambda_1 - \lambda_3 + 2)}{2}.$$

If $\lambda = (k, 0, 0)$ with $k \geq 0$, the next row is $(j, 0)$, where $0 \leq j \leq k$, and it contributes $j + 1$ lower patterns. Thus the total is $\sum_{j=0}^{k}(j+1) = (k+1)(k+2)/2$.

For $\lambda = (k, k, 0)$, the next row is (k, j) and contributes $k - j + 1$, giving the same total. These two weights are related by duality and a central shift, so their equal dimensions are also immediate from (101).

24.7 Symmetric and exterior powers

For the weight $(k, 0, \dots, 0)$, the product formula telescopes to

$$D_n(k, 0, \dots, 0) = \binom{k+n-1}{n-1}.$$

This agrees with the dimension of the kth symmetric power of an n-dimensional space: its monomial basis is indexed by tuples of nonnegative exponents summing to k. For the weight with r entries equal to one and the rest zero,

$$D_n(\underbrace{1, \dots, 1}_{r}, 0, \dots, 0) = \binom{n}{r},$$

matching the basis of the rth exterior power indexed by r-element subsets. These cases connect the general dimension product directly to the symmetric powers and jet modules already used in local coordinates.

The coincidence of a numerical dimension formula with a geometric rank is not by itself an identification of the spaces. To apply it to a restriction map, one needs the specified algebraic type or a basis indexed by the corresponding patterns. Once that identification is available, the counting calculation supplies its rank exactly.

24.8 The alternant expression

For pairwise distinct nonzero complex numbers $x_1, \dots, x_n$, define

$$s_\lambda(x_1, \dots, x_n) = \frac{\det[x_j^{\lambda_i+n-i}]_{i,j}}{\det[x_j^{n-i}]_{i,j}}.$$

After a common central shift makes $\lambda_n \geq 0$, the numerator and denominator are alternating polynomials in the x_j. Divisibility by the Vandermonde shows that the quotient extends to a symmetric polynomial. Undoing the shift gives a symmetric Laurent polynomial. Its value at $x_1 = \cdots = x_n = 1$ is $D_n(\lambda)$.

To compute that limit, set $x_j = e^{ty_j}$ with distinct y_j and expand $e^{ta_i y_j}$ in powers of t. The first nonzero determinant term uses the distinct powers $0, 1, \dots, n-1$. It is

$$t^{n(n-1)/2} \frac{\det[a_i^k]_{i,\,0\leq k<n} \det[y_j^k]_{0\leq k<n,\,j}}{\prod_{k=0}^{n-1} k!}.$$

Taking the ratio with the same expression for $a_i = n - i$ cancels the powers of t, the y-Vandermonde, and the factorials. The remaining ratio is exactly (101).

This limiting argument explains why the shift by $n - i$ appears in the dimension formula. It is the sequence of exponents in the alternating denominator. Replacing it by the unshifted weight would lose the strictly positive factors $j - i$ and give an incorrect answer at repeated weight coordinates.

24.9 Uniform bounds in a weight box

Suppose $0 \leq \lambda_1 - \lambda_n \leq R$. Then every numerator in (101) lies between $j - i$ and $R + j - i$. Consequently

$$1 \leq D_n(\lambda) \leq \frac{(R+n)^{n(n-1)/2}}{\prod_{k=1}^{n-1} k!}. \tag{104}$$

A coarser but convenient logarithmic bound is

$$\log D_n(\lambda) \leq \frac{n(n-1)}{2} \log(R+n).$$

This remains valid when both R and n vary.

Large absolute weights do not necessarily give large dimension. If all coordinates equal c, the dimension is one regardless of $|c|$. By contrast, for the evenly spaced weight $\lambda_i = (n-i)k$ with $k \geq 0$,

$$D_n(\lambda) = (k+1)^{n(n-1)/2}. \tag{105}$$

Here every ratio in the product is exactly $k + 1$. The example separates growth coming from a determinant twist from growth coming from gaps between weights.

24.10 Stability under bounded perturbations

Suppose two dominant weights λ, λ' satisfy $|\lambda'_i - \lambda_i| \leq h$, and suppose $\lambda_i - \lambda_j \geq \tau(j-i)$ for $i < j$, with $\tau \geq 0$. Write $d_{ij} = \lambda_i - \lambda_j + j - i \geq (\tau+1)(j-i)$. If $2h \leq (\tau+1)/2$, then the relative change of each numerator has absolute value at most one half. Since $|\log(1+u)| \leq 2|u|$ for $|u| \leq 1/2$,

$$|\log D_n(\lambda') - \log D_n(\lambda)| \leq \frac{4h}{\tau+1} \sum_{i<j} \frac{1}{j-i}.$$

The final sum equals $nH_{n-1} - (n-1)$, where $H_m = \sum_{j=1}^{m} 1/j$. Hence

$$|\log D_n(\lambda') - \log D_n(\lambda)| \le \frac{4hn}{\tau+1}(1+\log n). \tag{106}$$

This is more informative than a bound on each dimension separately when the weight gaps are large compared with the perturbation.

The hypothesis concerns every pair, but it follows from adjacent gap bounds $\lambda_i - \lambda_{i+1} \ge \tau$ by summing. It is therefore easy to check on a regular family of weights. Without such a gap condition, a bounded perturbation can have a much larger relative impact on a small numerator.

24.11 Rank growth versus exponential error decay

Suppose $n = n(x)$ and the weight spread is at most $R(x)$. The box bound gives

$$\log D_n(\lambda) \le \tfrac{1}{2} n(x)^2 \log(R(x) + n(x)).$$

If

$$n(x)^2 \log(R(x) + n(x)) = o(x^a),$$

then $D_n(\lambda) e^{-x^a} \to 0$. Indeed, take logarithms of the product and compare the two terms. This is the precise kind of calculation needed to turn an exponentially small spectral error into an absolute corank bound less than one.

If instead $n(x)^2 \log(R(x) + n(x))$ is comparable with or larger than x^a, the conclusion need not hold. A phrase such as “exponentially small error” has no quantitative force until the growing dimensions have been put on the same logarithmic scale.

24.12 Restriction and multiplicity

The interlacing recursion is multiplicity free at a single branching step in the pattern model: each allowed next row appears once. After several restrictions, the same final weight may be reached by many intermediate rows. Its multiplicity is the number of compatible partial patterns. The total dimension is the sum of each final dimension multiplied by that multiplicity.

For a restriction estimate, this means that summing only over distinct final weights can undercount the source rank. The indexing set must retain multiplicities or record them as coefficients. Conversely, if a selected family has a proven multiplicity-one property, the simpler set sum is valid. This bookkeeping is parallel to the need to combine weights when several finite-group characters induce the same primitive Dirichlet character.

24.13 Compatibility with the slope and frame arguments

Once an algebraic type identifies a source space V_i with dimension $D_{n_i}(\lambda^{(i)})$, these formulas supply the integers used in the frame normalization and in the slope-share estimate. They also control changes in dimension when the weights vary over a selected family. The operator kernels and integral lattices still need to be constructed on those same spaces.

Thus the full comparison requires three compatible descriptions: the branching data determine the ranks, the Hermitian realization determines the archimedean norms, and the integral structure determines the local lengths. None of these numerical calculations alone supplies the other two structures. Their role is to make a proposed comparison quantitative once the common representation has been identified.

25 Lattice indices and local length budgets

25.1 An index as a product of local contributions

Let $L' \subset L$ be free abelian groups of the same finite rank, with finite quotient $Q = L/L'$. For each rational prime ℓ, let Q_ℓ be the subgroup of elements annihilated by some power of ℓ. The elementary primary decomposition gives $Q = \bigoplus_\ell Q_\ell$, with only finitely many nonzero terms. As a finite $\mathbb{Z}_\ell$-module, Q_ℓ has finite length, and Lemma 18.1 gives

$$\log[L : L'] = \sum_\ell (\log \ell) \operatorname{length}_{\mathbb{Z}_\ell} Q_\ell. \tag{107}$$

Thus logarithmic prime weights are exactly the weights that turn local lengths into a global index.

The same identity can be expressed using tensor products: $Q \otimes_{\mathbb{Z}} \mathbb{Z}_\ell \simeq Q_\ell$. To see this, decompose Q into cyclic groups of prime-power order. Tensoring a group of order prime to ℓ gives zero, since that order is invertible in $\mathbb{Z}_\ell$; tensoring a cyclic ℓ-power group preserves it.

25.2 A determinant calculation over a discrete valuation ring

Let $\mathcal{O}$ be a discrete valuation ring with uniformizer ϖ, fraction field K, and valuation normalized by $v(\varpi) = 1$. Suppose $A \in M_n(\mathcal{O})$ has nonzero determinant. Then

$$\operatorname{length}_{\mathcal{O}}(\mathcal{O}^n / A\mathcal{O}^n) = v(\det A). \tag{108}$$

A proof uses diagonal reduction over the principal ideal domain $\mathcal{O}$: multiplication on the left and right by invertible integral matrices reduces A to $\operatorname{diag}(\varpi^{a_1}, \dots, \varpi^{a_n})$, up to units, with $a_i \geq 0$. Such operations preserve the quotient up to isomorphism and change the determinant only by a unit. The quotient is then a direct sum of $\mathcal{O}/\varpi^{a_i}\mathcal{O}$, whose lengths sum to $\sum_i a_i = v(\det A)$.

This formula depends on an injective map between lattices of equal rank. If the determinant vanishes, the quotient has positive rank and is not a finite-length module. A determinant estimate cannot be used as a finite-index estimate until rank equality and injectivity have been established.

25.3 Finite residue fields and normalized logarithms

If the residue field of $\mathcal{O}$ has cardinality ℓ^f, a finite-length module of length s has cardinality ℓ^{fs}. Its length after restriction of scalars to $\mathbb{Z}_\ell$ is fs. Consequently

$$\log |M| = s \log(\ell^f) = (\log \ell) \operatorname{length}_{\mathbb{Z}_\ell} M.$$

This is the normalization needed to compare lengths over local rings of a number field with sums indexed by rational primes. Omitting the residue degree would undercount the cardinality.

If several primes of a number field lie above one rational prime, the contributions from all of them enter the same $\mathbb{Z}_\ell$-length. The resulting rational-prime weighted sum is therefore compatible with a global index, provided the decomposition into local modules has been proved.

25.4 Nested lattices and addition of budgets

For $L'' \subset L' \subset L$ with finite indices, there is a short exact sequence

$$0 \longrightarrow L'/L'' \longrightarrow L/L'' \longrightarrow L/L' \longrightarrow 0.$$

Hence

$$[L : L''] = [L : L'][L' : L''], \qquad \log[L : L''] = \log[L : L'] + \log[L' : L''].$$

The same equality holds prime by prime for the local lengths. Thus successive modifications consume additive logarithmic budgets. If a proof bounds each modification separately, the total bound must include their sum. Reusing one budget for several independent modifications requires a common estimate that already controls that sum.

For two lattices L_1, L_2 in the same rational vector space, the intersection is a common sublattice of finite index. The quotient $(L_1 + L_2)/L_2$ is naturally isomorphic to $L_1/(L_1 \cap L_2)$. Such identities are useful for moving between intersection and sum constructions without changing the index being measured.

25.5 Integral maps and obstruction groups

Suppose $T : L \to M$ is a homomorphism of free abelian groups, and the induced rational map has image of rank r. The saturation of the integral image is

$$\mathrm{Sat}(T(L)) = (T(L) \otimes \mathbb{Q}) \cap M.$$

The quotient $\mathrm{Sat}(T(L))/T(L)$ is finite and measures the failure of the image to be primitive in M. This quotient can be nontrivial even when the rational image has the expected dimension.

In a diagonal form, with nonzero diagonal entries $d_1, \dots, d_r$, the torsion quotient is a direct sum of $\mathbb{Z}/d_i\mathbb{Z}$. Its order is $\prod_i |d_i|$, and its local length at ℓ is $\sum_i v_\ell(d_i)$. Thus rank estimates and divisibility estimates control different aspects of the same map: rank counts the nonzero diagonal entries, while local lengths count their valuations.

This distinction is relevant to the spectral estimate in Proposition 15.1. A large rational rank does not bound the integral index of the image. A cohomological obstruction bound can enter an integral argument only after a map from the obstruction group to the relevant saturation quotient, or another precise comparison, has been constructed.

25.6 A conditional transfer from cohomology to an index

Suppose that for each prime ℓ a finite module Q_ℓ is a subquotient of

$$\bigoplus_{1 \le j < t} \mathrm{Sha}_\ell(E, j)^{\oplus m_j}, \qquad 0 \le m_j \le m.$$

Length additivity and monotonicity give

$$\mathrm{length}_{\mathbb{Z}_\ell} Q_\ell \le m \sum_{1 \le j < t} \mathrm{length}_{\mathbb{Z}_\ell} \mathrm{Sha}_\ell(E, j).$$

If $\mathcal{V}(E, t, B)$ holds, then for every finite set of primes,

$$\sum_{\ell \in P} (\log \ell)\, \mathrm{length}\, Q_\ell \le mB.$$

If these Q_ℓ are the primary components of a lattice quotient L/L', equation (107) yields $[L : L'] \le e^{mB}$.

All the numerical steps are elementary once the subquotient comparisons are available. The content of an arithmetic application is often the construction of those comparisons, including the multiplicities m_j and their uniform bounds. A bare coincidence between the dimensions of two spaces would not imply the subquotient statement.

25.7 Compatibility with a varying twist range

Suppose a construction at scale x needs twists only up to $r(x)$, with $r(x) \le \lfloor x^{401} \rfloor$. The monotonicity of $\mathcal{V}$ permits the larger-range cohomological bound to be used for that construction. If a multiplicity bound also varies with x, its growth enters the transferred index bound explicitly.

For instance, if $m(x) \le Cx^a$ and the cohomological budget is $\exp(\frac{1}{2}x^\varepsilon)$, then the logarithm of the resulting index is bounded by $Cx^a \exp(\frac{1}{2}x^\varepsilon)$. For fixed $a, C, \varepsilon > 0$, this is at most $\exp(x^\varepsilon)$ for sufficiently large x, because $\log C + a \log x \le \frac{1}{2}x^\varepsilon$ eventually. This is a different absorption from replacing a factor by x^ε: here there is room inside an exponential budget.

The distinction between polynomial and exponential budgets should be maintained throughout the construction. A factor harmless inside $\exp(x^\varepsilon)$ may be significant in the final power bound $x^{1/2+A\varepsilon}$. The normalization step has to account for that change of scale before the analytic criterion is invoked.

26 Euclidean lattices and slope filtrations

Logarithmic indices measure finite-place defects. At the infinite places, the analogous quantities are logarithms of covolumes and operator norms. Slope theory organizes these quantities across subspaces of different ranks. We first work over $\mathbb{Z}$, where all normalizations can be seen directly. The subsequent passage to number fields adds discriminant and archimedean multiplicity factors.

26.1 Covolumes and degree

A Euclidean lattice L is a free abelian subgroup of a finite-dimensional real inner-product space whose real span $L_{\mathbb{R}}$ has dimension $r = \operatorname{rank} L$, and which is discrete in that span. If $e_1, \dots, e_r$ is a $\mathbb{Z}$-basis, set

$$\operatorname{covol}(L) = \det(\langle e_i, e_j \rangle)^{1/2}, \qquad \deg L = -\log \operatorname{covol}(L).$$

A change of integral basis has determinant ± 1, so the definition is independent of the basis. We set $\operatorname{covol}(0) = 1$ and $\deg 0 = 0$. For $r > 0$, the slope is $\mu(L) = \deg L / r$.

A subgroup $A \subset L$ is saturated if L/A is torsion free. Equivalently, $A = L \cap A_{\mathbb{R}}$. To verify this, if $x \in L \cap A_{\mathbb{R}}$, write x in a basis of L adapted to the rational subspace spanned by A; the coefficients show that a positive multiple of x belongs to A. Saturation then gives $x \in A$. The converse follows because $nx \in A$ implies $x \in A_{\mathbb{R}}$.

For an arbitrary A, its saturation is $A^{\mathrm{sat}} = L \cap A_{\mathbb{R}}$. The quotient A^{sat}/A is finite, and

$$\deg A^{\mathrm{sat}} = \deg A + \log[A^{\mathrm{sat}} : A]. \tag{109}$$

Thus saturation raises degree without changing rank. Extremal slopes among sublattices can therefore be sought among saturated sublattices.

26.2 Quotient metrics and additivity

For saturated $A \subset L$, identify $L_{\mathbb{R}}/A_{\mathbb{R}}$ with the orthogonal complement of $A_{\mathbb{R}}$. Give L/A the lattice structure obtained by orthogonal projection. It is discrete: extend a basis of A to a basis of L, and project the additional basis vectors. These projections form a real basis of the quotient space and generate the quotient lattice.

The determinant of the Gram matrix, or the volume of a parallelepiped computed by orthogonal projection, gives

$$\operatorname{covol}(L) = \operatorname{covol}(A)\operatorname{covol}(L/A), \qquad \deg L = \deg A + \deg(L/A). \tag{110}$$

The formula uses the quotient metric. An arbitrary metric chosen independently on L/A would introduce a correction term.

If A is not saturated, L/A has torsion. Its free quotient is L/A^{sat}, and the finite torsion order must be kept separately. Using (109) and (110),

$$\deg L = \deg A + \log[A^{\mathrm{sat}} : A] + \deg(L/A^{\mathrm{sat}}).$$

This separates an integral index from an archimedean covolume. It is exactly the separation needed when local cohomology bounds the torsion term while operator norms control the free quotient.

26.3 Existence of maximal slope

For $L \neq 0$, define

$$\mu_{\max}(L) = \sup_{0 \neq A \subset L} \mu(A), \qquad \mu_{\min}(L) = \inf_{A \subsetneq L \ msaturated} \mu(L/A).$$

Both extrema are finite and attained. For the maximum, fix a rank k. A basis $a_1, \dots, a_k$ of a rank-k sublattice determines a nonzero vector $a_1 \wedge \cdots \wedge a_k$ in the lattice $\bigwedge^k L$. Its Euclidean norm is the covolume of the sublattice. A lattice has only finitely many vectors in a bounded ball and a positive minimum norm among its nonzero vectors. Hence the covolumes of rank-k sublattices are bounded away from zero, and their minimum is attained if that rank occurs. There are only r possible positive ranks. The largest slope among these finitely many rankwise maxima is therefore attained.

For the minimum, use additivity:

$$\mu(L/A) = \frac{\deg L - \deg A}{r - \operatorname{rank} A}.$$

For each fixed rank of A, minimizing this quantity means maximizing $\deg A$, which is attained by the same exterior-power argument. The zero sublattice is allowed. In particular,

$$\mu_{\min}(L) \leq \mu(L) \leq \mu_{\max}(L). \tag{111}$$

These statements depend on discreteness. If one replaces integral sublattices by arbitrary real subspaces and allows arbitrary rescaling of their bases, the covolume is not an intrinsic lattice quantity and the extrema no longer express the intended arithmetic information.

26.4 Supermodularity of degree

For saturated sublattices $A, B \subset L$ one has

$$\deg A + \deg B \leq \deg(A + B) + \deg(A \cap B). \tag{112}$$

The rank identity is an equality:

$$\operatorname{rank} A + \operatorname{rank} B = \operatorname{rank}(A + B) + \operatorname{rank}(A \cap B).$$

To understand the degree inequality, first pass to the quotient by the real span of $A \cap B$. Extend a basis of $A \cap B$ to bases of A and B. After orthogonal projection, the remaining vectors span subspaces with zero intersection. The volume of their union is at most the product of their separate volumes: successive orthogonal projection can only decrease vector lengths, and hence the Gram–Schmidt volume. Multiplying back the common intersection volume gives

$$\operatorname{covol}(A + B) \operatorname{covol}(A \cap B) \le \operatorname{covol}(A) \operatorname{covol}(B).$$

Taking negative logarithms proves (112).

The possible strictness records the angle between the two subspaces. For two independent vectors a, b in a plane, the formula reduces to $\|a \wedge b\| = \|a\| \|b\| \sin \theta$. The degree of their sum is larger than the sum of their one-dimensional degrees by $-\log \sin \theta$. Nearly parallel vectors therefore create a large positive correction. Orthogonality is the case of equality.

26.5 The first maximal-slope sublattice

Put $m = \mu_{\max}(L)$. Suppose A and B both have slope m. They are saturated: otherwise saturation would strictly raise the degree at the same rank, contradicting maximality. By the definition of m,

$$\deg(A + B) \le m \operatorname{rank}(A + B), \qquad \deg(A \cap B) \le m \operatorname{rank}(A \cap B),$$

with the zero intersection interpreted by degree and rank zero. Supermodularity and the rank identity force equality in both inequalities. Thus $A + B$ again has slope m.

Choose a sublattice L_1 of slope m with maximal possible rank and saturate it. Saturation cannot raise its slope above m, so it was already saturated up to an index-one equality. For every other maximal-slope sublattice A, the sum $L_1 + A$ has slope m and no larger rank. Its saturation equals L_1, so $A \subset L_1$. Consequently L_1 is the unique largest sublattice of maximal slope.

Every nonzero sublattice of L_1 has slope at most $m = \mu(L_1)$. Such a lattice is called semistable. The maximal-slope sublattice is therefore semistable by construction; semistability here is a statement about covolumes of all integral sublattices.

26.6 Construction of the slope filtration

If $L_1 \neq L$, apply the same construction to the quotient lattice L/L_1 with its quotient metric. Its maximal slope is strictly less than m. Indeed, let Q be a nonzero saturated sublattice of L/L_1 and let $B \subset L$ be its inverse image. Additivity gives

$$\deg B = \deg L_1 + \deg Q.$$

If $\mu(Q) > m$, the weighted average $\mu(B)$ exceeds m, a contradiction. If $\mu(Q) = m$, then B is a maximal-slope sublattice strictly containing L_1, again a contradiction. Since the maximal slope of the quotient is attained, it is strictly less than m.

Iterating gives a finite filtration by saturated sublattices

$$0 = L_0 \subsetneq L_1 \subsetneq \cdots \subsetneq L_t = L \tag{113}$$

whose quotients $Q_i = L_i / L_{i-1}$ are semistable and have strictly decreasing slopes $\mu_1 > \mu_2 > \cdots > \mu_t$. Termination follows because rank increases at every step. The first step was unique, and the same is true recursively in each quotient, so the entire filtration is canonical for the Euclidean lattice.

Repeat each μ_i exactly $\operatorname{rank} Q_i$ times to obtain a decreasing list $s_1 \geq \cdots \geq s_r$. Degree additivity gives

$$\sum_{j=1}^{r} s_j = \deg L. \tag{114}$$

The graph with vertices $(\operatorname{rank} L_i, \deg L_i)$ is the slope polygon. Its segments have slopes μ_i, and it is concave because these slopes decrease.

26.7 The polygon bounds all sublattices

For a sublattice $A \subset L$, intersect A with each L_i. The successive quotients inject into Q_i as abelian groups. Their quotient metric from A is at least as large as the metric obtained by mapping into Q_i, because taking a quotient by the larger subspace $L_{i-1,\mathbb{R}}$ can only decrease norms. Therefore their degrees are at most the degrees of their images in Q_i. Semistability of Q_i gives

$$\deg A \leq \sum_i \mu_i a_i, \quad a_i = \operatorname{rank}(A \cap L_i) - \operatorname{rank}(A \cap L_{i-1}),$$

where $0 \leq a_i \leq \operatorname{rank} Q_i$ and $\sum_i a_i = \operatorname{rank} A$. The right side is maximized by using the largest available slopes first. Thus, if $P_L(k) = \sum_{j=1}^{k} s_j$,

$$\deg A \leq P_L(\operatorname{rank} A). \tag{115}$$

Equality holds at the filtration steps $A = L_i$.

This shows that the polygon is the least concave polygon with endpoints $(0,0)$ and $(r, \deg L)$ lying above all points $(\operatorname{rank} A, \deg A)$. It also identifies the first and last slopes with $\mu_{\max}(L)$ and $\mu_{\min}(L)$ respectively. For the last claim, any quotient L/A has degree $\deg L - \deg A \geq \deg L - P_L(\operatorname{rank} A)$, whose average over the remaining ranks is at least s_r. The last filtration quotient attains s_r.

26.8 Duality

The dual lattice is

$$L^\vee = \{y \in L_{\mathbb{R}} : \langle y, x \rangle \in \mathbb{Z} \text{ for all } x \in L\}.$$

If a basis of L is represented by a real matrix B in orthonormal coordinates, the dual basis is represented by $B^{-\mathsf{T}}$. Consequently

$$\operatorname{covol}(L^\vee) = \operatorname{covol}(L)^{-1}, \qquad \deg L^\vee = -\deg L, \qquad (L^\vee)^\vee = L.$$

For saturated $A \subset L$, the dual of L/A identifies isometrically with the sublattice of $L^\vee$ annihilating A. Conversely every saturated sublattice of $L^\vee$ arises this way from its annihilator in L. Taking extreme slopes gives

$$\mu_{\max}(L^\vee) = -\mu_{\min}(L), \qquad \mu_{\min}(L^\vee) = -\mu_{\max}(L). \tag{116}$$

The slope list of the dual is $-s_r, \ldots, -s_1$. This can also be seen by dualizing the filtration in reverse order; the quotient metrics and subspace metrics are dual to one another.

26.9 The slope inequality for a nonzero map

Let $T : L \to M$ be a nonzero homomorphism of Euclidean lattices, extended to a real-linear map of norm at most $C > 0$. Its kernel K is saturated because M is torsion free. Give L/K the quotient metric. The induced isomorphism of real spaces onto the image has all singular values at most C. If $q = \operatorname{rank} T$, its determinant on these spaces has absolute value at most C^q. Therefore

$$\operatorname{covol}(T(L)) \le C^q \operatorname{covol}(L/K),$$

and hence

$$\mu(T(L)) \ge \mu(L/K) - \log C \ge \mu_{\min}(L) - \log C.$$

Since $T(L) \subset M$, its slope is at most $\mu_{\max}(M)$. We obtain

$$\mu_{\min}(L) \le \mu_{\max}(M) + \log C. \tag{117}$$

In particular, if the reverse strict inequality holds, every such map is zero.

If T is injective, apply the determinant argument to a maximal-slope sublattice of L to obtain

$$\mu_{\max}(L) \le \mu_{\max}(M) + \log C.$$

For a surjective map, the same argument applied to each quotient of M shows $\mu_{\min}(L) \le \mu_{\min}(M) + \log C$. These are different consequences of the same volume inequality; they should not be used interchangeably without the corresponding rank condition.

26.10 Metric rescaling

If every norm on $L_{\mathbb{R}}$ is multiplied by e^{-a}, a rank-k sublattice has its covolume multiplied by e^{-ak}. Thus its degree increases by ak, every slope increases by a, and

$$P_{L,\mathrm{new}}(k) = P_{L,\mathrm{old}}(k) + ak.$$

The underlying saturated filtration is unchanged. Only the numerical thresholds at which its steps are selected move.

If both source and target metrics are rescaled, the logarithm of an operator norm changes by the difference of the two scale parameters. This precisely compensates for the change in the two slopes in (117). A slope inequality is therefore invariant under simultaneous consistent renormalization. It is not invariant under changing a metric in one place while retaining an operator norm computed in another.

26.11 An explicit rank-two example

Take $L = \mathbb{Z}e_1 \oplus 2\mathbb{Z}e_2$ with e_1, e_2 orthonormal. Then $\deg L = -\log 2$ and $\mu(L) = -\frac{1}{2}\log 2$. Every nonzero vector has length at least one, so every rank-one sublattice has degree at most zero. The sublattice $\mathbb{Z}e_1$ attains zero. The quotient by it has covolume two and slope $-\log 2$. Hence the slope list is $(0, -\log 2)$.

The average slope does not reveal either extreme. For example, a map from L to a rank-one lattice can kill the high-slope direction and be nonzero on the low-slope direction. This is why the nonzero-map inequality uses the minimum slope of the source. An injective map cannot kill the high-slope direction, which explains the stronger maximum-slope inequality in that case.

26.12 Counting low-slope directions

Suppose a list of N real numbers $s_1, \dots, s_N$ satisfies

$$\sum_{j=1}^{N} s_j \ge Ns, \qquad s_j \le s+u, \qquad u \ge 0.$$

For $\tau > 0$, let b be the number of indices with $s_j < s-\tau$. Bounding the low terms by $s-\tau$ and all other terms by $s+u$ gives

$$Ns \le \sum_{j} s_j \le b(s-\tau) + (N-b)(s+u).$$

Thus

$$b \le N\frac{u}{u+\tau}. \tag{118}$$

If $u = 0$, the average and upper bound force every $s_j = s$, so $b = 0$. The same inequality applies to the union of the slope lists of several lattices; only the total rank and total degree enter.

For a lattice, the span of the filtration factors with slope at least $s-\tau$ has minimum slope at least $s-\tau$. Its codimension is bounded by (118). One may therefore discard a controlled number of low-slope directions and retain a sublattice on which the nonzero-map inequality is strong. This is the bridge between an average arithmetic slope estimate and a spectral statement about the rank of many restriction maps.

26.13 Why a slope threshold must be optimized

A larger τ decreases the discarded proportion $u/(u+\tau)$, but weakens the retained minimum-slope bound $s-\tau$. In a later argument, the first effect improves a rank estimate while the second worsens an operator-height comparison. The parameter cannot be sent to infinity without cost. It must be chosen by balancing these two contributions.

For example, suppose the final upper bound has the schematic form

$$A\tau + B\frac{u}{u+\tau}, \qquad A, B, u > 0.$$

Its derivative is $A - Bu/(u+\tau)^2$. If $Bu > Au^2$, the stationary point is $\tau = \sqrt{Bu/A} - u > 0$, giving a bound $2\sqrt{ABu} - Au$. Otherwise the minimum over $\tau \ge 0$ occurs at $\tau = 0$. This elementary optimization explains the square-root losses that often appear when a slope-distribution estimate is combined with a rank estimate. The actual constants and admissible range must come from the maps under consideration.

27 Hermitian lattices over number fields

The preceding Euclidean theory applies to a number-field lattice after restriction of scalars. Its numerical slopes then contain a contribution from the discriminant of the field. Removing that contribution gives the normalization naturally associated with Hermitian Gram determinants. We derive the conversion explicitly, because the arithmetic comparison uses both local lengths and archimedean heights.

27.1 The real realization of a totally complex field

Let K be a totally complex number field of degree $d = 2s$. Choose one embedding $\sigma_v : K \hookrightarrow \mathbb{C}$ from each conjugate pair, where v ranges over the s infinite places. The map

$$\iota : K \longrightarrow \mathbb{C}^s, \qquad a \longmapsto (\sigma_v(a))_v$$

identifies $K \otimes_{\mathbb{Q}} \mathbb{R}$ with $\mathbb{C}^s$ as real algebras. Give $\mathbb{C}^s$ the usual real Euclidean metric $\sum_v |z_v|^2$ and its associated Lebesgue measure.

If $\omega_1, \dots, \omega_d$ is an integral basis, let S be the complex matrix of all d embeddings evaluated on that basis. The trace-pairing matrix is $S^{\mathsf{T}}S$, so $|\det S| = |D_K|^{1/2}$, where D_K is the field discriminant. Let M be the real matrix whose rows are the real and imaginary parts of the chosen embeddings. For each pair, the passage from those two real rows to the complex embedding and its conjugate uses the matrix

$$\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix},$$

whose determinant has absolute value 2. Hence $|\det S| = 2^s |\det M|$, and therefore

$$\operatorname{covol}(\iota(\mathcal{O}_K)) = 2^{-s} |D_K|^{1/2}. \tag{119}$$

In particular the image is a full Euclidean lattice. Discreteness can also be seen directly from the invertibility of M: it is the image of $\mathbb{Z}^d$ under an invertible real-linear map.

The factor 2^{-s} is not optional. It comes from using ordinary Lebesgue measure on each copy of $\mathbb{C}$. A convention that gives the complex plane twice its ordinary measure would remove this factor from the covolume and move it elsewhere in the height formulas.

27.2 Real dimensions and Hermitian Gram matrices

Let V be an n-dimensional K-vector space. At each infinite place choose a positive-definite Hermitian form h_v on $V_v = V \otimes_{K,\sigma_v} \mathbb{C}$. The direct sum $V_{\mathbb{R}} = \bigoplus_v V_v$, viewed as a real vector space, has dimension $dn = 2sn$. Its Euclidean metric is the sum of the real parts of the Hermitian forms.

If H is a positive-definite Hermitian matrix on $\mathbb{C}^n$, its real metric matrix has determinant $(\det H)^2$. To prove this, choose a unitary basis diagonalizing H with positive eigenvalues $\lambda_1, \dots, \lambda_n$. Each complex coordinate is a real plane, and its two real directions both have metric coefficient λ_j. Thus the real determinant is $\prod_j \lambda_j^2$. Consequently the volume scaling factor is $\det H$, rather than $(\det H)^{1/2}$.

Equivalently, a complex-linear map B has real determinant $|\det_{\mathbb{C}} B|^2$. This follows from the same calculation applied to B^*B, or from the block real matrix of its real and imaginary parts. These squares account for the exponent two in number-field operator heights.

27.3 Free lattices and normalized height

Suppose $e_1, \dots, e_n$ is a K-basis and $L_0 = \bigoplus_j \mathcal{O}_K e_j$. At the place v, let $G_v = (h_v(e_i, e_j))_{ij}$ be its Hermitian Gram matrix. Combining (119) in each coordinate with the metric volume factor gives

$$\operatorname{covol}(L_0) = 2^{-sn} |D_K|^{n/2} \prod_{v \mid \infty} \det G_v. \tag{120}$$

All determinants on the right are positive real numbers. For any full $\mathcal{O}_K$-lattice $L \subset V$, define

$$H(L) = 2^{sn}|D_K|^{-n/2}\,\mathrm{covol}(L), \qquad \mathrm{size}(L) = H(L)^{1/(dn)}.$$

The first quantity is the normalized height, and the second is its geometric mean per real dimension. For the free lattice above,

$$H(L_0) = \prod_v \det G_v. \tag{121}$$

This formula also applies to a free lattice in a lower-dimensional K-subspace, with n replaced by its K-rank and the induced metrics.

A full lattice need not be free as an $\mathcal{O}_K$-module. The covolume definition remains meaningful and avoids choosing a free basis that may not exist. When a free sublattice $L_0 \subset L$ of finite index is chosen, volume gives

$$H(L_0) = [L : L_0]H(L). \tag{122}$$

Thus a free frame can compute the height of L only after its index has been included.

27.4 Normalized slopes

Let $\mu_{\mathbb{R}}(L) = -\log \mathrm{covol}(L)/(dn)$ denote the slope of the underlying real lattice. Define the Hermitian normalized slope by

$$\widehat{\mu}(L) = -2\log \mathrm{size}(L).$$

Substitution yields the exact identity

$$\widehat{\mu}(L) = 2\mu_{\mathbb{R}}(L) - \log 2 + \frac{\log|D_K|}{d}. \tag{123}$$

The correction is independent of the rank. Therefore the same affine change converts maximal and minimal slopes, provided the extrema on the real side are taken over $\mathcal{O}_K$-submodules and their saturated quotients. If one instead takes the maximum over all $\mathbb{Z}$-sublattices, the collection being optimized is larger and the number may change.

The term $\log|D_K|/d$ is the logarithm of the root discriminant. It can vary with the field in an arithmetic family. A uniform inequality for raw Euclidean slopes is therefore not automatically a uniform inequality for normalized slopes unless the discriminant correction is included consistently on both sides.

For the lattice $\mathcal{O}_K^n$ with standard metrics, each G_v is the identity, so $H = 1$, size is one, and $\widehat{\mu} = 0$. The underlying real slope is nevertheless generally nonzero. Formula (123) makes these two statements compatible.

27.5 Fractional ideals as rank-one examples

Let I be a nonzero integral ideal. Since $I \subset \mathcal{O}_K$ has index $N(I) = |\mathcal{O}_K/I|$, its covolume is $N(I)2^{-s}|D_K|^{1/2}$. With standard metrics,

$$H(I) = N(I), \qquad \widehat{\mu}(I) = -\frac{2}{d}\log N(I).$$

For a fractional ideal, multiply by an integer to reduce to the integral case. The identity extends with the usual positive rational ideal norm.

If $I = a\mathcal{O}_K$ is principal, the map of real realizations given by multiplication by a has determinant $\prod_v |\sigma_v(a)|^2 = |N_{K/\mathbb{Q}}(a)|$. Thus

$$H(a\mathcal{O}_K) = |N_{K/\mathbb{Q}}(a)|.$$

A unit has norm ± 1 and preserves the covolume even though it can expand some archimedean directions and contract others. Covolume sees the product of these changes, while the operator norm sees the largest one at each place. This is a simple instance of the distinction between average slope and extreme behavior.

27.6 Operator heights at the infinite places

Let $T : V \to W$ be a nonzero K-linear map between Hermitian spaces. Write $C_v = \|T_v\|_{\mathrm{op}} > 0$ at each infinite place. Define

$$h(T) = \frac{1}{s} \sum_{v|\infty} \log C_v^2. \tag{124}$$

If $L \subset V$ is a lattice and $T(L)$ has K-rank q, give $L/(L \cap \ker T)$ the quotient metrics. At each place, the induced map on its q-dimensional complex space has determinant of absolute value at most C_v^q. The real volume factor is at most C_v^{2q}. The discriminant and power-of-two factors are the same for the source quotient and the image, since their K-ranks agree. Hence

$$H(T(L)) \le H(L/(L \cap \ker T)) \prod_v C_v^{2q}.$$

Taking $1/(dq)$ powers and then -2 logarithms gives

$$\widehat{\mu}(T(L)) \ge \widehat{\mu}(L/(L \cap \ker T)) - h(T). \tag{125}$$

Thus a nonzero integral map $T : L \to M$ satisfies

$$\widehat{\mu}_{\min}(L) \le \widehat{\mu}_{\max}(M) + h(T).$$

This is the normalized number-field version of (117).

The average of logarithmic norms is essential. Replacing it by the logarithm of an average norm gives a different quantity. Jensen's inequality relates the two in one direction, but that replacement usually loses information and may alter a sharp comparison constant.

27.7 Composition, scalar multiplication, and dual maps

For nonzero composable maps with nonzero composite, $\|(ST)_v\| \le \|S_v\|\|T_v\|$, so

$$h(ST) \le h(S) + h(T).$$

If $a \in K^\times$, then

$$h(aT) = h(T) + \frac{1}{s} \sum_v \log |\sigma_v(a)|^2 = h(T) + \frac{2}{d} \log |N_{K/\mathbb{Q}}(a)|.$$

An algebraic unit therefore leaves this archimedean average height unchanged. Multiplication by a nonunit changes the height, and its finite-place valuations supply the compensating terms in a global product formula.

The dual map has the same operator norm at each place, provided the dual metrics are used. Thus $h(T^\vee) = h(T)$. Combining this with slope duality converts an inequality about maximal slopes and injective maps into one about minimal slopes and surjective dual maps. The integral dual lattice must be chosen with the relevant coefficient ring; trace duals over $\mathbb{Z}$ and linear duals over $\mathcal{O}_K$ need not be identical.

27.8 Local lengths and residue degrees

Let $\mathfrak{p}$ be a finite prime of K above the rational prime p. Write $\mathcal{O}_{K,\mathfrak{p}}$ for the discrete valuation ring and $k_\mathfrak{p}$ for its residue field, of cardinality $p^{f_\mathfrak{p}}$. If A is a finite-length module over this ring, a composition series has successive factors $k_\mathfrak{p}$. Therefore

$$\log |A| = f_\mathfrak{p} \log p \, \operatorname{length}_{\mathcal{O}_{K,\mathfrak{p}}} A = \log N\mathfrak{p} \, \operatorname{length}_{\mathcal{O}_{K,\mathfrak{p}}} A. \tag{126}$$

Viewed as a $\mathbb{Z}_p$-module, A has length $f_\mathfrak{p} \operatorname{length}_{\mathcal{O}_{K,\mathfrak{p}}} A$. There is no additional ramification factor in this formula.

A ramification factor enters when a power of p is converted into a power of a local uniformizer ϖ. If $p = u\varpi^{e_\mathfrak{p}}$ with u a unit, then

$$\operatorname{length}_{\mathcal{O}_{K,\mathfrak{p}}}(\mathcal{O}_{K,\mathfrak{p}}/p^n) = e_\mathfrak{p} n,$$

and its $\mathbb{Z}_p$-length is $e_\mathfrak{p} f_\mathfrak{p} n$. Keeping these two conversions separate prevents the ramification index from being inserted twice.

27.9 Global indices from local modules

Let $L_0 \subset L$ be full lattices of finite index. The quotient is a finite $\mathcal{O}_K$-module and decomposes into its primary pieces. Localization at $\mathfrak{p}$ retains exactly the corresponding primary piece. Applying (126) and multiplying cardinalities gives

$$\log[L : L_0] = \sum_\mathfrak{p} \log N\mathfrak{p} \, \operatorname{length}_{\mathcal{O}_{K,\mathfrak{p}}}(L_\mathfrak{p}/(L_0)_\mathfrak{p}). \tag{127}$$

Only finitely many terms are nonzero because the quotient is finite. Substituting this identity into (122) yields

$$\log H(L_0) - \log H(L) = \sum_\mathfrak{p} \log N\mathfrak{p} \, \operatorname{length}(L_\mathfrak{p}/(L_0)_\mathfrak{p}).$$

This is the exact bridge between local divisibility and global height. A bound on a weighted sum of local obstruction lengths becomes a bound on the difference of two logarithmic heights once the local quotients have been identified with those obstructions or their subquotients.

27.10 Determinants and the product formula

For $a \in K^\times$, factor the fractional ideal $a\mathcal{O}_K = \prod_\mathfrak{p} \mathfrak{p}^{v_\mathfrak{p}(a)}$. Taking ideal norms gives

$$\log |N_{K/\mathbb{Q}}(a)| = \sum_\mathfrak{p} v_\mathfrak{p}(a) \log N\mathfrak{p}.$$

The same field norm is the product over all complex embeddings, so

$$\sum_v 2 \log |\sigma_v(a)| - \sum_\mathfrak{p} v_\mathfrak{p}(a) \log N\mathfrak{p} = 0. \tag{128}$$

This is the product formula in the present normalization. For an invertible K-linear map between equal-rank spaces, apply it to its determinant. The archimedean determinant change and the finite index change are two expressions for the same rational scaling.

An operator norm inequality replaces a determinant by the product of upper bounds for singular values. The resulting inequality is weaker than the exact product formula but has the advantage of applying to maps of nonmaximal rank after passing to their image. The slope argument is a systematic way of making that passage while retaining control of all possible kernels.

27.11 A concrete imaginary quadratic calculation

Take $K = \mathbb{Q}(i)$, with discriminant -4, so $s = 1$ and $d = 2$. The standard lattice $\mathbb{Z}[i] \subset \mathbb{C}$ has covolume one, agreeing with $2^{-1}\sqrt{4} = 1$. For the ideal $(1+i)$, the covolume is two and $H((1+i)) = 2$. Its normalized slope is $-\log 2$. Multiplication by $1 + i$ has complex operator norm $\sqrt{2}$, so its height is $\log 2$.

The map $\mathbb{Z}[i] \to (1 + i)\mathbb{Z}[i]$ is surjective onto its target lattice. The normalized slope inequality reads $0 \le -\log 2 + \log 2$, an equality. If the same map is instead viewed as an endomorphism of $\mathbb{Z}[i]$, its image has index two. The operator height is unchanged, but the integral index now appears in the target comparison. The example shows why a map, its image lattice, and its ambient target lattice must be distinguished even in rank one.

27.12 Applying a cohomological budget

Suppose a comparison supplies finite local modules $C_{\mathfrak{p}}$ and an injection

$$L_{\mathfrak{p}}/(L_0)_{\mathfrak{p}} \hookrightarrow C_{\mathfrak{p}}$$

for every finite prime. Then

$$\log[L : L_0] \le \sum_{\mathfrak{p}} \log N\mathfrak{p} \ \operatorname{length} C_{\mathfrak{p}}.$$

If the right-hand modules are themselves subquotients of a direct sum of cohomological obstruction groups, length additivity yields the next comparison. Every multiplicity in that direct sum multiplies the corresponding length. Every change of coefficient ring contributes its residue-degree factor. These are numerical parts of the theorem, not merely notation.

The arithmetic window hypothesis can bound the final weighted sum. The resulting inequality then feeds into (125) through the height-index identity. This passage requires actual local comparison homomorphisms. Equality of the orders of two groups, or an informal similarity between their definitions, does not supply the injections used above.

28 From a family of restriction maps to a slope bound

We now combine the spectral and arithmetic arguments. The mechanism has two independent parts. A slope estimate selects large sublattices in the sources of the restriction maps. A frame estimate shows that the images of those sublattices still span most of the target. The common maps are what make these two conclusions interact.

28.1 A normalized family of maps

Let H be a complex Hilbert space of dimension $D > 0$, and let V_i be complex Hilbert spaces of dimensions n_i, indexed by a finite set. Put $N = \sum_i n_i > 0$. Let $A_i : V_i \to H$ be linear maps. Define

$$K_i = A_i^* A_i, \qquad Q = \sum_i A_i A_i^*, \qquad t = \operatorname{Tr} Q.$$

The two quantities

$$E = \sum_i \|K_i - I_{V_i}\|_{\mathrm{HS}}^2, \qquad \Delta = \frac{D \operatorname{Tr}(Q^2)}{(\operatorname{Tr} Q)^2} - 1$$

measure different defects. The first measures failure of the individual maps to preserve norms. The second measures failure of the aggregate operator Q to be a scalar multiple of the identity. The latter is defined when $t > 0$ and is nonnegative by Cauchy–Schwarz on the D eigenvalues, including any zeros.

Individual near-isometries do not imply aggregate isotropy. For example, if all V_i are one-dimensional and all maps have the same unit image vector, then $E = 0$ while Q has rank one. Conversely Q can be a scalar operator even if the individual maps have very different norms. The arithmetic application needs control of both defects.

28.2 Trace control from the diagonal defect

Suppose $E \le eN$ with $e \ge 0$. Since $|\operatorname{Tr}(K_i - I)| \le \sqrt{n_i}\|K_i - I\|_{\mathrm{HS}}$, Cauchy–Schwarz over the index i gives

$$|t - N| \le \sqrt{\sum_i n_i}\sqrt{\sum_i \|K_i - I\|_{\mathrm{HS}}^2} \le \sqrt{e}\, N.$$

Thus

$$(1 - \sqrt{e})N \le t \le (1 + \sqrt{e})N. \tag{129}$$

In particular $t > 0$ if $e < 1$. The square root is unavoidable at this level of information because a Hilbert–Schmidt square controls a linear trace through Cauchy–Schwarz.

The estimate is invariant under decomposing one V_i into orthogonal pieces only if the corresponding defect is recomputed. Off-diagonal blocks of K_i contribute to its Hilbert–Schmidt norm and cannot be discarded by changing the indexing.

28.3 Removing source directions

Choose subspaces $U_i \subset V_i$ and let P_i be the orthogonal projection onto $U_i^{\perp}$. Put

$$b = \sum_i \operatorname{codim}_{V_i} U_i.$$

The total energy carried by the discarded directions is

$$L = \sum_i \|A_i P_i\|_{\mathrm{HS}}^2 = \sum_i \operatorname{Tr}(P_i K_i).$$

Since $\operatorname{Tr} P_i = \operatorname{codim} U_i$ and $\|P_i\|_{\mathrm{HS}} = \sqrt{\operatorname{codim} U_i}$,

$$L \le b + \sum_i \sqrt{\operatorname{codim} U_i}\|K_i - I\|_{\mathrm{HS}} \le b + \sqrt{bE}.$$

If $b \le \varepsilon N$ and $E \le eN$, this becomes

$$L \le N(\varepsilon + \sqrt{\varepsilon e}). \tag{130}$$

A rank loss is converted into an energy loss using the diagonal near-isometry estimate. Without that estimate, a single removed direction could carry nearly all of the operator's energy.

28.4 Isotropy turns energy loss into codimension

Let $W \subset H$ contain every $A_i(U_i)$, and let P be the orthogonal projection onto $W^{\perp}$. Since PA_i vanishes on U_i,

$$\ell := \operatorname{Tr}(PQ) = \sum_i \|PA_i\|_{\mathrm{HS}}^2 = \sum_i \|PA_iP_i\|_{\mathrm{HS}}^2 \le L.$$

Write $r = \operatorname{codim}_H W$. If $P_W = I - P$, then $\operatorname{Tr}(P_W Q) = t - \ell$. The Hilbert–Schmidt Cauchy–Schwarz inequality and $\|P_W\|_{\mathrm{HS}}^2 = D - r$ give

$$(t - \ell)^2 \le (D - r)\operatorname{Tr}(Q^2).$$

Suppose $\operatorname{Tr}(Q^2) \le (1 + \eta)t^2/D$ with $\eta \ge 0$. Then

$$\frac{r}{D} \le 1 - \frac{(1 - \ell/t)^2}{1 + \eta}. \tag{131}$$

Here $0 \le \ell \le t$ because Q is positive semidefinite and P is an orthogonal projection. Using $1 - (1 - z)^2/(1 + \eta) \le \eta + 2z$ for $0 \le z \le 1$, we obtain

$$r \le D\left(\eta + \frac{2L}{t}\right). \tag{132}$$

The exact expression (131) is preferable when constants matter and L/t is known accurately. The simpler expression is convenient for collecting errors that tend to zero. Either bound may exceed D; in that case the trivial estimate $r \le D$ should be retained instead.

28.5 A corank theorem

Theorem 28.1 (Corank after restriction of the sources)**.** *Assume*

$$\sum_i \|A_i^* A_i - I\|_{\mathrm{HS}}^2 \le eN, \qquad \operatorname{Tr}(Q^2) \le \frac{1 + \eta}{D}(\operatorname{Tr} Q)^2,$$

where $0 \le e \le 1/4$ *and* $\eta \ge 0$*. If* $\sum_i \operatorname{codim} U_i \le \varepsilon N$ *with* $0 \le \varepsilon \le 1$*, then*

$$\operatorname{codim}_H \sum_i A_i(U_i) \le D\big(\eta + 4\varepsilon + 4\sqrt{e}\big).$$

A sharper bound replaces the final two terms by $2(\varepsilon + \sqrt{\varepsilon e})/(1 - \sqrt{e})$*.*

Proof. Take $W = \sum_i A_i(U_i)$. Equations (129), (130), and (132) give the sharper bound. Since $1 - \sqrt{e} \ge 1/2$ and $\sqrt{\varepsilon} \le 1$, it is at most the stated simpler expression. □

The theorem does not require pairwise orthogonality of the images. It requires the single aggregate second-moment inequality for Q. In practice that inequality is often proved by expanding

$$\operatorname{Tr}(Q^2) = \sum_{i,j} \|A_i^* A_j\|_{\mathrm{HS}}^2,$$

which returns to the diagonal and off-diagonal kernel estimates of Section 16.

28.6 An auxiliary target space

Sometimes the available kernel acts on a larger Hilbert space $\widetilde{H} = H \oplus H'$ while the arithmetic map lands in H. Let $\widetilde{A}_i : V_i \to \widetilde{H}$ be maps and put $A_i = P_H \widetilde{A}_i$. If $W \subset H$ contains all $A_i(U_i)$, then $W \oplus H'$ contains all $\widetilde{A}_i(U_i)$. Its codimension in $\widetilde{H}$ equals the codimension of W in H. Applying Theorem 28.1 upstairs yields

$$\operatorname{codim}_H W \le \dim \widetilde{H}\,(\eta + 4\varepsilon + 4\sqrt{e}). \tag{133}$$

The dimension on the right is that of the larger space. Replacing it by $\dim H$ requires a separate estimate comparing these dimensions.

This formulation is useful when a complete projection kernel has a clean spectral identity but the arithmetic construction sees only a specified summand. The projection to that summand is part of the common map and must be included in the comparison.

28.7 Scalar normalizations

Suppose the original maps are B_i and positive scalars c_i are used to normalize them. Set $A_i = \sqrt{c_i} B_i$. Then

$$A_i^* A_i = c_i B_i^* B_i, \qquad Q = \sum_i c_i B_i B_i^*.$$

Since each scalar is nonzero, $A_i(U_i) = B_i(U_i)$ as subspaces. Thus the corank conclusion concerns the original images even though the spectral hypotheses concern the normalized maps.

Integral image lattices behave differently: multiplication by an arbitrary positive real scalar is not an operation on a number-field lattice. The scalars here are used only at an infinite place to estimate ranks. The arithmetic slope comparison must still use the original K-linear maps and their actual operator heights. Identifying the complex image subspaces is sufficient for the rank step and does not assert equality of integral lattices.

28.8 Submodules of large minimum slope

For a Euclidean lattice L and a real threshold θ, let $L_{\ge\theta}$ be the filtration step containing precisely the successive factors with slopes at least θ. Then

$$\mu_{\min}(L_{\ge\theta}) \ge \theta$$

when this step is nonzero, and the quotient $L/L_{\ge\theta}$ has maximal slope strictly less than θ when it is nonzero.

Every sublattice $A \subset L$ with $\mu_{\min}(A) \ge \theta$ is contained in $L_{\ge\theta}$. To prove this, compose its inclusion with the orthogonal quotient map to $L/L_{\ge\theta}$. The map has operator norm at most one. If it were nonzero, the slope inequality would give

$$\theta \le \mu_{\min}(A) \le \mu_{\max}(L/L_{\ge\theta}) < \theta,$$

a contradiction. Hence the map vanishes.

It follows that sums of sublattices of minimum slope at least θ also have minimum slope at least θ. One may prove this by applying the same quotient argument to the bottom filtration quotient of their sum: every summand maps trivially to a hypothetical quotient of slope less than θ, which is impossible if the summands generate the sum. This proof avoids a factor from the norm of the addition map on a direct sum.

28.9 Image slopes under bounded maps

Let $B : L \to M$ be an integral map with operator norm at most e^h. For every quotient of $B(L)$ by a saturated sublattice, the composite map from L onto that quotient still has norm at most e^h, because orthogonal projection does not increase norms. The image determinant inequality therefore gives

$$\mu_{\min}(B(L)) \geq \mu_{\min}(L) - h$$

when the image is nonzero. The corresponding statement over a totally complex field is

$$\widehat{\mu}_{\min}(B(L)) \geq \widehat{\mu}_{\min}(L) - h(B),$$

with the normalized operator height of (124).

If a family $B_i : L_i \to M$ has $h(B_i) \leq h$ and one selects $L_{i,\geq s-\tau}$, every nonzero image has minimum slope at least $s - \tau - h$. Their sum therefore lies in $M_{\geq s-\tau-h}$. The sum can have a large index inside its saturation. For the rank conclusion that index is irrelevant; for a later height conclusion it must be bounded separately by the finite-place comparison.

28.10 Arithmetic selection followed by spectral comparison

We state the combined argument with its hypotheses separated by their mathematical role. Let K be totally complex, let L_i be Hermitian $\mathcal{O}_K$-lattices, and let $B_i : L_i \to M$ be integral maps. Write $n_i = \operatorname{rank}_K L_i$, $N = \sum_i n_i$, and use normalized slopes. Suppose the union of the source slope lists has average at least s and every source slope is at most $s + u$, with $u \geq 0$. Fix $\tau > 0$ and put

$$U_i = (L_i)_{\geq s-\tau} \otimes_{K,\sigma_v} \mathbb{C}$$

at one infinite place v. The slope-share estimate gives

$$\sum_i (n_i - \dim_{\mathbb{C}} U_i) \leq N \frac{u}{u+\tau}. \tag{134}$$

Assume also $h(B_i) \leq h$ for all i.

Suppose that, at the chosen place, normalized realizations of the same maps satisfy the frame hypotheses in an ambient space of dimension $\widetilde{D}$, possibly with an auxiliary summand as above. Then

$$\operatorname{rank}_K M - \operatorname{rank}_K M_{\geq s-\tau-h} \leq \widetilde{D}\left(\eta + 4\frac{u}{u+\tau} + 4\sqrt{e}\right). \tag{135}$$

Indeed, the selected image lattices lie in the displayed target filtration step by the image-slope argument. Their complex dimensions at the chosen place equal their K-ranks. Apply (133) and then enlarge the image to the filtration step.

The identity of the maps is indispensable. Spectral estimates for one family of maps and height estimates for another family with the same source and target dimensions do not imply (135). The image subspaces must be those of the maps to which the arithmetic estimate applies.

28.11 From corank to a scalar slope estimate

Let the target slope list be $a_1 \geq \cdots \geq a_D$, and suppose at least $D - r$ of these slopes are at least θ. If in addition all slopes are bounded below by m, then

$$\deg M \geq (D - r)\theta + rm.$$

Thus

$$\mu(M) \geq \theta - \frac{r}{D}(\theta - m) \qquad \text{when } m \leq \theta. \tag{136}$$

For normalized slopes, replace degree by the sum of the normalized slope list and obtain the same formula for the normalized average. A lower bound for the exceptional slopes is needed: a tiny proportion of very negative slopes can dominate the average.

Dually, an upper bound for all slopes together with a bound on the number above a threshold gives an upper bound for the average. The direction used in a particular arithmetic application depends on whether one compares the original lattice or its dual. Formula (116) converts the two forms exactly.

28.12 Small relative errors and full rank

If the right side of (135) is less than one, then the corank is zero because it is an integer. In a family with $\widetilde{D} = \widetilde{D}(x)$, the condition is

$$\widetilde{D}(x)\left(\eta(x) + 4\frac{u(x)}{u(x)+\tau(x)} + 4\sqrt{e(x)}\right) < 1.$$

The vanishing of each relative error as $x \to \infty$ is not sufficient if $\widetilde{D}(x)$ grows too quickly. For instance, an error e^{-x^a} dominates a polynomial dimension, but it does not dominate e^{2x^a}. Comparing logarithms gives the exact threshold: $\log \widetilde{D} = o(x^a)$ is sufficient when the relative error is bounded by a constant times e^{-x^a}.

The arithmetic comparison may need only a proportion-of-rank estimate, in which case the product with the dimension need not tend to zero. Distinguishing these two uses prevents a relative asymptotic statement from being silently promoted to exact surjectivity.

28.13 A numerical model

Take $\eta = e = 0$, so the maps are individually isometric and their aggregate operator is exactly scalar. If the arithmetic selection removes at most an ε fraction of the source directions, (131) gives

$$\frac{r}{D} \leq 1 - (1-\varepsilon)^2 = 2\varepsilon - \varepsilon^2.$$

The simpler theorem gives $r/D \leq 4\varepsilon$. The sharper bound shows the true first-order loss in this ideal model. The looser constant is useful when diagonal and isotropy errors must be combined without repeated optimization.

Now let $u > 0$ and choose $\tau = 9u$. Then the source slope-share bound removes at most one tenth of the source rank. Even in the ideal spectral case this only yields a positive-proportion image statement. To force an arbitrarily small exceptional proportion, one must take $\tau/u \to \infty$, while the retained threshold $s - \tau - h$ must remain adequate for the intended scalar inequality. This example makes the tradeoff in the arithmetic choice of parameters explicit.

29 Formal coordinates, integral jets, and character symmetries

The integral comparison uses local expansions of functions near an identity point. Their finite jets form modules of finite rank, while formal group laws describe how these expansions change under

addition. A finite-order symmetry can force many low-degree coefficients to vanish. This gives a concrete source of divisibility and rank savings at large primes. We develop the underlying algebra without assuming the existence of any particular geometric realization.

29.1 The order filtration on a power-series ring

Let R be a commutative ring and let $R[[X]] = R[[X_1, \dots, X_d]]$. For a multi-index $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$, write $X^\alpha = \prod_i X_i^{\alpha_i}$ and $|\alpha| = \sum_i \alpha_i$. Every series is uniquely $F = \sum_\alpha a_\alpha X^\alpha$. Let $I = (X_1, \dots, X_d)$. The ideal I^m consists exactly of the series whose coefficients of degree less than m vanish.

The order of a nonzero series is the smallest total degree of a nonzero coefficient. Thus $F \in I^m$ precisely when its order is at least m. One always has

$$\operatorname{ord}(FG) \geq \operatorname{ord} F + \operatorname{ord} G,$$

with equality over an integral domain. The inequality follows because a monomial of total degree less than the sum cannot arise in the product. Equality over a domain follows from the product of the two nonzero leading homogeneous polynomials.

The mth associated graded piece is

$$I^m / I^{m+1} \simeq \operatorname{Sym}_R^m(R^d).$$

Its standard basis consists of the monomials of degree m. Counting weak compositions of m into d parts gives

$$\operatorname{rank}_R(I^m/I^{m+1}) = \binom{m+d-1}{d-1}. \tag{137}$$

This count follows by placing $d-1$ separators among m indistinguishable units, with empty groups allowed.

29.2 Finite jets and their dimensions

The jet module of order less than t is $J_{<t} = R[[X]]/I^t$. It is free with basis the monomials of total degree less than t. Summing (137), or adjoining a slack coordinate to a tuple of total degree at most $t-1$, gives

$$\operatorname{rank}_R J_{<t} = \binom{t+d-1}{d}. \tag{138}$$

For $d = 1$ this is t; for $d = 2$ it is $t(t+1)/2$. For fixed d and large t,

$$\binom{t+d-1}{d} = \frac{t^d}{d!} + O_d(t^{d-1}).$$

If d varies, the implied constant cannot be treated as uniform. The product expression $\binom{t+d-1}{d} = \prod_{j=0}^{d-1}(t+j)/d!$ is often safer in a family with changing dimension.

Let R be a discrete valuation ring with residue field of cardinality q, and let ϖ be a uniformizer. Then

$$|J_{<t}/\varpi^n J_{<t}| = q^{n\binom{t+d-1}{d}}.$$

This is simply the cardinality of a free module of the displayed rank over R/ϖ^n. It makes explicit how jet order and dimension affect an integral counting estimate.

29.3 Substitution with zero constant term

Given $T_1, \dots, T_d \in I$, the substitution $F \mapsto F(T)$ is well-defined as a formal power series. For a fixed degree bound m, only finitely many monomials of F can contribute modulo I^m, since $T^\alpha \in I^{|\alpha|}$. Thus the coefficient of each monomial is a finite sum. This argument supplies formal convergence without using an analytic norm.

Substitution preserves the order filtration: $F \in I^m$ implies $F(T) \in I^m$. Its action on the degree-m graded piece depends only on the linear terms of the T_i. If these linear terms form a matrix A, the action is the degree-m symmetric power of the corresponding linear substitution.

The zero-constant-term condition cannot be dropped over an arbitrary ring. Substituting $T = 1 + X$ into an unrestricted formal series would make even the constant coefficient an infinite sum with no prescribed meaning. Analytic convergence can sometimes repair that problem, but it is a separate hypothesis.

29.4 An integral inverse-function theorem

Suppose $T = (T_1, \dots, T_d)$ has zero constant term and its linear matrix A lies in $\mathrm{GL}_d(R)$. Then there exists a unique formal tuple S with zero constant term such that

$$T(S(X)) = X, \qquad S(T(X)) = X.$$

The inverse has coefficients in R.

To construct it, first take the linear part $A^{-1}X$. Suppose S has been constructed through degree $m-1$ so that $T(S(X)) - X \in I^m$. Let E_m be the degree-m part of this error. Adding the homogeneous correction $-A^{-1}E_m$ to S changes the composition in degree m by exactly $-E_m$; all nonlinear terms of T contribute only higher-degree changes. This determines the next homogeneous term uniquely. Induction constructs a right inverse. The same construction gives a left inverse, and associativity of substitution shows that the two coincide.

Every step uses only A^{-1} and ring operations. Therefore a unit determinant ensures an integral inverse, while a merely nonzero determinant over the fraction field can introduce denominators. Consequently an integral coordinate change with unit Jacobian induces an automorphism on every finite jet module.

29.5 Determinants of symmetric powers

Let A be an invertible $d \times d$ matrix over a field. Its action on homogeneous polynomials of degree m has determinant

$$\det(\mathrm{Sym}^m A) = (\det A)^{\binom{m+d-1}{d}}. \tag{139}$$

For a diagonal matrix with diagonal entries $a_1, \dots, a_d$, the monomial X^α is multiplied by $\prod_i a_i^{\alpha_i}$. In the product over all $|\alpha| = m$, symmetry gives the same exponent for each a_i. The sum of those d exponents is $m\binom{m+d-1}{d-1}$, so each exponent is

$$\frac{m}{d}\binom{m+d-1}{d-1} = \binom{m+d-1}{d}.$$

The identity for a general matrix follows by triangularizing over an algebraic closure; the monomial basis then gives an upper triangular symmetric-power matrix with the same diagonal calculation.

For a formal coordinate change with linear matrix A, the determinant on $J_{<t}$ is the product over the graded pieces. Hence

$$\det(T^*|J_{<t}) = (\det A)^{\binom{t+d-1}{d+1}}. \tag{140}$$

The exponent is the sum of $\binom{m+d-1}{d}$ for $0 \le m < t$, with the degree-zero term interpreted as zero. The binomial summation identity gives the stated answer. For $d = 1$, it is $t(t-1)/2$, as one sees from the diagonal entries $1, a, a^2, \dots, a^{t-1}$.

29.6 Jet indices over a discrete valuation ring

Let R be a discrete valuation ring, and suppose a linear substitution has integral matrix A with nonzero determinant. Its action on $J_{<t}$ is injective and has a finite-length cokernel. Smith normal form and (140) give

$$\operatorname{length}_R \operatorname{coker}(T^*|J_{<t}) = \binom{t+d-1}{d+1} v_\varpi(\det A). \tag{141}$$

The same formula holds for an integral formal substitution with that linear part. Its matrix in the monomial basis is block triangular for the degree filtration, so its determinant is unchanged by higher-order terms. Over the fraction field the substitution is invertible, hence the cokernel is torsion and the determinant computes its length.

This identity explains why a determinant defect can be magnified by a large jet order. It also explains the advantage of a unit Jacobian: when $v_\varpi(\det A) = 0$, no such index loss occurs at any jet order. A high-order nonlinear term may affect individual coefficients, but it does not change this determinant calculation.

29.7 Formal group laws

A commutative formal group law of dimension d over R is a tuple $F(X, Y) \in R[[X, Y]]^d$ satisfying

$$F(X, 0) = X, \quad F(0, Y) = Y, \quad F(F(X, Y), Z) = F(X, F(Y, Z)), \quad F(X, Y) = F(Y, X).$$

Its linear term is $X + Y$. A formal inverse $\iota(X)$ satisfying $F(X, \iota(X)) = 0$ is constructed degree by degree: its linear term is $-X$, and at each later degree the unknown term appears with coefficient one. Thus it exists integrally and is unique.

The additive law is $F_a(X, Y) = X + Y$. In one variable the multiplicative law is

$$F_m(X, Y) = X + Y + XY,$$

which corresponds to multiplication of $1 + X$ and $1 + Y$. Its inverse is $-X/(1 + X) = -X + X^2 - X^3 + \cdots$. These two examples already show why a formal group law is more than its tangent space: both have the same linear term, but their nonlinear addition laws differ.

29.8 Formal logarithms over characteristic zero

Over a field of characteristic zero, a formal logarithm is a tuple $\ell(X) = X + \text{higher terms}$ satisfying

$$\ell(F(X, Y)) = \ell(X) + \ell(Y).$$

When such a logarithm is given, its formal inverse $\exp_F$ exists by the inverse-function construction, and

$$F(X, Y) = \exp_F(\ell(X) + \ell(Y)).$$

This identity lets one perform addition in logarithmic coordinates. It does not imply that ℓ or $\exp_F$ is integral.

For the multiplicative law,

$$\ell(X) = \log(1+X) = \sum_{n\ge 1} \frac{(-1)^{n+1}}{n} X^n, \qquad \exp_F(Z) = e^Z - 1 = \sum_{n\ge 1} \frac{Z^n}{n!}.$$

Both contain denominators, while the recovered group law $X + Y + XY$ is integral. Consequently the integrality of an addition law must be checked for the composite expression, not inferred from the individual logarithmic coordinate series.

29.9 Coefficient valuations of the exponential

Let v_p be normalized by $v_p(p) = 1$. Counting the multiples of p^j in $n!$ gives

$$v_p(n!) = \sum_{j\ge 1} \left\lfloor \frac{n}{p^j} \right\rfloor = \frac{n - s_p(n)}{p-1},$$

where $s_p(n)$ is the sum of the base-p digits of n. The second identity follows by writing $n = \sum_i a_i p^i$ and summing the finite geometric series for each digit. In particular $v_p(n!) \le (n-1)/(p-1)$ for $n \ge 1$.

Let $B(X)$ have zero constant term and coefficients in a valued field of residue characteristic p. If every coefficient has valuation at least $c \ge 1/(p-1)$, then every coefficient of $B^n/n!$ has valuation at least

$$nc - v_p(n!) \ge \frac{s_p(n)}{p-1} > 0.$$

For a fixed monomial only finitely many n contribute, because B has positive order. Hence $\exp B$ has integral coefficients and constant term one. The same is true of $\exp(-B)$, so multiplication by $\exp B$ is an integral automorphism of every finite jet module.

This is a sufficient condition, not a necessary one. Special cancellations may give an integral exponential even when the displayed uniform coefficient bound fails. In a geometric application, the integrality assertion can be more delicate than this elementary test.

29.10 Finite-order diagonal symmetries

Suppose a field of characteristic zero contains a primitive Mth root of unity ζ. Give the variables integer weights $w_1, \dots, w_d$ and consider the substitution $X_j \mapsto \zeta^{w_j} X_j$. If a series $B = \sum_\alpha b_\alpha X^\alpha$ is invariant, comparison of coefficients gives

$$(\zeta^{\sum_j w_j \alpha_j} - 1) b_\alpha = 0.$$

Thus

$$b_\alpha \ne 0 \quad \Longrightarrow \quad \sum_j w_j \alpha_j \equiv 0 \pmod{M}. \tag{142}$$

Conversely, support on these congruence classes makes the series invariant. The statement is purely coefficientwise and does not require analytic convergence.

The root of unity must have exact order M. Invariance under a root whose order is a proper divisor gives only the weaker congruence for that divisor. Likewise, invariance after reduction modulo p may lose information if the order becomes divisible by the characteristic. The characteristic-zero formulation avoids that ambiguity.

29.11 Alternating Frobenius weights

A useful pattern has an even number d of variables, indexed by $j = 0, \dots, d-1$, weights $w_j = (-1)^j p^j$, and $M = p^d - 1$. If a nonconstant monomial satisfies (142), then its total degree is a positive multiple of $p+1$. Indeed, $p \equiv -1 \pmod{p+1}$, so $(-1)^j p^j \equiv 1 \pmod{p+1}$. Since d is even, $p+1$ divides $p^d - 1$. Reducing the resonance congruence modulo $p+1$ therefore gives

$$|\alpha| \equiv 0 \pmod{p+1}.$$

In particular, every invariant series with zero constant term belongs to I^{p+1}.

The lower bound is attained. For adjacent indices $j, j+1$ before the end of the cycle, the monomial $X_j^p X_{j+1}$ has weight zero. At the end of the cycle, $X_{d-1}^p X_0$ has weight $1 - p^d$, which is zero modulo M. All these monomials have degree $p+1$. Thus the symmetry alone forces a sharp prime-dependent gap in the Taylor expansion.

If d is odd, the divisibility $p + 1 \mid p^d - 1$ fails, and this argument cannot be used. The parity of the local orbit length is therefore a substantive hypothesis in this particular vanishing mechanism.

29.12 The two-variable case

For $d = 2$, the symmetry is $(X, Y) \mapsto (uX, u^{-p}Y)$ with $u^{p^2-1} = 1$. A monomial $X^a Y^b$ is invariant exactly when $a - pb \equiv 0 \pmod{p^2 - 1}$. Among mixed monomials of degree $p + 1$, the possibilities are precisely $X^p Y$ and XY^p. To see this, write $a + b = p + 1$. Then $a - pb = (p + 1)(1 - b)$, so divisibility by $(p - 1)(p + 1)$ means $b \equiv 1 \pmod{p - 1}$. With $1 \le b \le p$, this gives $b = 1$ or $b = p$.

A series satisfying the symmetry and vanishing on the coordinate axes therefore begins in the form

$$B(X, Y) = aX^p Y + bXY^p + \text{terms of degree at least } 2(p + 1).$$

The next allowed total degree is another multiple of $p + 1$. Additional functional identities may relate a and b or force them to vanish, but symmetry alone does not do so.

29.13 Additivity defects and multiplicative correction factors

For a series $B(X, Y)$ with $B(X, 0) = B(0, Y) = 0$, define its two additivity defects by

$$\Delta_1 B = B(X_1 + X_2, Y) - B(X_1, Y) - B(X_2, Y),$$
$$\Delta_2 B = B(X, Y_1 + Y_2) - B(X, Y_1) - B(X, Y_2).$$

Over characteristic zero, both defects vanish exactly when B is bilinear. For example, differentiating the first identity with respect to X_2 at zero shows that B is linear in X, and the second makes its coefficient linear in Y. The same conclusion follows by comparing binomial coefficients of monomials.

For a general B, exponentiation gives

$$e^{B(X_1+X_2,Y)} = e^{B(X_1,Y)} e^{B(X_2,Y)} e^{\Delta_1 B},$$

and the analogous identity in the second variable. Thus integrality of $e^{\Delta_1 B}$ and $e^{\Delta_2 B}$ provides integral correction factors for addition, even when e^B itself has denominators. This distinction is useful in local comparison formulas involving a phase and an integral unit factor.

For the term $aX^p Y$, the first defect contains $a\binom{p}{j} X_1^j X_2^{p-j} Y$ for $1 \le j < p$. Every binomial coefficient $\binom{p}{j}$ is divisible by p, since p is prime and the denominator $j!(p - j)!$ is prime to p. Thus the additivity defect gains a factor of p relative to the original coefficient. This is an elementary source of improved integrality in the correction factors.

29.14 Consequences for finite jets

If $B \in I^{p+1}$, then $e^B \equiv 1 \pmod{I^{p+1}}$ as a formal series over the fraction field. Hence an identity

$$F = Ue^B$$

with U integral implies that every coefficient of F of total degree at most p is integral and agrees with the corresponding coefficient of U. No assertion about integrality in higher degrees follows from this truncation alone.

In particular, for a fixed jet order t, all primes $p \geq t-1$ fall within the range in which the phase contributes nothing to $J_{<t}$. The remaining primes form a finite set. If t varies with the scale, that exceptional set also varies and its total contribution must still be estimated. Large-prime vanishing and a uniform global sum are separate steps.

When the local coordinates are changed by an integral substitution with unit Jacobian, these statements persist because the order filtration and integral jet lattices are preserved. If the Jacobian is not a unit, (141) quantifies the extra index that must be included in the local budget. This completes the elementary chain from formal symmetry to a finite-length comparison term.

30 Prime classes and corrected character expansions

The arithmetic comparison ultimately produces a scalar expression in prime classes. Finite Fourier analysis converts that expression into a combination of Dirichlet character sums. Two corrections are then needed: characters must be replaced by their primitive inducing characters, and prime-power contributions must be normalized consistently. We give exact formulas for both operations.

30.1 Indicators of cosets

Let A be a finite abelian group and $B \subset A$ a subgroup. Its annihilator is

$$B^\perp = \{\chi \in \widehat{A} : \chi(b) = 1 \text{ for every } b \in B\}.$$

Characters in $B^\perp$ are precisely the characters of A/B, so $|B^\perp| = [A : B]$. Orthogonality on this quotient gives

$$\mathbf{1}_{aB}(x) = \frac{1}{[A:B]} \sum_{\chi \in B^\perp} \overline{\chi(a)}\chi(x). \tag{143}$$

In particular the indicator of the subgroup B has nonnegative real Fourier coefficients, all equal to $1/[A : B]$ on its annihilator. For a translated coset, the coefficients acquire phases $\overline{\chi(a)}$ and need not be positive.

That distinction matters for the analytic pole argument. A positive function on a finite group need not have positive Fourier coefficients. For example, the indicator of a nontrivial singleton on a cyclic group has coefficients of different phases. Positivity of the weights in the analytic character family must therefore be checked after Fourier expansion, rather than inferred from nonnegativity of a prime-class indicator.

30.2 Two sign coordinates

Consider

$$G = \{\pm 1\} \times \{\pm 1\} \times A,$$

and a subgroup $A_0 \subset A$. Put $o = [A : A_0]$. A character trivial on A_0 has the unique form

$$\theta(a, b, x) = a^{\alpha} b^{\beta} \chi(x), \qquad \alpha, \beta \in \{0, 1\}, \quad \chi \in A_0^{\perp}.$$

There are therefore $4o$ such characters. Requiring triviality on the second sign coordinate leaves $2o$ characters. Requiring triviality on all of A as well leaves exactly two: the trivial character and $\xi(a, b, x) = a$.

These counts are exact and use no independence assumption beyond the specified direct-product decomposition. If a group is only presented as a quotient of this product, some characters may fail to descend and the count can decrease. A comparison using the factor $4o$ must either establish the product description or use $4o$ only as an upper bound.

Complex conjugation of characters preserves all these restrictions. Indeed, a value equal to 1 remains 1, and the sign characters are real. Thus the retained set can be made conjugation stable without changing the sign-coordinate conditions.

30.3 A positive weighted retained family

Let $\mathcal{K}$ be a set of primitive characters containing the trivial character of conductor one, and optionally a distinguished real character ξ. Consider weights

$$a_1 = 1, \qquad a_\xi = 1 \text{ when } \xi \in \mathcal{K}, \qquad a_\chi = \frac{1}{3o + 1} \text{ for all other members.}$$

Every weight is strictly positive, and $a_\chi \le 1$. The corresponding corrected sum is

$$G_{\mathcal{K}}(x) = E_1(x) + \sum_{\chi \in \mathcal{K} \setminus \{1\}} a_\chi \operatorname{Re} E_\chi(x).$$

At a zero ρ of any retained L-function, the symmetrized Mellin transform has multiplicity weight

$$M_\rho = \frac{1}{2} \sum_{\chi \in \mathcal{K}} a_\chi (m_\chi(\rho) + m_\chi(\overline{\rho})) > 0.$$

Thus no retained zero can disappear by cancellation of residues. The denominators $3o + 1$ affect the size of the residue but not whether it is zero.

Repeated primitive characters should be handled by adding their weights. A list of distinct finite-group characters can induce the same primitive character if the comparison map does not distinguish them. If injectivity of that map is available, the set description above applies directly. Otherwise the mathematically correct object is a weighted multiset, and the combined weights remain positive.

30.4 A residual family bound

Suppose at most $4o$ characters are discarded and their corrected expressions satisfy $|E_\chi(x)| \le x^{1/2+\varepsilon}$ at a given scale. Then any real sum of those expressions with coefficients of absolute value at most one satisfies

$$|\Sigma(x)| \le 4o\, x^{1/2+\varepsilon}.$$

The same conclusion holds if the sum uses their absolute values. This estimate uses no cancellation and is compatible with a selected set of scales or a common multiplicative interval.

The words “at a given scale” carry a quantifier requirement. To use the estimate on a window, every residual bound must hold throughout that same window. Bounds on separate windows for separate characters cannot be added unless those windows have a common interval of the required length.

30.5 Removing Euler factors exactly

Let χ modulo q be induced by a primitive character χ^* of conductor f. Set $P = \{p : p \mid q,\ p \nmid f\}$. Then, for $\operatorname{Re} s > 1$,

$$L(s,\chi) = L(s,\chi^*) \prod_{p\in P} (1 - \chi^*(p)p^{-s}).$$

Logarithmic differentiation gives

$$\frac{L'}{L}(s,\chi) = \frac{L'}{L}(s,\chi^*) + \sum_{p\in P} \frac{\chi^*(p)\log p}{p^s - \chi^*(p)}. \tag{144}$$

Every factor in the finite product is nonzero when $\operatorname{Re} s > 0$, since $|\chi^*(p)p^{-s}| = p^{-\operatorname{Re} s} < 1$. The finite Euler factors therefore do not change the zero set in that half-plane.

They do change the logarithmic derivative at one. If the primitive character is nonprincipal, put

$$c_P = \sum_{p\in P} \frac{\chi^*(p)\log p}{p - \chi^*(p)}.$$

Then the correction constants for χ and χ^* differ by c_P. Failing to alter the correction constant would leave an unwanted term of size x in the comparison of their smoothed sums.

30.6 The exact corrected ramp discrepancy

For $x \geq 1$, put

$$R_\chi(x) = \sum_{n\leq x} \Lambda(n)\chi(n)(x/n - 1).$$

For a nonprincipal character, use $E_\chi(x) = R_\chi(x) + x(L'/L)(1,\chi)$. Fix $p \in P$, set $z = \chi^*(p)$, and let $k_p = \lfloor \log x/\log p \rfloor$. The contribution of the missing Euler factor to $R_{\chi^*} - R_\chi$ is

$$\log p \sum_{k=1}^{k_p} z^k (xp^{-k} - 1).$$

Subtracting the correction $x\log p\, z/(p-z)$ and summing the geometric series gives

$$E_{\chi^*}(x) - E_\chi(x) = -\sum_{p\in P} \log p \left(\sum_{k=1}^{k_p} z^k + \frac{x(z/p)^{k_p+1}}{1 - z/p} \right). \tag{145}$$

The first finite sum is zero when $k_p = 0$. Since $x < p^{k_p+1}$ and $|1 - z/p| \geq 1 - 1/p \geq 1/2$,

$$|E_{\chi^*}(x) - E_\chi(x)| \leq |P| \log x + 2\sum_{p\in P} \log p. \tag{146}$$

Thus the corrected discrepancy is logarithmic in x for fixed modulus.

For principal characters, define the corrected expression at modulus q by

$$E_{0,q}(x) = R_{0,q}(x) - x\log x + \left(1 + \gamma + \sum_{p\mid q} \frac{\log p}{p-1}\right) x.$$

The same formula and bound hold relative to E_1, with $z = 1$. This extension agrees with the conductor-one normalization and shows explicitly how the principal local factors are accounted for.

30.7 Prime powers and the square-root scale

Let χ be any Dirichlet character and isolate the contribution of prime powers of exponent at least two:

$$Q_\chi(x) = \sum_{\substack{p^k \le x \\ k\ge 2}} \log p\, \chi(p)^k (x/p^k - 1).$$

The series

$$c_\chi^{\mathrm{pow}} = \sum_p \sum_{k\ge 2} \frac{\log p\, \chi(p)^k}{p^k}$$

converges absolutely, since its absolute value is bounded by $\sum_{n\ge 2} \log n/[n(n-1)]$. The normalized prime-power contribution satisfies

$$Q_\chi(x) - x c_\chi^{\mathrm{pow}} = O(\sqrt{x}\log(2x)), \tag{147}$$

with an absolute implied constant.

To prove it, subtract the full convergent series:

$$Q_\chi(x) - x c_\chi^{\mathrm{pow}} = -\sum_{\substack{p^k \le x \\ k\ge 2}} \log p\, \chi(p)^k - x \sum_{\substack{p^k > x \\ k\ge 2}} \frac{\log p\, \chi(p)^k}{p^k}.$$

For the first sum, each prime is at most $\sqrt{x}$ and its number of allowed exponents times $\log p$ is at most $\log x$. There are at most $\sqrt{x}$ such primes, so the contribution is at most $\sqrt{x}\log x$.

For the second sum, first take $p \le \sqrt{x}$. With $k_p = \lfloor \log x/\log p \rfloor \ge 2$, the geometric tail gives

$$x \sum_{k>k_p} \frac{\log p}{p^k} = \frac{x \log p\, p^{-k_p-1}}{1-p^{-1}} \le 2\log p.$$

Summing gives at most $2\sqrt{x}\log x$. For $p > \sqrt{x}$, all exponents $k \ge 2$ are in the tail, and their contribution is bounded by

$$2x \sum_{n>\sqrt{x}} \frac{\log n}{n^2} \ll \sqrt{x}\log(2x),$$

using the integral of $(\log t)/t^2$. This proves (147) without a prime number theorem.

The linear term $x c_\chi^{\mathrm{pow}}$ is necessary because the ramp weights contain x/p^k. An uncorrected assertion that all higher prime powers contribute only a square-root error would be false. After the linear term is removed, the elementary square-root scale emerges from the inequality $p \le \sqrt{x}$ for $k \ge 2$.

30.8 Prime-class sums

Let G be a finite abelian Galois group and let $a(p) \in G$ denote the Frobenius element of an unramified rational prime. For a function $F : G \to \mathbb{C}$, its finite Fourier expansion is

$$F(g) = \sum_{\theta \in \widehat{G}} \widehat{F}(\theta)\theta(g), \qquad \widehat{F}(\theta) = \frac{1}{|G|} \sum_{g \in G} F(g)\overline{\theta(g)}.$$

Therefore every finite smoothed prime sum satisfies the exact identity

$$\sum_{\substack{p \le x \\ p\ munramified}} \log p\, F(a(p))(x/p - 1) = \sum_{\theta} \widehat{F}(\theta) \sum_{\substack{p \le x \\ p\ munramified}} \log p\, \theta(a(p))(x/p - 1).$$

All sums are finite, so no convergence issue occurs. Identifying $\theta(a(p))$ with a Dirichlet character value requires the appropriate abelian reciprocity description of the field. Once that identification is given, the preceding prime-power and Euler-factor formulas convert the right side into corrected primitive character expressions.

Ramified primes are a finite exceptional set for a fixed field. Their raw ramp contribution can have size x, so they must be included in the correction constants before being treated as a small error. The same cancellation seen in (145) is the model for that bookkeeping.

30.9 Subgroup proportions versus prime densities

The proportion $|C|/|G|$ of a subset $C \subset G$ is a finite group calculation. It is the constant Fourier coefficient of $\mathbf{1}_C$. Interpreting it as an asymptotic density of primes is an additional analytic theorem. The exact finite Fourier identity above needs only the group proportion and makes no density assertion.

This distinction is useful in a conditional argument about zeros. One can derive an identity among smoothed character sums without already assuming an equidistribution theorem of the strength one hopes to prove. All analytic information then enters through explicit bounds on the nontrivial Fourier modes, rather than through an unstated density assumption.

30.10 The final finite-dimensional bookkeeping

Suppose a comparison has fixed a retained positive family, a residual family of cardinality at most $4o$, and an exact identity

$$\mathcal{B}(x) = (3o + 1)(4G_{\mathcal{K}}(x) - \Sigma(x)).$$

Assume all local Euler and prime-power corrections have already been included in these definitions. Then the scalar upper bound for $\mathcal{B}$ and the absolute bound for Σ imply an upper bound for $G_{\mathcal{K}}$ by division by the positive constant $4(3o + 1)$. The positivity of that divisor determines the direction of the inequality.

If instead the comparison is initially written with imprimitive characters, (146) contributes only a fixed logarithmic error after normalization. For every fixed $\varepsilon > 0$, such an error is eventually dominated by x^{ε}. Formula (147) contributes an error of size at most $x^{1/2+\varepsilon}$ for sufficiently large x. For a varying modulus, the dependence on the missing prime set must also be bounded, as detailed in Section 31. Under that additional control, these elementary estimates explain why finite local changes and higher prime powers can be absorbed into the exponent losses of the final window bound.

31 Scale calculus and simultaneous selection

The last arithmetic implication combines estimates proved at several scales. Their constants may depend on a fixed character and on a small positive exponent, but they must remain fixed as the main scale tends to infinity. We make those dependencies explicit and prove the selection principle that allows several comparison statements to be used on one family.

31.1 Three levels of growth

Fix $\varepsilon > 0$ and put

$$y = x^{\varepsilon}, \qquad b = e^{y}, \qquad x > 1.$$

There are three distinct growth levels: powers of $\log x$, powers of x, and exponentials of x^{ε}. For fixed $A, B, c > 0$,

$$(\log x)^{A} = o(x^{B}), \qquad x^{A} = o(e^{cx^{\varepsilon}}).$$

For the first, write $u = \log x$ and compare u^{A} with e^{Bu}. The exponential series has a term of degree greater than A, giving the limit. For the second, take logarithms and use $A \log x = o(x^{\varepsilon})$.

These limits hold with ε fixed. They are not uniform as $\varepsilon \downarrow 0$. For instance, if $\varepsilon(x) = 1/\log x$, then $x^{\varepsilon(x)} = e$ and $e^{x^{\varepsilon(x)}}$ is constant. Therefore the order of choices must be: first fix ε, then increase the scale x.

31.2 Explicit absorption of a logarithm

For $u \geq 0$ and an integer $m \geq 1$, the exponential series gives $e^{u} \geq u^{m}/m!$. Taking $u = \delta \log x$ yields

$$(\log x)^{m} \leq \frac{m!}{\delta^{m}} x^{\delta}.$$

To obtain a bound with coefficient one, use a smaller exponent first. For example,

$$(\log x)^{m} \leq C_{m,\delta} x^{\delta/2} \leq x^{\delta}$$

when $x \geq C_{m,\delta}^{2/\delta}$. This supplies a concrete threshold rather than only an asymptotic symbol. A finite sum of terms $C_j x^{a_j} (\log x)^{m_j}$ is therefore at most $x^{\max a_j + \delta}$ after increasing the threshold, with all parameters fixed before x varies.

The same reasoning absorbs a fixed positive constant C into x^{δ} once $x \geq \max(1, C^{1/\delta})$. The threshold can be very large when δ is small. The window theorem permits this because it requires arbitrarily late windows, rather than a uniform starting point for every ε.

31.3 Subpower errors in an exponential scale

Suppose $\delta(x) \to 0$. Then for every $\eta > 0$,

$$b^{\delta(x)} \leq b^{\eta}$$

for sufficiently large x. This is subpower growth relative to b. It does not imply subpower growth relative to x. For example, $\delta(x) = x^{-\varepsilon/2}$ tends to zero but

$$b^{\delta(x)} = e^{x^{\varepsilon/2}},$$

which grows faster than every fixed power of x.

On the other hand, a normalization by a fixed positive power of b makes such an error small. For every $c > 0$, eventually $\delta(x) \le c/2$, so

$$\frac{b^{\delta(x)}}{b^c} \le e^{-cx^{\varepsilon}/2}.$$

This decays faster than every negative power of x. Thus a term described as $b^{o(1)}$ can be harmless after division by b, while being far too large to absorb directly into x^{ε}. The normalization must be displayed before the error is simplified.

31.4 Rounding the cyclotomic level

Let

$$a(x) = \left\lceil \frac{(2+8\varepsilon)\log x}{\log 2} \right\rceil.$$

Then

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

The lower inequality uses $a(x)$ at least the real argument of the ceiling. The strict upper inequality uses $a(x)$ less than that argument plus one. If a field degree is at most $4r2^{a(x)}$, it is therefore less than $8rx^{2+8\varepsilon}$.

The function $a(x)$ is piecewise constant and jumps when the logarithm crosses an integer multiple of $\log 2/(2+8\varepsilon)$. A window bound that quantifies over the fields attached to $a(x)$ must cover these jumps. A proof restricted to a single fixed field between two successive jumps would not give the full window statement.

31.5 The cohomological budget and a larger normalization

The cohomological budget is $B = e^{y/2} = \sqrt{b}$. Suppose the comparison normalization is $B_0 = fmby$ with $f, m \ge 1$. Then

$$\frac{B}{B_0} \le \frac{e^{-y/2}}{y},$$

which tends to zero faster than every negative power of x. This is stronger than the coarse inequality $B \le B_0$ used in the normalized division estimate. The coarse version is useful when only a polynomial upper bound is needed; the sharper version can be retained if an exponentially small error matters later.

If a local comparison contains an additional term $Cfmby^2$, dividing by B_0 leaves $Cy = Cx^{\varepsilon}$. If it contains $C(\log x)^2 N_c$, dividing by $B_0 + N_c$ bounds this term by $C(\log x)^2$, which is at most Cx^{ε} for large x. These are two different cancellations: one uses the explicit factorization of B_0, while the other uses $N_c/(B_0 + N_c) \le 1$.

31.6 Comparing a conductor bound with the window budget

An estimate of the form

$$V \le C\exp(C(\log Q)^2)$$

for a conductor-like parameter Q does not automatically imply $V \le e^{x^{\varepsilon}/2}$. Taking logarithms shows that one needs

$$\log C + C(\log Q)^2 \le \tfrac{1}{2}x^{\varepsilon}.$$

For fixed C, a sufficient condition is $\log Q = o(x^{\varepsilon/2})$. By contrast, if $\log Q$ is comparable with x^{ε}, the displayed upper bound is much larger than the required budget.

This is a comparison of bounds, not a lower bound for V. A weak upper bound being too large does not prove the desired inequality false. It proves only that the desired inequality does not follow from that estimate alone. A sharper arithmetic theorem or a different relation between the conductor scale and the window scale is then needed.

31.7 Varying finite Euler factors

For a fixed modulus, the corrected Euler-factor discrepancy is $O(\log x)$ by (146). If the missing prime set $P = P(x)$ varies, the actual bound is

$$|P(x)| \log x + 2 \sum_{p \in P(x)} \log p.$$

Suppose $|P(x)| \le Cx^{a}$ and $\sum_{p \in P(x)} \log p \le Cx^{b}$, with fixed $a, b \ge 0$. For every $\eta > 0$, this discrepancy is at most $x^{\max(a,b)+\eta}$ for sufficiently large x. If the set contains only a bounded number of auxiliary primes but those primes are exponentially large, their logarithms still enter and must be estimated.

In particular, the mere finiteness of $P(x)$ at every scale gives no uniform error bound. The dependence on its size and on the logarithms of its members is part of the conductor bookkeeping.

31.8 Common thresholds for finitely many estimates

Suppose $P_1(x), \dots, P_m(x)$ each hold for all $x \ge X_j$. Then they all hold for $x \ge \max_j X_j$. This elementary fact allows a finite list of eventual estimates to be combined once all parameters on which the thresholds depend have already been fixed.

It does not combine separately existing objects. Statements $\exists D_1\, P_1(D_1)$ and $\exists D_2\, P_2(D_2)$ do not imply $\exists D\, (P_1(D) \wedge P_2(D))$. A simple example is an integer required to be both even and odd. In a comparison theorem the objects are spaces, maps, lattices, measures, and parameter choices, rather than integers, but the logical issue is the same.

The appropriate hypothesis is uniform validity over every sufficiently large admissible object, followed by an existence theorem producing one such object. The next proposition makes that order precise.

31.9 A simultaneous-selection principle

Let $\mathscr{D}_{\varepsilon}$ be a class of admissible data for each $\varepsilon > 0$. Let $\mathscr{K}$ be a partially ordered set in which every finite set has an upper bound. Write $D \succeq K$ for a largeness condition, assumed monotone in the sense that $K \le K'$ and $D \succeq K'$ imply $D \succeq K$.

Suppose, for each $j = 1, \dots, m$, there exists $\varepsilon_j > 0$ such that for every $0 < \varepsilon \le \varepsilon_j$ there is a threshold $K_j(\varepsilon)$ with

$$D \in \mathscr{D}_{\varepsilon}, quad D \succeq K_j(\varepsilon) \quad \Longrightarrow \quad P_j(D).$$

Suppose also there is $\varepsilon_* > 0$ such that, for every $0 < \varepsilon \le \varepsilon_*$ and every $K \in \mathscr{K}$, there exists $D \in \mathscr{D}_\varepsilon$ with $D \succeq K$ and an additional property $Q(D)$. Then, for every

$$0 < \varepsilon \le \min(\varepsilon_*, \varepsilon_1, \dots, \varepsilon_m),$$

there is a single admissible D satisfying $Q(D)$ and all the $P_j(D)$.

The proof is direct. Choose all $K_j(\varepsilon)$, take a common upper bound K, and invoke existence at K. Monotonicity lets each uniform implication apply to the resulting same object. The proposition remains valid when K is a finite tuple of constants ordered coordinatewise; the common upper bound is then the coordinatewise maximum.

31.10 Preserving selected windows

Suppose a set S contains windows $[Y, Y^{1+\delta}]$ with arbitrarily large Y, and a statement holds for all $x \in S$ above a threshold X_0. Then the set on which the statement holds still contains arbitrarily late windows: choose the left endpoint beyond both the requested threshold and X_0.

If several statements hold eventually on the same S, take the maximum of their thresholds. If they hold on different window sets, there need not be a common window. One quantitative sufficient condition can be written on the logarithmic scale. Suppose the available intervals are $[U_i, (1+\delta)U_i]$ in $u = \log x$. Their intersection contains $[U, (1+\delta')U]$ with $U = \max_i U_i$ whenever

$$(1+\delta') \max_i U_i \le (1+\delta) \min_i U_i.$$

Thus a controlled alignment of the starting points can compensate for using different initial intervals, at the cost of reducing the width. Without such alignment, separate unbounded collections of windows can be disjoint.

31.11 A fixed family and a small exponent

Suppose a fixed positive character family satisfies, for every sufficiently small $\varepsilon > 0$, an upper bound

$$G(x) \le x^{1/2 + A\varepsilon}$$

on arbitrarily late windows, where A is fixed independently of ε. To exclude a hypothetical zero with real part $\beta > 1/2$, choose ε so small that $A\varepsilon < \beta - 1/2$ and so that all construction hypotheses hold. The window pole criterion gives a contradiction.

If $A = A(\varepsilon)$, it is enough that $A(\varepsilon)\varepsilon \to 0$ along a sequence for which the same family is admissible. It is not enough to know merely that $A(\varepsilon)$ is finite for each ε. For example, $A(\varepsilon) = 1/\varepsilon$ never yields an exponent approaching $1/2$. The limiting quantity is the product controlling the actual error exponent.

31.12 The complete direction of dependence

The character and its positive retained family are fixed first. Next one fixes the exponent losses and selects a sufficiently small ε. Uniform comparison theorems provide a finite collection of largeness thresholds. A common existence theorem then supplies one compatible family of geometric and arithmetic data beyond those thresholds. Its cohomological hypothesis selects arbitrarily late windows. On each such window, the local indices, heights, kernel estimates, and prime-class identity yield the scalar bound for G. Only then is the analytic pole criterion applied.

Reversing these choices can change the statement. A family selected after seeing a hypothetical zero may fail to contain the original fixed character. A window selected separately for each comparison map may fail to be common. A constant allowed to grow with the window endpoint may no longer be absorbable into a fixed exponent loss. The simultaneous-selection principle records exactly the uniformity needed to avoid these failures.

32 Cohomological windows and conditional arithmetic comparison

32.1 The cohomological window hypothesis

We now give a quantitative condition that can serve as an arithmetic input. The notation $\mathcal{V}(E,t,B)$ is as in (76), and associated cyclotomic fields are those defined by (84).

Hypothesis 32.1 (Cohomological windows)**.** For every primitive character ψ, every $0<\varepsilon<1/52$, and every function $a(x)$ satisfying (85), there exist $J\ge 1$ and X_0 such that the following holds. For every $X\ge X_0$ and every prime pair (q,d) admissible at $(\varepsilon, X^{1/J})$, there is a real number Y satisfying

$$X^{1/J}\le Y, \qquad Y^2\le X,$$

such that, for all $Y\le x\le Y^2$ and every field E associated to $(\psi,a(x),q,d)$,

$$\mathcal{V}\left(E,\lfloor x^{401}\rfloor,\exp\left(\tfrac{1}{2}x^{\varepsilon}\right)\right) \tag{148}$$

holds.

This hypothesis does not assert existence of an admissible prime pair. Without such existence its universal statement about pairs could be vacuous at some scales. Prime supply must therefore be added whenever one uses the hypothesis to obtain actual windows.

The constant J is chosen before X and before the pair. It may depend on ψ,ε, and the function a. The same q,d are then used for every x in the selected interval. The twist range varies with x, and the field may vary with $a(x)$. These dependencies are part of the strength of the hypothesis.

32.2 Recovering arbitrarily late intervals

Assume admissible pairs exist for all sufficiently large values of X in Hypothesis 32.1. Fix a requested lower endpoint $Y_0>1$. Choose $X\ge \max\{X_0,Y_0^J\}$ large enough for prime supply and select an admissible pair. The resulting Y then satisfies $Y\ge Y_0$. Thus the cohomological bounds hold on arbitrarily late intervals $[Y,Y^2]$.

For any fixed $1<\lambda\le 2$, the smaller interval $[Y,Y^{\lambda}]$ lies inside $[Y,Y^2]$. A later comparison may therefore use a window exponent such as $3/2$. This restriction costs no new arithmetic estimate. It does require that all other comparison objects be defined on the smaller interval with the same selected data.

There is also a simple consistency observation about J. If an admissible pair exists for some $X>1$ and the conclusion gives Y, then $X^{2/J}\le Y^2\le X$, so $J\ge 2$. The weaker displayed condition $J\ge 1$ is compatible with this: nonvacuous instances impose the stronger inequality automatically. It would be incorrect to choose an arbitrary J in the interval $[1,2)$ and assume that such a Y exists.

32.3 The size of the cohomological budget

Write

$$b = \exp(x^{\varepsilon}), \qquad y = x^{\varepsilon}.$$

The budget in (148) is $\sqrt{b}$. For $x \geq 1$, one has $y \geq 1$, $b \geq 1$, and $\sqrt{b} \leq b$. The twist cutoff $\lfloor x^{401} \rfloor$ grows polynomially in x, whereas the budget grows like the exponential of a small power of x. For every fixed $A > 0$,

$$x^{A} = o(\exp(\tfrac{1}{2}x^{\varepsilon})) \quad \text{as } x \to \infty$$

when ε is fixed and positive. This follows by taking logarithms and using $\log x = o(x^{\varepsilon})$.

The budget is therefore large compared with a fixed polynomial. That observation does not make the hypothesis automatic: it simultaneously controls many twists, all primes in a uniform finite-set inequality, and varying number fields. A pointwise finiteness theorem for each individual cohomology group would not give the stated uniform bound.

32.4 A normalized division estimate

The following elementary calculation isolates how a cohomological budget can enter a later comparison. Let $f, m \geq 1$ and put $B_0 = fmby$. Suppose $d_j \geq 0$, $N_c \geq 0$, and a nonnegative quantity η_j satisfies

$$\eta_j \leq d_j \left[V + C(fmby^2 + (\log x)^2 N_c) \right], \qquad 0 \leq V \leq \sqrt{b}, \tag{149}$$

where C is a fixed real constant. For fixed $\varepsilon > 0$ and sufficiently large x, one has $(\log x)^2 \leq y$. Since $V \leq b \leq B_0$ and $fmby^2 = yB_0$, it follows that

$$\begin{aligned} \eta_j &\leq d_j \left[B_0 + |C| y (B_0 + N_c) \right] \\ &\leq d_j (|C| + 1) y (B_0 + N_c). \end{aligned}$$

Consequently

$$\frac{\eta_j}{B_0 + N_c} \leq d_j (|C| + 1) x^{\varepsilon}. \tag{150}$$

The denominator is positive because $B_0 > 0$.

This calculation is conditional on (149). The cohomological bound supplies a possible value of V; it does not by itself identify the division quantity η_j, the rank d_j, or the common normalization $B_0 + N_c$. Those identifications belong to the geometric comparison. Once they are established, the displayed normalization is a direct inequality with no hidden limiting step.

32.5 Residual character sums

Let $o \geq 0$ be an integer, and suppose a residual family contains at most $4o$ character expressions, each bounded in absolute value by $x^{1/2+\varepsilon}$. If $\Sigma(x)$ is their real sum, then

$$|\Sigma(x)| \leq 4o\, x^{1/2+\varepsilon}. \tag{151}$$

This follows from the triangle inequality; no cancellation among the residual terms is needed.

Suppose a common arithmetic-geometric comparison supplies a real quantity $\mathcal{B}(x)$ with the exact identity

$$\mathcal{B}(x) = (3o + 1)\big(4G_{\mathcal{F}}(x) - \Sigma(x)\big). \tag{152}$$

The coefficient $3o+1$ is positive. Hence any upper bound for $\mathcal{B}$ transfers to an upper bound for $G_{\mathcal{F}}$, after the residual estimate is inserted. The identity is essential: bounds on some unrelated geometric quantity would not control the character sum.

Proposition 32.2 (Absorbing fixed comparison constants)**.** *Fix $A \ge 0$, $C \ge 0$, and $o \ge 0$. Suppose* (151) *and* (152) *hold, and*

$$\mathcal{B}(x) \le C x^{1/2+(A+6)\varepsilon}$$

on a set of scales $x \ge 1$. Then, for all sufficiently large x in that set,

$$G_{\mathcal{F}}(x) \le x^{1/2+(A+7)\varepsilon}.$$

The new threshold may depend on C, o, ε.

Proof. The balance identity and residual estimate give

$$G_{\mathcal{F}}(x) \le \frac{C}{4(3o+1)} x^{1/2+(A+6)\varepsilon} + o\, x^{1/2+\varepsilon}.$$

Since $A+6 \ge 1$ and $x \ge 1$, the second power is at most the first. Thus $G_{\mathcal{F}}(x) \le C' x^{1/2+(A+6)\varepsilon}$ with $C' = C/[4(3o+1)] + o$. If $C' = 0$ there is nothing to prove. Otherwise choose x large enough that $C' \le x^{\varepsilon}$. □

The six losses in the preliminary exponent must come from actual comparison estimates; Proposition 32.2 does not derive them from a cohomological bound. It proves the final seventh loss used to absorb a fixed coefficient. This is the place where a constant depending on a fixed character family can be harmless even when no uniform bound in the conductor is known.

32.6 A complete conditional assembly statement

Theorem 32.3 (Arithmetic comparison implies nonvanishing)**.** *For each primitive character ψ, suppose there are a finite admissible positively weighted family $\mathcal{F}_\psi$ containing ψ, a fixed integer $o \ge 0$, a fixed real $A \ge 0$, and constants $\varepsilon_0, \delta > 0$ with the following property. For each $0 < \varepsilon < \varepsilon_0$ there is a set S_ε containing arbitrarily large windows $[Y, Y^{1+\delta}]$, and common comparison data on S_ε satisfying* (152), (151), *and*

$$\mathcal{B}(x) \le C_\varepsilon x^{1/2+(A+6)\varepsilon}$$

for all sufficiently large $x \in S_\varepsilon$, with a finite $C_\varepsilon \ge 0$. Then every Dirichlet L-function is nonzero in $\operatorname{Re} s > 1/2$, apart from the principal pole, and every zero in the open critical strip lies on its central line.

Proof. Proposition 32.2 gives a bound with exponent $1/2 + (A+7)\varepsilon$ after increasing the lower threshold. The set of admissible scales still contains arbitrarily late windows: choose each left endpoint beyond that threshold. All family data and the coefficient $A+7$ are fixed before ε is made small. Theorem 11.1 applies to each primitive target character. □

Hypothesis 32.1, together with prime supply, can provide a family of intervals on which the required cohomological input is available. A geometric construction must still produce the common comparison data and prove their scalar estimates on those intervals. Theorem 32.3 states exactly what is obtained once that construction is supplied. The theorem is not an unconditional assertion that such data exist.

32.7 The order of the mathematical tasks

The construction problem naturally separates into several obligations. One must first fix the target character family and the exponent losses. For a chosen small ε, one must then obtain auxiliary primes and associated fields on sufficiently late intervals. Next come the cohomological bounds, the geometric spaces and lattices, and the maps that identify the quantities used by the different estimates. Finally one proves the balance identity and the upper bound for its geometric side.

The last analytic step selects ε from a hypothetical zero and uses one sufficiently late interval. It does not require a limit of fields, lattices, or representations as $\varepsilon \downarrow 0$. Conversely, the simplicity of that last step does not reduce the content of the construction obligations. A proof of the unconditional critical-line statement would have to discharge them, with the stated quantifier order and common-data requirements.