Notes on Alpöge's Proof of The Riemann Hypothesis

Printed by Taylor Dupuy
AI GENERATED

21 November 2026

Contents

Preface vii

A remark by Levent Alpöge ix

Introduction 1
The main theorem 1
The idea of the argument 2
The structure of the proof 3
What is proved in these notes 5
Prerequisites and conventions 6
Organization of the notes 6
AI and Software Discussion 7

Chapter 1. The basic idea 9
1.1. Stepanov's method over a finite field 9
1.2. The dictionary 13
1.3. The quantity to bound 14
1.4. Divisibility is at most size 20
1.5. Multiplicity is the bottleneck 25
1.6. Theta lifts instead of differentiating integers 26
1.7. What the outline omits, and where to read on 29

Chapter 2. The function-field model: Stepanov–Bombieri and the sign-only proof 31
2.1. Setting 32
2.2. The zeta function 35
2.3. The upper bound 38
2.4. Route 1: the lower bound by Galois averaging 50
2.5. The Riemann hypothesis 56
2.6. Route 2: a one-sided bound and a Landau–Pringsheim lemma 57
2.7. Stepanov's derivative form over a prime field 66
2.8. The sign-only identity $qN = M^k + R$ 71
2.9. Iterated Parseval: the bound on R 74
2.10. The sign-only theorem 77
2.11. Multiplicity versus sign 82
2.12. What the identity is for 85
2.13. Dictionary with Alpöge's argument 88

2.14. Remarks and questions 90

Chapter 3. The argument assembled 95
3.1. The main theorem 95
3.2. The skeleton of the proof 96
3.3. The vocabulary of the black boxes 97
3.4. Black box: Galois representations 101
3.5. Black box: the Sha bound on windows 102
3.6. Black box: the supply of hosts 103
3.7. Black box: the covolume bound 104
3.8. Black box: packages, Green's parametrisation, tied families 105
3.9. Black box: the Chebyshev inequality on windows 107
3.10. Black box: the windowed Landau criterion 112
3.11. Black box: from a half-plane to the critical line 113
3.12. Proof of the main theorem 113
3.13. Chebyshev's comparison and the function-field model 116
3.14. The status of the inputs 119

Chapter 4. Smoothed Chebyshev functions and a windowed Landau criterion 125
4.1. Conventions for Chapters 3–12 125
4.2. Smoothed Chebyshev functions 126
4.3. The windowed Landau criterion 127
4.4. Levels 132
4.5. From a half-plane to the critical line 133

Chapter 5. Host fields and averaging over auxiliary primes 137
5.1. Frames and hosts 137
5.2. Kept and bounded characters 139
5.3. Scales and the explicit-formula condition 141
5.4. Averaging over the auxiliary prime 143
5.5. The strength of the analytic input 147

Chapter 6. The balance identity and Fourier analysis on the Galois group 151
6.1. The additive model and arithmetic data 151
6.2. Thick and thin weights 157
6.3. Averaging over shifts and top places 159

Chapter 7. Hermitian lattices, the theta tower and slope records 165
7.1. The tower of unitary groups and the theta chain 165
7.2. Hermitian lattices and slopes 168
7.3. The slope record 170
7.4. The frame theorem 175
7.5. The archimedean side 178
7.6. The finite side 179

Chapter 8. Families and the covolume bound 181
8.1. Families and closed properties 181
8.2. The covolume bound 182
8.3. How arithmetic geometry enters 185

Chapter 9. The Chebyshev inequality on windows 189
9.1. Window hypotheses and ties 189
9.2. The gap bound 192
9.3. The square-root bookkeeping 195
9.4. The Chebyshev inequality on windows 196

Chapter 10. Tate–Shafarevich groups of Tate twists 199
10.1. The groups and the bound 199
10.2. Eigenspaces and the case split 202
10.3. The minus part 205
10.4. The plus part 207
10.5. The Sha bound on windows 213

Chapter 11. Galois representations and local–global compatibility 217
11.1. Weak local Langlands maps and truncated compatibility 217
11.2. A proposition on compact groups 222
11.3. What the Sha bound uses 224

Chapter 12. Existence of families 225
12.1. The supply of hosts 225
12.2. Green's parametrization 228
12.3. Admissible families 229
12.4. Zero-freeness for primitive characters 232
12.5. What this chapter supplies to the main proof 233

Appendix A. Software and AI methods 237
A.1. Typesetting 238
A.2. Agent management 239
A.3. Natural language proofs 240
A.4. Adversarial reviews 241
A.5. AI assisted computations 242
A.6. AI statement generation 243
A.7. Proofreading and adversarial review 243
A.8. Literature review 245
A.9. Formalization 246
A.10. Record keeping and token accounting 247

Bibliography 251

Preface

These are lecture notes on the argument for the Riemann Hypothesis announced by L. Alpöge in the note reproduced after this preface [**Alp26b**], together with a Lean archive [**Alp26a**]. They are written for professional mathematicians and graduate students. The aim is that the reader first sees how the argument fits together and then meets the tools one chapter at a time.

Chapter 1 explains the idea of the argument. Chapter 2 works out the model case over a finite field: the Stepanov–Bombieri proof of the Riemann hypothesis for curves, and the conditional sign-only argument through the identity $qN = M^k + R$. Chapter 3 states the main theorem and gives its whole proof, with each tool used as a black box and a reference to the chapter where it is treated. Chapters 4 to 12 develop the tools. Appendix A describes the software and AI methods used to produce the text.

AI notice. AI agents wrote this text, and the AI and Software Discussion at the end of the Introduction (page 7) and Appendix A describe how.

A remark by Levent Alpöge

The Riemann Hypothesis is the classic problem in number theory. It asks for optimal control on the count of prime numbers of a given size. Its shadows have guided mathematicians for 150 years, and its many rhymes and reflections, even reaching physics, taught us the unity of our fields.

I believe I have, working with Claude, discovered a proof[1] of the Riemann Hypothesis. In fact, the method easily extends much further, resolving the so-called Generalized Riemann Hypothesis, and I expect, with more time, the Extended Riemann Hypothesis too. Claude and I have also formalized the arguments without assumptions. It is natural to feel that this is an end, in my case a loss of a principal guiding light, but what I saw during the course of progress so amazed me that I believe it is one dawn across science of something far greater.

It was my secret in my student days that my plans were for this problem in decades, which meant then that I tried to become familiar with all tools that I could possibly recognize in the problem when it was time. Claude and I have had a great time together with such breadth in recent months, but this time Claude far outpaced me near the finish line. No, Claude flew, and danced, and sparkled near that finish line. I had us on the line of combining two quite different ideas present in a toy setting, one of which was already new and shocking, making very simple one of the shadows of Riemann over which we mathematicians felt most triumph. I guided us to the right situation analogous to that toy setting, and then Claude began to leave me in awe.

Tens of such occurred, but I will give one line which will suffice to make fellow mathematicians fall out of their chairs. In Stepanov's argument resolving Riemann over function fields one uses a divisor involving multiplicity. This argument is a descendant of Stepanov's, but multiplicity is hard to

[1] `https://alpo.ge/lean.tar.gz.part01`,...,`https://alpo.ge/lean.tar.gz.part14`.

find over the integers, corresponding to something like differentiating whole numbers. The simple, new, Fourier-analytic ideas in the toy setting led us to examine chaining theta lifts if we were to have the same tricks recur but work in a natively automorphic manner. The line: for the von Mangoldt sum of length X, Claude's suggested length of chain of theta lifts, in view of each mimicking multiplicity, was $\sqrt{X}$. (And yet, what else could it be?)

An astounding amount of mathematics is present in the argument, far more than I had imagined would be necessary as a student. My intuition insists it will be massively simplified, likely immediately. But I hope in having the experience of seeing all of number theory's beauty brought together I am just one of the first, and those who love the mathematics will all be able to enjoy such transcendence in the future.

The argument is incredibly long at the moment, and while its formalization is finished and it is routine to produce a huge, unreadable document containing sufficient detail, in my view of my path I have just taken a step. I intend to do my best to understand and impart understanding of this argument and its certain simplifications, to speak about it, to write it properly.

In a time of revolution in mathematics I hope the argument provides a glimpse towards the miracles that lie ahead, and shows us the joy and wonder we are about to find in the simplest things.

1

It is obvious to me that I should remark on the following.

The precise timing is horrible. If I could have, I'd have finished this all much earlier. Hell, strictly before this age would have sure been nice. This all happened for me late Sunday night, and I messaged Dario, "Hypothetically, if I were to have . . . ", eventually asking about my ability to announce. We went and figured it out and the fact is that there was a constraint to respect, though I did ask about progressively more bizarre ideas. This problem has been *the* sacred object in mathematics for me since my childhood, and I am very touchy about the treatment of great intellectual goals, though seriously imperfect. Our leadership and the few trusted mathematicians I consulted hold extremely similar views on the right things to do in these circumstances (it is occasionally possible for physicists and mathematicians to agree. . .), and I am clearly unilaterally responsible for feeling

I just could not live if I waited even longer. There is an obvious cynical view, to do with the timing, but it is false, and I am pleased that I am in circumstances where it is obviously false.

This puts us in the rather extraordinary situation this age has made possible, in which I am presenting the world with a complete argument before having mastered a massive neighbourhood around its mathematics. It being obviously inappropriate to attempt to post some massive unreadable travesty of "slop"[2], I decided to only present Lean.

In a different vein, I have a lot of people I want to praise, and I will do a proper job in a proper writeup, alas pardon the accidental omissions — among many others, Manjul Bhargava, Kai Gill, Hamish Kerr, Max Leiter, John Ma, Tal Magdish, Chris Paika, Maksym Radziwiłł, Jascha Sohl-Dickstein, Peter Sarnak, Jarred Sumner, Amir Tajdini, and especially David Dohan, Ben Michel, Ralph Furman, Eric Price, and Cora Vinther, and my miraculous colleagues. I want them to be proud.

I have benefited from the good taste of the people at OpenAI involved with various mathematical results, so I am quite happy to praise them too. The Ehrhart conjecture was one such tasteful choice some time ago; playing further and working out the equality case I saw accurate handling of solutions of the Monge-Ampère equation and immediately understood Navier-Stokes was going to fall. Similarly, their quiet inclusion of the Siegel argument, whose result was strictly dominated by the $(2, 3, 6)$-triangle group trick, again showed great taste; though I did not comment so publicly, I immediately recognized it as Stepanov without multiplicity, though it was only clear to me during this rather harebrained tornado through the whole of number theory that this problem would fall given enough all-nighters and physical toll when I had the $qN = M^k + R$ sign-only proof of Weil and Deligne.

2

I had a stream-of-consciousness jaunt through the experience of finding all the pieces of the argument through incredible confusion with Claude, but a trusted mathematician advised me to just get to the point and have some sort of timeline. I emphasize that once we saw multiplicity 2 on $O(2, n)$ via the

[2]I of course amused myself, and hit $250,000+$ pages.

Claude cutely called it the GRH ... "book" — I looked at Theorem 1 and it began by noting that because the reals are totally ordered if $\sigma \neq \frac{1}{2}$ then $\sigma > \frac{1}{2}$ or $\sigma < \frac{1}{2}$.

"natively-automorphic" determinant that I'd been searching for via analogy to the definite sign of the left-hand side of[3] $qN = M^k + R$ meaning holo/automorphy over a function field, winning more multiplicity with more theta lifts, we were going to win but I was also, so to speak, out of my element, in that I personally have little mastery of automorphic forms on higher rank groups (youthful foolish stubbornness). I also emphasize that I may have baked lots of strange suboptimal choices into the argument, and expect short arguments to follow fairly quickly, though I have not had time to search for such myself (yet!).

So here is a rough timeline Claude provided after a thorough search.[4]

1. **10/6/26:** Stepanov determinant.
2. **10/9/26:** $D^p \equiv 0$ on a CM curve for p inert in K, D the invariant derivation, and $\frac{D^a \cdot \theta(0)}{\Omega^a}$ is divisible by $p^{\frac{a}{p}+\frac{a}{p^3}+\cdots-O(\log a)}$.
3. **10/19/26+:** $qN = M^k+R$, using $x_1\cdot x_2+\cdots+x_{k-1}\cdot x_k$, wanting "automorphy" of N since μ is automorphic from θ on sixfold cover of GL_5 as a period integral of thetas, quadric has natural interpretation on corresponding CM e.g. $\mathbb{C}/\mathbb{Z}[\omega]$ as Weil[5], put Poincaré bundle on the edges to get θs.
4. **10/6/26+:** $W(\mathbb{1}_S) \geq 0$ via a determinant, interpret as Stepanov with multiplicity one again; bootstrap!!

---

[3] $$\sum_{h,x_1,\ldots,x_k} t(x_1)\cdots t(x_k)\cdot e_q\left(h\left(x_1\cdot x_2+\cdots+x_{k-1}\cdot x_k\right)\right),$$
sum over h first for the left-hand side, with e.g. quadric coefficients too Bertini gives a definite sign from middle cohomology, one off the one on M^k, which is the zero-mode on the right-hand side with R square-root-quality up to an inch or so as Katz would say by iterating Parseval k times, opposite signs meaning no cancellation of inverse roots, QED; vectorize to go up dimensions. Gabber's hypersurface in one of Katz's arguments reminds one of the quadric, which comes from looking for a simplification of averaging over the family $\chi \mapsto \chi(\mathrm{disc})$, corresponding to $a \mapsto y(\cdot)^a$ with y the normalized Gauss sum of order $\#|\mu_K|$ who lives via θ on the $\#|\mu_K|$-fold cover of GL_2, the reciprocity law of $y(c\cdot c')$ vs. $y(c)\cdot y(c')$ showing some sort of complete graph, simplifying to the above "path" quadric, Gauss sum cancellation replaced by iterated Parseval.

[4] The parts in which my vague intuition worked out due to Claude's technical mastery are of course less familiar to me today, and probably I'll have to edit some of these descriptions on reflection. The parts in which Claude is the expert and not me I am sort of paraphrasing a "slop" summary, which might make zero sense, and I apologize, though the Lean should suffice to clarify everything. I'm just too inexpert in the relevant stuff at the moment, I hope the reader sees the breadth of this argument!

[5] A lot is obviously being lossily compressed here, but especially after gritting with Sussman through Bałaban's fantastic induction I had to think to myself about how hard fighting for some sort of $\mathbb{F}_1$ à la e.g. Borger would be.

5. **10/20/26:** definitively switch to Λ aka $\sum_{n\leq X}\Lambda(n)\cdot\left(\frac{X}{n}-1\right)_+ - X\log X+(1+\gamma)\cdot X$; Guinand-Weil had the sign, variance test function trick to make things look trace-formula-like (but how can I pigeonhole / Turán over n? X? *it*?), Mehta-like integral over twists or at infinity...
6. **10/8/26:** $W\left(\mathbb{1}_{(0,1]}+(\sim X^{\eta})\cdot\mathbb{1}_{(0,1/X]}\right)\geq 0$ gives $>1-\eta$.
7. **10/14/26:** Determinants for $W(\cdot)$'s containing the desired periods, indeed GGP periods, with values $0,1$ meaning interlacing weights, or up to m non-interlacing ones.
8. **11/15/26:** Iterated theta lifts $U(X_0)\to U(X_1)\to\cdots\to U(X_m)$ between totally definite unitary groups, $\pi\mapsto\pi\boxplus\mu_j[t]$, wins multiplicity m. Time to take m like $\sqrt{X}$...
9. **11/15/26:** Norms of the lifts and level and cancellation of order-m terms with proven divisibility via Chowla-Selberg, Weyl, and Siegel mass.
10. **11/15/26:** Sufficient lower bound on divisibility of the m-fold iterated lift.
11. **10/6/26:** Iwasawa tower via analogy to pigeonhole on $\mathbb{F}_{q^n}/\mathbb{F}_q$.
12. **11/15/26:** Seesaw: the period of the iterated lift is a Bessel period independent of the Hecke characters, Rallis.
13. **10/6/26:** Bost slope inequality.
14. **11/15/26:** precise computation of the "cost" of each lift, showing that one would stop at $2/3$ a priori, attempt to get a main term and transfer constant to a function field or kill via twisting, need for "division".
15. **11/15/26:** Glue together an integral structure across Hecke characters for a "gcd" i.e. extra averaging.
16. **10/14/26:** Relative trace formula, seen in a different half-plane argument via sixfold cover of GL_5 automorphy of μ and reducing via Cauchy-Schwartz to a bound on the norm of a suitably projected Poincaré series, approximate orthogonality of the columns of the integral structure.

17. **11/15/26:** for the classes in $H^1\left(\mathfrak{o}_E\left[\frac{1}{\ell}\right], \mathbb{Q}_\ell/\mathbb{Z}_\ell(-e)\right)$ that we couldn't kill by twisting, fight for a significant weakening of Vandiver.[6]
18. **11/15/26:** computation of covolumes, theta series' index, congruences between lifts and forms.
19. **10/19/26:** precise choice of the Iwasawa tower, specifically using Behrend!![7]
20. **11/16/26:** GRH.

— Levent Alpöge, untitled note.

Source: `https://alpo.ge/remark.pdf`, retrieved 2026-11-21 [**Alp26b**]; PDF creation date 2026-11-20. Reproduced verbatim; the bold dates, the part headings and the list layout are the transcriber's formatting.

[6]This is still pretty strong!! I believe vanishing of a number of K-groups of $\mathbb{Z}$ follows, and cyclicity up to 100ish maybe? I had Lichtenbaum, so I knew we would have a lot of slack in fighting these K's.

[7]This is the sort of trick that intuitively to me is obviously going to be unnecessary.

Introduction

The Riemann hypothesis states that every zero of the Riemann zeta function $\zeta(s)$ in the strip $0 < \operatorname{Re} s < 1$ has real part $1/2$ [**Rie59**]. The same statement for the L-function $L(s, \chi)$ of every Dirichlet character χ is the generalized Riemann hypothesis for Dirichlet L-functions. The first statement is a sharp bound on the oscillation of the primes around their expected density. The second is the corresponding bound for the primes in residue classes.

In November 2026 L. Alpöge published a note [**Alp26b**] and an archive of Lean source files [**Alp26a**]. The note is reproduced before this introduction. It states that its author believes that he has discovered, working with Claude, a proof of the Riemann hypothesis. It also states that the method gives the generalized Riemann hypothesis and that the arguments are formalized in Lean without assumptions. The note contains a timeline of twenty steps. It gives no proof of the theorem. The argument is published only as the Lean archive.

These notes explain that argument for professional mathematicians and graduate students. They state the main theorem, show how the proof is put together from its parts, and then treat the parts one chapter at a time. They do not verify the argument or the Lean archive (Remark 1.7.1).

The main theorem

Dirichlet characters and their L-functions are defined in Section 3.1. The character modulo 1 is the primitive principal character, and its L-function is ζ. The statement that Alpöge claims is Theorem 3.1.1 of Chapter 3.

Main Theorem (Alpöge; Theorem 3.1.1). *Let χ be a Dirichlet character. Then $L(s, \chi) \neq 0$ for every s with $\operatorname{Re} s > 1/2$, except $s = 1$ when χ is principal, and every zero of $L(s, \chi)$ with $0 < \operatorname{Re} s < 1$ has real part $1/2$. Every zero of $\zeta(s)$ other than $-2, -4, -6, \ldots$ has real part $1/2$.*

The theorem concerns ζ and Dirichlet L-functions only (Remark 3.1.2). Chapter 3 proves it relative to twenty statements that are used as black boxes (Table 3.3). Fifteen of them are statements of Alpöge's argument, two are cited from the literature and three are unattributed. These notes give complete proofs of two of the fifteen and of the three unattributed black boxes, each relative to results that it quotes. For the other thirteen they give an outline of the proof, a description of its inputs or an explanatory account

of a standard method. For some of these black boxes, parts have none of the three (Table 3.3). So these notes do not establish the Main Theorem.

The idea of the argument

The model is the proof of the Riemann hypothesis for curves over finite fields by the method of Stepanov [**Ste69**], in the form that Bombieri gave it [**Bom74**]. Let C be a smooth, projective, geometrically irreducible curve of genus g over $\mathbb{F}_q$, and let $N = \#C(\mathbb{F}_q)$. One constructs a nonzero rational function on C that vanishes to order at least m at every rational point except one, and that has a single pole of controlled order. A nonzero function has as many zeros as poles, so $m(N-1)$ is at most the order of the pole. For q a square with $q > (g+1)^4$ and $m = \sqrt{q}$ this gives

$$N \le q + (2g+1)\sqrt{q}$$

(Proposition 1.1.1). The multiplicity m costs only one vanishing condition at each point, because m is a power of the characteristic and the map $f \mapsto f^m$ is additive. In general the error is at most the sum of two terms, $(g+1)q/m$ from the vanishing conditions and $2gm$ from the linear algebra that produces the function (Proposition 2.3.9). The first term falls as m grows and the second rises. They have the same order at $m \approx \sqrt{q}$, and no choice of the parameters gives an error $o(\sqrt{q})$ (Proposition 2.3.10).

Over the integers the counterpart of the comparison of zeros with poles is the product formula. For a nonzero integer Δ,

$$\sum_p \operatorname{ord}_p(\Delta) \log p = \log|\Delta|.$$

If Δ is known to be divisible by p^{v_p} for many primes p, then $\sum_p v_p \log p \le \log|\Delta|$. Chebyshev's bound for the primes between n and $2n$ is this inequality for the binomial coefficient $\Delta = \binom{2n}{n}$ (Proposition 1.2.1). It has multiplicity one, and it gives the right order of magnitude without a square-root error. The ring $\mathbb{Z}$ has no endomorphism with the properties of the map $f \mapsto f^m$ (item (B1) of Section 2.13). Alpöge's note says that multiplicity is "hard to find over the integers" [**Alp26b**, p. 1].

The quantity that the argument bounds is a smoothed Chebyshev function. With Λ the von Mangoldt function and γ Euler's constant, put

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

A one-sided bound for G_ζ is sufficient. If $0 < \theta < 1$ and $G_\zeta(x) \le x^\theta$ for all large x, then $\zeta(s) \ne 0$ for $\operatorname{Re} s > \theta$ (Proposition 1.3.4). The proof uses a theorem of Landau on Mellin transforms of nonnegative functions [**Lan05**]. The target is therefore an upper bound $G_\zeta(x) \le x^{1/2+\varepsilon}$ for every $\varepsilon > 0$, together with its analogue for Dirichlet characters.

The integers Δ are replaced by periods of theta lifts of algebraic automorphic forms on definite unitary groups. The groups are defined over

an abelian CM field that contains an imaginary quadratic field K. On the archimedean side the periods have a size that is given by a Stirling term and a power of a CM period. On the finite side they are divisible, to an order of the type of Legendre's formula, by every prime that is inert in K. A slope inequality for the period map between two Hermitian lattices takes the place of the comparison of zeros with poles. Fourier analysis on the Galois group of the CM field brings in the primes that are not inert in K and the Dirichlet characters.

Multiplicity comes from a chain of m theta lifts. According to the note, each lift collects the forced divisibility again (Section 1.6.3). After normalisation one part of the error is of size x/m and another is of size m, and the two are equal at $m \approx \sqrt{x}$ (Proposition 9.3.1, Remark 9.4.3). This is the chain length $\sqrt{X}$ for a von Mangoldt sum of length X that the note names [**Alp26b**, p. 1].

Chapter 1 gives this overview with more detail and with proofs of its elementary steps. Chapter 2 gives the function-field model in full. It contains a complete proof of the Riemann hypothesis for curves by the method of Stepanov and Bombieri (Theorem 2.5.3). It also treats the second argument that the note mentions for finite fields, the sign-only identity

$$qN = M^k + R$$

(Proposition 2.8.2). In this identity M is the character sum to be bounded, N is a weighted point count on a quadric in k variables, and R is bounded by Fourier analysis. The bound for M that Chapter 2 derives from the identity is conditional (Theorem 2.10.2).

The model and the argument differ in the direction of the inequality. An upper bound for the function-field counterpart of G_ζ is a lower bound for point counts (Proposition 2.6.10). The method of Stepanov and Bombieri gives an upper bound for point counts. The model and the argument also differ in the auxiliary object. The argument constructs no auxiliary function and no space of small sections. Its slope inequality is applied to the period map between two given lattices (Remark 7.2.6).

The structure of the proof

Chapter 3 gives the proof of Theorem 3.1.1 with each tool as a black box. Figure 3.1 shows the dependencies. Fix a primitive Dirichlet character ψ. The proof has four parts (Figure I.1).

The first part is the analytic end. To ψ is attached a combination $G_{\mathscr{L}_\psi}$, with positive weights, of smoothed Chebyshev functions of ζ and of the primitive characters that induce the powers of ψ (Definition 4.4.1). Suppose that $G_{\mathscr{L}_\psi}(x) \le x^{\sigma_0}$ on windows $[Y, Y^{3/2}]$ with Y arbitrarily large. Then $L(s,\psi) \neq 0$ for $\operatorname{Re} s > \sigma_0$, except at $s = 1$ when ψ is trivial. This is the windowed Landau criterion (Theorems 4.3.3 and 4.4.3). Because the weights are positive, an upper bound is sufficient (Remark 4.3.7). Chapter 4 treats the criterion. It also contains the step from primitive characters and

the half-plane $\operatorname{Re} s > 1/2$ to all Dirichlet characters and the critical line (Proposition 4.5.2), which uses the functional equation (Theorem 4.5.1).

The second part is the inequality. The Chebyshev inequality on windows (Theorem 9.4.1) concerns a family of automorphic data that is *tied* to a slope record at every large point of its window set. For the families that Chapter 12 produces it gives $G_{\mathscr{L}_\psi}(x) \leq x^{1/2+9\varepsilon}$ on the windows of the family.

Five chapters lead to the inequality. Chapter 5 constructs the CM fields, which are called host fields (Theorem 5.2.4). It bounds the characters of their Galois groups that are not wanted by an average over two auxiliary primes (Theorem 5.4.8). Chapter 6 proves the balance identity (Proposition 6.1.4). The identity writes a prime sum of the type of Legendre's formula as a main term plus a combination of twisted Chebyshev functions. Chapter 6 also contains the core bound (Theorem 6.3.4), which turns an upper bound for the prime sum into an upper bound for $G_{\mathscr{L}_\psi}$.

Chapter 7 defines the tower of unitary groups, the chain of theta lifts, Hermitian lattices and slope records. Its main statement is the frame theorem (Theorem 7.4.2), a slope comparison between many source lattices and one target lattice. Chapter 8 states the covolume bound for lattices of integral automorphic forms (Theorem 8.2.3). This theorem is the one place where arithmetic geometry enters the argument. Chapter 9 combines these results in the gap bound (Theorem 9.2.1) and deduces the inequality. The exponent $1/2$ comes from Proposition 9.3.1.

The third part consists of two arithmetic inputs. The lattices of the chain differ from the ideal ones by congruence defects. The defects are bounded by Tate–Shafarevich groups of Tate twists over the host fields. Chapter 10 states the bound that is needed, the Sha bound on windows (Theorem 10.5.1). The bound is obtained by a pigeonhole argument that gives it only on windows (Remark 10.5.2). For this reason the Landau criterion is needed on windows (Remark 4.3.8). The Sha bound uses Galois representations attached to automorphic representations over CM fields, with a truncated form of local–global compatibility. Chapter 11 states this input (Theorem 11.1.4).

The fourth part is existence. Chapter 12 states that hosts are supplied on windows (Theorem 12.1.2), that packages of automorphic data exist (Theorem 12.3.4) and that tied families exist (Theorem 12.3.5). It uses Green's parametrisation of the cuspidal characters of $\mathrm{GL}_n(\mathbb{F}_p)$ [**Gre55**]. From these statements, the Chebyshev inequality on windows and the Landau criterion it then deduces that $L(s,\psi) \neq 0$ for $\operatorname{Re} s > 1/2$, except at $s = 1$ when ψ is trivial (Theorem 12.4.1).

The proof of Theorem 3.1.1 in Section 3.12 is a chain of six implications in this order: Galois representations, the Sha bound on windows, the supply of hosts, the existence of tied families, zero-freeness for primitive characters, and the step to all characters and the critical line. The deduction contains no substantial argument of its own (Remark 3.12.2).

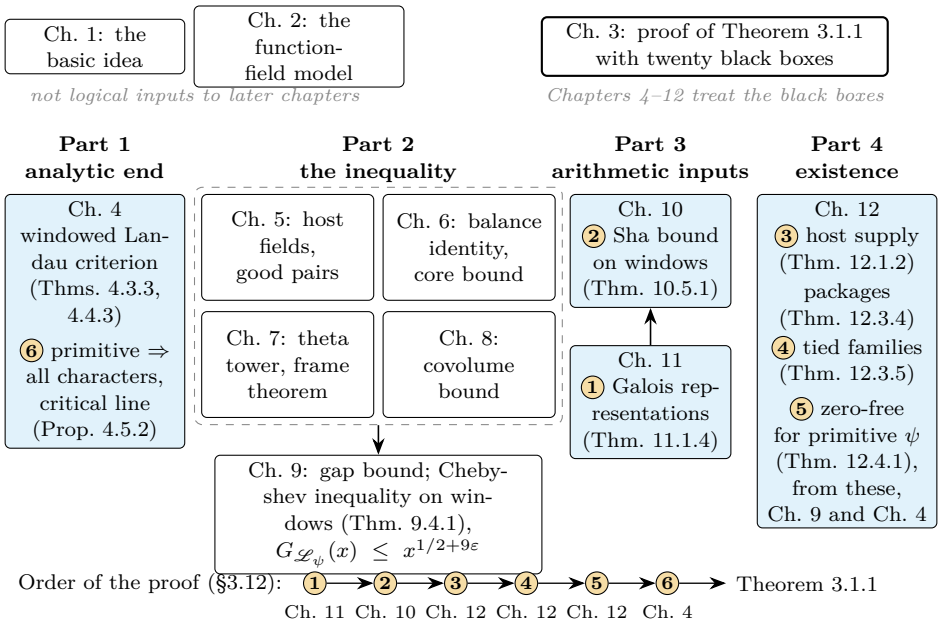


FIGURE I.1. A reading map of the book. Chapters 4–12 are grouped by the four parts of the proof. The numbered marks give the order of the six steps of the proof of Theorem 3.1.1 in Section 3.12, which is not the order of the chapters. Figure 3.1 gives the dependencies between the theorems. Schematic.

What is proved in these notes

For each result of the argument these notes give the statement and the place where it enters the proof. They then give one of the following: a complete proof, an outline of Alpöge's route, a description of the inputs of the proof, an explanatory account of a standard method, or a citation. The attribution of each statement and the kind of each proof are marked where the statement occurs (Section 4.1). Table 3.3 shows what the notes give for each black box, and Remark 3.14.8 lists the points that the later chapters leave open. Two open points are the main ones. The first is the step from the windows $[Y, Y^2]$ of the Sha bound to the windows $[W, W^{3/2}]$ of the supply of hosts: the smaller window lies in the larger one and Theorem 12.4.1 accepts every window exponent $\lambda > 1$, so the open question is only whether the scale also changes between the two statements (Question 12.1.3, Remark 12.1.4). The second is Hypothesis (H2) of the sign-only argument, a vanishing statement in étale cohomology that the bound of Theorem 2.10.2 assumes and that is not proved (Question 2.10.10).

Prerequisites and conventions

Chapters 1 and 2 use the algebraic geometry of curves: divisors, the Riemann–Roch theorem and function fields. The sections on the sign-only argument also use ℓ-adic sheaves on varieties over finite fields, the Grothendieck–Lefschetz trace formula, Artin's vanishing theorem and Poincaré duality.

Chapters 4, 5 and 6 assume the theory of Dirichlet characters and their L-functions at the level of a first course in analytic number theory [**Dav00, MV07**], which includes the functional equation and the explicit formula. Remark 5.5.2 lists the further classical results that Chapter 5 cites. Chapters 7, 9 and 12 use Hermitian lattices and their slopes [**Bos96**], algebraic automorphic forms on definite unitary groups [**Gro99**] and the theta correspondence for unitary dual pairs [**How89, Ral84**]. Chapter 8 cites results on integral models and compactifications of Shimura varieties and on Arakelov geometry. Chapter 10 uses Galois cohomology of number fields, which includes Poitou–Tate duality [**Tat63, Poi67**]. Chapter 11 cites the construction of Galois representations attached to automorphic representations of GL_n over CM fields [**HT01, Shi11, CH13**]. Results from these subjects are cited where they are used.

Theorems, propositions, definitions, remarks and questions share one counter within each section: Theorem 3.1.1 is the first numbered statement of Section 3.1. Equations are numbered within sections. The conventions for attributions, for the four kinds of proof block and for schematic definitions are in Section 4.1. They hold in Chapters 3–12. Section 3.3 recalls the terms in which the black boxes are stated: levels, windows, scales, hosts, packages, families and ties. From Section 2.8.3 to the end of Section 2.12 the letters M, N, R, k and t have the meaning that they have in the identity $qN = M^k + R$. In the first part of Chapter 2 they have a different meaning.

Organization of the notes

The front matter contains a preface and Alpöge's note, which is reproduced in full. Chapter 1 is the informal overview. Table 1.1 gives, for each step of the overview, the chapter that develops it. Chapter 2 is the function-field model. Its Section 2.13 compares each step of the model with the part of the argument that has the same role, and rates each comparison. Chapter 3 states the main theorem and all black boxes and gives the proof. Chapters 4–12 treat the black boxes in the order of the section on the structure of the proof, above. Appendix A describes the software and the AI methods that produced the text.

Chapters 1 and 2 are not logical inputs to the later chapters. A reader who wants only the structure of the argument can start with Chapter 3. Chapters 4–12 use from Chapter 3 only the statement of Theorem 3.1.1, and each of them opens with a paragraph that says where its results enter the proof. The lattices and automorphic forms of Chapter 7 enter the

analytic argument of Chapters 6 and 9 only through the slope record (Definitions 7.3.1 and 7.3.2).

AI and Software Discussion

AI agents assembled and rewrote the text of these notes on 2026-11-21 from drafts that AI agents wrote on 2026-11-20 and 2026-11-21. The models for the assembly were Claude Fable 5.2 (`claude-fable-5-2`) [**Ant26b**] and Claude Opus 5.5 (`claude-opus-5-5`) [**Ant26c**], in Claude Code [**Ant26a**]. The model `claude-fable-5-2` drafted this introduction. Appendix A gives the methods, the citations for all software that was used, which includes the software used for the drafts, and the token accounting. No human has reviewed the text.

CHAPTER 1

The basic idea

This chapter is an overview of the whole book. It describes the shape of the argument for the Riemann Hypothesis announced by L. Alpöge [**Alp26b**], and it does so by following one classical proof. Over a finite field the Riemann hypothesis for curves has an elementary proof, due to Stepanov [**Ste69**] and put in general form by Bombieri [**Bom74**]. One builds a nonzero function that vanishes to order at least m at every rational point and whose number of poles is controlled. Since a nonzero function has as many zeros as poles, one gets

$$(\text{number of rational points}) \cdot m \ \leq \ (\text{number of poles}),$$

and the choice $m \approx \sqrt{q}$ gives the Weil bound.

Alpöge describes his argument as a descendant of Stepanov's. He names the main difficulty: over the integers multiplicity is "hard to find", and corresponds to something like differentiating whole numbers [**Alp26b**, p. 1]. The purpose of the chapter is to make that sentence precise enough to read the rest of the book. Nothing here is used later as a proved input: the chapter states what each later chapter is for, and every step ends with a pointer to the place where it is developed. Table 1.1 collects these pointers.

Three kinds of statement occur, and they should be kept apart. Statements proved here are propositions with proofs. Statements proved elsewhere in the book are quoted with a reference to the chapter that proves them. Statements about Alpöge's argument that are only reports, of the note [**Alp26b**] or of statements read in the Lean archive [**Alp26a**], are marked as reports, and the interpretation that connects them is marked as an interpretation.

1.1. Stepanov's method over a finite field

Throughout this section C is a smooth, projective, geometrically irreducible curve of genus g over a finite field $\mathbb{F}_q$ of characteristic p, and $N = \#C(\mathbb{F}_q)$. For a divisor D on C defined over $\mathbb{F}_q$, let $\mathcal{L}(D)$ be the $\mathbb{F}_q$-vector space of rational functions f on C defined over $\mathbb{F}_q$ with $f = 0$ or $\operatorname{div}(f) + D \geq 0$, and let $\ell(D) = \dim \mathcal{L}(D)$.

Two facts from the theory of curves are used; for both see [**Vak25**, Exercises 16.3.E and 18.4.B, §18.5.3]. The first is Riemann's inequality $\ell(D) \geq \deg D + 1 - g$ for every D, with equality when $\deg D \geq 2g - 1$; both statements are part of the Riemann–Roch theorem. The second is

Step of the overview	Where it is developed
Stepanov's method over $\mathbb{F}_q$ (Section 1.1)	Chapter 2, Sections 2.3–2.5
The sign-only argument and $qN = M^k + R$	Chapter 2, Sections 2.8–2.12
The dictionary (Section 1.2)	Chapter 2, Section 2.13; Chapter 3
The quantity to bound, a one-sided criterion (Section 1.3)	Chapter 4
Seeing every prime: characters and host fields	Chapters 5 and 6
Divisibility is at most size (Section 1.4)	Chapters 6, 7 and 8
Multiplicity and the exponent 1/2 (Section 1.5)	Chapter 9
Theta lifts (Section 1.6)	Chapter 7
The costs the outline omits (Section 1.7)	Chapters 10, 11 and 12
The argument in one place	Chapter 3

TABLE 1.1. The steps of this chapter and the chapters that develop them.

that a nonzero rational function has as many zeros as poles, counted with multiplicity and degree.

PROPOSITION 1.1.1 (Stepanov–Bombieri bound). *Let C be a smooth, projective, geometrically irreducible curve of genus g over $\mathbb{F}_q$. Suppose that q is a square and that $q > (g+1)^4$. Then*

$$\#C(\mathbb{F}_q) \;\le\; q + (2g+1)\sqrt{q}.$$

This is the case $\sigma = \mathrm{id}$ of Theorem 2.3.12, and the complete proof is given there. What follows is a sketch. It is meant to show where each of the three ideas enters, because each of them has a counterpart in the rest of the book.

Sketch of the proof. Fix a rational point x_0, write $m = \sqrt{q}$, a power of p, and let A and B be nonnegative integers with $mA < q$, to be chosen at the end. The whole sketch rests on the additivity of $f \mapsto f^m$, which holds because m is a power of the characteristic. This additivity replaces differentiation.

The first idea is independence. Choose a basis $s_1, \dots, s_r$ of $\mathcal{L}(Ax_0)$ whose elements have pairwise distinct pole orders $n_1 < \cdots < n_r \le A$ at x_0. If $w_1, \dots, w_r$ are rational functions with $\sum_i s_i^m w_i^q = 0$, then every w_i is zero. Indeed a nonzero term has valuation $-mn_i + q\,\mathrm{ord}_{x_0}(w_i)$ at x_0, which

is congruent to $-mn_i$ modulo q, and the integers mn_i are distinct and lie in $[0, q)$. So the nonzero terms have pairwise distinct valuations, and they cannot cancel. In Chapter 2 this is Propositions 2.3.2 and 2.3.4.

The second idea is that linear algebra produces an auxiliary function. Consider the $\mathbb{F}_q$-linear map

$$\Phi\colon \mathcal{L}(Bx_0)^r \to \mathcal{L}\big((mA+B)x_0\big), \qquad (v_1,\dots,v_r) \mapsto \sum_{i=1}^{r} s_i^m v_i.$$

By Riemann–Roch its source has dimension at least $(A+1-g)(B+1-g)$ and its target has dimension $mA+B+1-g$ when $mA+B \geq 2g-1$. So if

$$(A+1-g)(B+1-g) > mA+B+1-g, \qquad A+1-g>0, \tag{1.1.1}$$

then Φ has a nonzero kernel element $(v_1,\dots,v_r)$. Put

$$f = \sum_{i=1}^{r} s_i^m v_i^q.$$

By the first idea $f \neq 0$. Its only pole is at x_0, of order at most $mA+qB$.

The third idea is to count zeros against poles. Let $P \neq x_0$ be a rational point. The values $v_i(P)$ lie in $\mathbb{F}_q$, so $v_i(P)^q = v_i(P)$ and $f(P) = \Phi(v_1,\dots,v_r)(P) = 0$. But $f = \big(\sum_i s_i v_i^{q/m}\big)^m$ is an m-th power of a function that is regular at P, so one vanishing condition gives vanishing to order m. Counting zeros against poles,

$$m\,(N-1) \;\leq\; mA+qB. \tag{1.1.2}$$

With $A = m-1$ and $B = m+2g$, condition (1.1.1) becomes $m > (g+1)^2$, which is the hypothesis, and (1.1.2) becomes $N \leq q + (2g+1)\sqrt{q}$. (Remark 2.3.13 checks this choice against the general statement of Chapter 2.)

Figure 1.1 shows the count behind (1.1.2) for small numbers.

The balance. The same balance of parameters reappears in Section 1.5, so we isolate it here. Inequality (1.1.2) says $N \leq 1 + A + qB/m$. The constraint $mA < q$ forces $A < q/m$, and the kernel condition needs B of size about m plus a constant times g. A short calculation with $A = q/m - 1$ and $B = m + 2g$ shows that (1.1.1) holds only when $m^2 \lesssim q(g+1)/g$, and then

$$N \;\leq\; q + (2g+1)\,\frac{q}{m} + O(1).$$

So the error term is q/m: more multiplicity helps. With these values of A and B the kernel condition fails when m exceeds about $\sqrt{q}$. A larger B restores the condition and adds to the order of the pole. In general the error is at most $(g+1)q/m + 2gm$ (Proposition 2.3.9): the first term is the price of the vanishing conditions and falls as m grows, and the second is the price of the kernel condition and rises. The two terms have the same order at $m \approx \sqrt{q}$, and no choice of the parameters makes the method give an error $o(\sqrt{q})$ (Proposition 2.3.10). Figure 1.2 computes this balance exactly in one case.

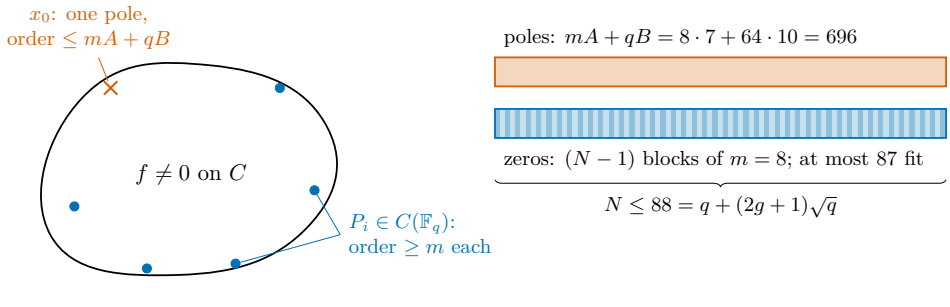


FIGURE 1.1. Zeros against poles in the sketch of Proposition 1.1.1. The auxiliary function f vanishes to order at least m at every rational point other than x_0 and has a single pole, of order at most $mA + qB$, at x_0. The ledger on the right is drawn to scale for $q = 64$, $g = 1$, $m = 8$, $A = 7$, $B = 10$, which satisfy the hypotheses and condition (1.1.1): at most 87 blocks of length 8 fit under 696, so $N \le 88 = q + (2g+1)\sqrt{q}$. The curve is schematic.

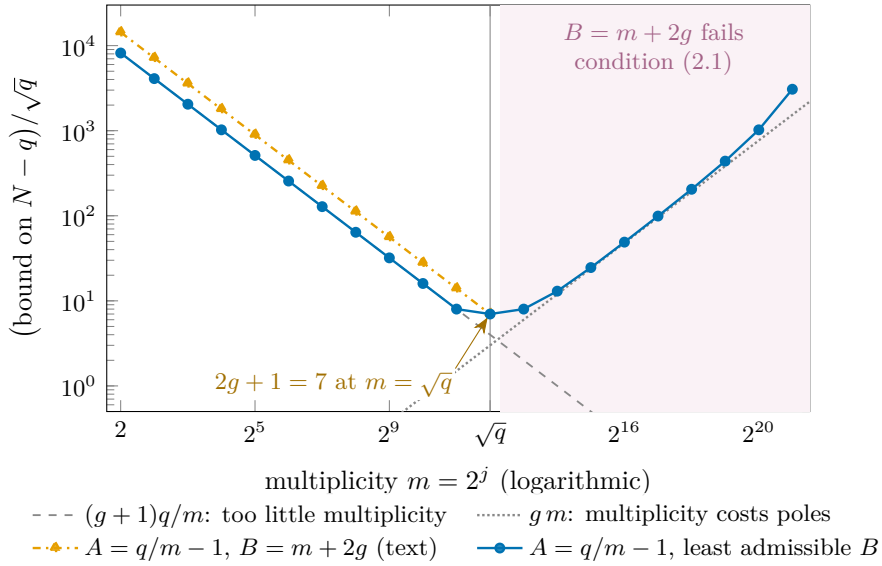


FIGURE 1.2. The bound on $N - q$ given by (1.1.2), divided by $\sqrt{q}$, as a function of the multiplicity $m = 2^j$, for $p = 2$, $q = 2^{24}$, $g = 3$ and $A = q/m - 1$, in exact integer arithmetic. With $B = m + 2g$ (triangles) the bound is $(2g+1)q/m$, but condition (1.1.1) fails for $m > \sqrt{q}$ (shaded). With the least B satisfying (1.1.1) (dots) the bound is about $(g+1)q/m + gm$ (grey lines). Both are smallest at $m = \sqrt{q}$, where the bound is $q + (2g+1)\sqrt{q}$.

From the point count to the zeta function.

REMARK 1.1.2 (From the upper bound to the Riemann hypothesis). Let $\alpha_1, \dots, \alpha_{2g}$ be the reciprocal roots of the numerator of the zeta function of C, so that $\#C(\mathbb{F}_{q^n}) = q^n + 1 - \sum_i \alpha_i^n$ for all $n \geq 1$ (Theorem 2.2.6). Proposition 1.1.1, applied over $\mathbb{F}_{q^n}$ for even n large, gives an *upper* bound $\#C(\mathbb{F}_{q^n}) \leq q^n + (2g+1)q^{n/2}$. Bombieri obtains a matching lower bound $\#C(\mathbb{F}_{q^n}) \geq q^n - c\, q^{n/2}$, with c independent of n, by applying the upper bound to the twists of a Galois cover of C [**Bom74**]. This is Section 2.4. With both bounds, $\sum_i (\alpha_i^2)^k = O(q^k)$ for k large. The power series $\sum_k (\sum_i \alpha_i^{2k}) t^k = \sum_i \alpha_i^2 t/(1-\alpha_i^2 t)$ then converges for $|t| < 1/q$, so it has no pole there, and $|\alpha_i| \leq \sqrt{q}$ for every i. The complete argument is Theorem 2.5.3.

The argument that works for ζ in Proposition 1.3.4 below does not derive $|\alpha_i| \leq \sqrt{q}$ from the upper bound alone. The reason is the first essential difference between the finite-field model and the integers.

REMARK 1.1.3 (The Pringsheim argument does not suffice for curves). Let $\alpha_1, \dots, \alpha_{2g}$ be the reciprocal roots of the numerator of the zeta function of C. For even n large, the upper bound $\#C(\mathbb{F}_{q^n}) \leq q^n + (2g+1)q^{n/2}$ of Remark 1.1.2 says only that $\sum_i \alpha_i^n \geq 1 - (2g+1)q^{n/2}$. If one α_i were real with $|\alpha_i| > \sqrt{q}$, then for even n the sum would be large and *positive*, so the upper bound would not be contradicted. Pringsheim's theorem on power series with nonnegative coefficients, applied to $\sum_k (\sum_i \alpha_i^{2k} + (2g+1)q^k - 1)t^k$ after changing finitely many coefficients, shows only this: if some $|\alpha_i| > \sqrt{q}$, then some α_i of maximal modulus is real. It cannot rule that case out. For ζ a one-sided bound suffices (Proposition 1.3.4). The difference lies in the direction of the inequality. The proof for ζ does not use the absence of zeros of ζ on the real segment $(0,1)$. In the function-field model an upper bound for the counterpart of G_ζ is a *lower* bound for point counts (Proposition 2.6.10). For a bound in that direction a real root of maximal modulus enters with a sign that the bound forbids, and the one-sided bound suffices (Theorem 2.6.9). The proof of Proposition 1.3.4 excludes real zeros of ζ in the same way, by the sign of a residue. The Stepanov–Bombieri bound is the bound in the other direction, and Bombieri uses a second argument for curves. The remark shows that this method fails. It does not show that no other argument derives the conclusion from the upper bound. Section 2.6 makes all of this exact. There the one-sided bound proves "the Riemann hypothesis or a real zero of the L-polynomial in one explicit interval" (Theorems 2.6.2 and 2.6.3), and a formal example shows that the second case cannot be excluded by the formal properties of zeta functions (Proposition 2.6.5).

1.2. The dictionary

The strategy is to translate each line of the sketch of Proposition 1.1.1 into arithmetic.

Curve over $\mathbb{F}_q$	Integers
rational point P	prime p
vanishing to order m at P	divisibility by $p^{m(\cdots)}$
number of poles of f	log of the size of an algebraic number
zeros = poles for $f \neq 0$	product formula; slope inequality
upper bound for N	upper bound for a prime sum $G_\zeta(x)$
$v^q = v$ on $\mathbb{F}_q$-points	reduction modulo p
$f = (\cdots)^m$ with m a power of p	a chain of m theta lifts

Each row is explained below. The guiding principle is the rank-one case of the third and fourth rows. Let Δ be a nonzero integer. Then

$$\sum_p \operatorname{ord}_p(\Delta) \log p \;=\; \log|\Delta|. \tag{1.2.1}$$

Read the left side as the number of zeros of Δ, each prime counted with its multiplicity $\operatorname{ord}_p(\Delta)$ and weight $\log p$. Read the right side as the number of poles. If we know in advance that Δ is divisible by p^{v_p} for many primes, and we can bound $|\Delta|$ from above, then

$$\sum_p v_p \log p \;\leq\; \log|\Delta|. \tag{1.2.2}$$

This is "divisibility is at most size", the arithmetic analogue of (1.1.2). The whole method consists of choosing Δ, or a lattice of such numbers, so that the left side is a prime sum we want to bound. Figure 1.3 draws (1.2.1) next to its function-field counterpart.

PROPOSITION 1.2.1 (Chebyshev). *Let $n \geq 1$ be an integer. Then $\sum_{n<p\leq 2n} \log p < 2n \log 2$, the sum running over primes.*

PROOF. Every prime p with $n < p \leq 2n$ divides $(2n)!$ and not $n!^2$, so it divides $\Delta = \binom{2n}{n}$. Apply (1.2.2) with $v_p = 1$ for those primes, and $\Delta < (1+1)^{2n}$. □

Chebyshev's argument is Stepanov's method with multiplicity one and without a good main term. It gives upper bounds of the right order of magnitude for prime sums [**Tch52**], but no asymptotic. Two things must be added: an exact cancellation of main terms, and multiplicity. The four steps below supply them.

The same comparison organises Chapter 3: its Table 3.1 sets Chebyshev's factorial argument beside the parts of the argument that replace it, and Remark 7.6.2 states in what sense the argument is Chebyshev's, twisted by characters and sharpened by a chain of lifts. A second dictionary, from the function-field side and with each row rated, is in Section 2.13.

1.3. The quantity to bound

DEFINITION 1.3.1. Let Λ be the von Mangoldt function and γ Euler's constant. For real $x \geq 1$ put

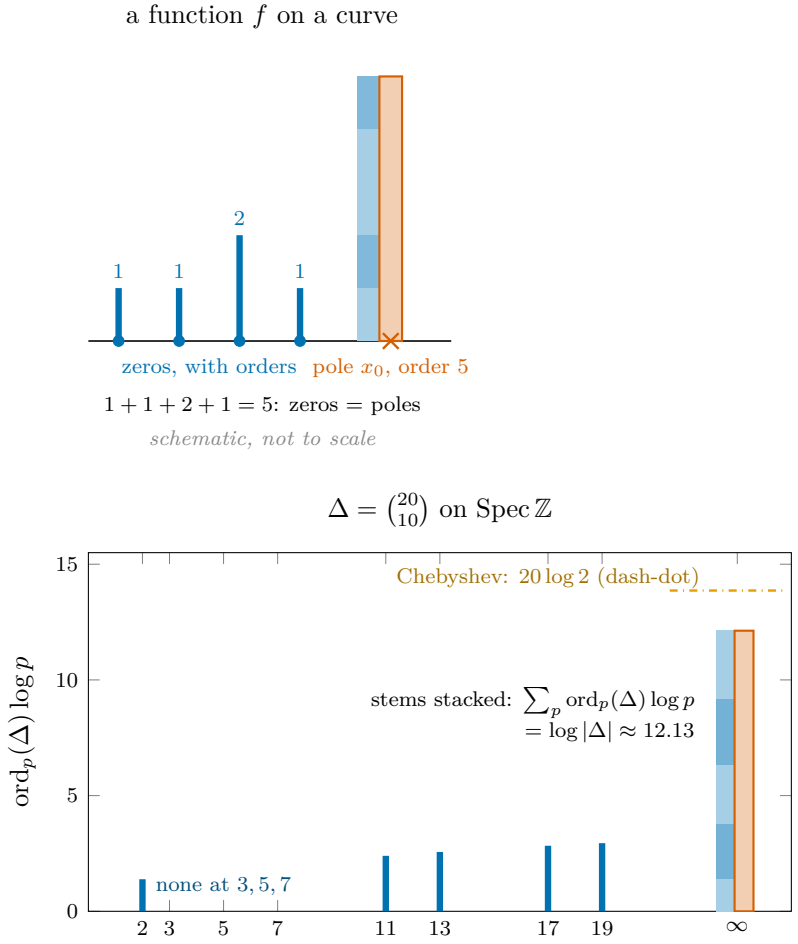


FIGURE 1.3. Zeros equal poles, on a curve and on Spec $\mathbb{Z}$. Top (schematic): a nonzero function has as many zeros as poles, counted with order. Bottom: for $\Delta = \binom{20}{10} = 2^2 \cdot 11 \cdot 13 \cdot 17 \cdot 19$, the stems $\operatorname{ord}_p(\Delta)\log p$ stacked at the archimedean place ∞ add up to $\log|\Delta| \approx 12.13$, which lies below Chebyshev's bound $20\log 2$ (Proposition 1.2.1).

(i) $S_\zeta(x) = \sum_{n\le x} \Lambda(n)\,(x/n - 1)$;
(ii) $G_\zeta(x) = S_\zeta(x) - x\log x + (1+\gamma)\,x$.

These are the smoothed Chebyshev function of ζ and its normalised form. Definition 4.2.1 repeats them together with their twists by Dirichlet characters. Figure 1.4 shows G_ζ and the size the Riemann Hypothesis predicts for it.

Alpöge's note calls the switch to this sum one of the decisive steps: "Guinand–Weil had the sign" [**Alp26b**, p. 3, step 5]. Since $\sum_{n\le x}\Lambda(n)(x/n - 1) = x\sum_{n\le x}\Lambda(n)\int_n^x t^{-2}\,dt$, we have $S_\zeta(x) =$

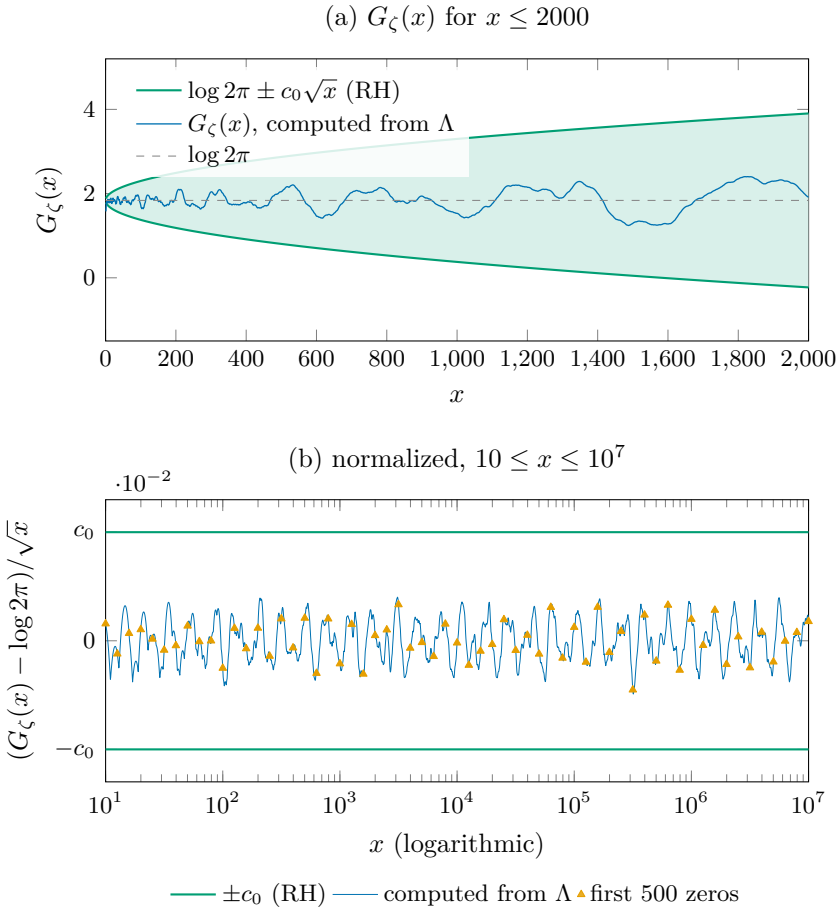


FIGURE 1.4. The quantity to bound. (a) $G_\zeta(x)$ for $x \le 2000$, computed from the von Mangoldt function, inside the band $\log 2\pi \pm c_0\sqrt{x}$, where $c_0 = 2 + \gamma - \log 4\pi \approx 0.046$ equals $\sum_\rho |\rho|^{-2}$ under the Riemann Hypothesis. (b) $(G_\zeta(x) - \log 2\pi)/\sqrt{x}$ for $10 \le x \le 10^7$ (logarithmic axis), with the explicit formula evaluated from the first 500 zeros (triangles). The plot samples 2500 values of x; it illustrates the prediction and proves nothing.

$x \int_1^x \mathrm{Ch}(t)\, t^{-2}\, dt$, where $\mathrm{Ch}(t) = \sum_{n \le t} \Lambda(n)$ is Chebyshev's function. So S_ζ is a smoothed form of Ch.

PROPOSITION 1.3.2 (Mellin transform). *Let s be a complex number with $\mathrm{Re}\, s > 1$. Then*

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

and

$$\int_1^\infty G_\zeta(x)\, x^{-s-1}\, dx \;=\; -\frac{\zeta'(s)}{\zeta(s)}\cdot\frac{1}{s(s-1)} - \frac{1}{(s-1)^2} + \frac{1+\gamma}{s-1}.$$

PROOF. For each n, $\int_n^\infty (x/n-1)\, x^{-s-1}\, dx = n^{-s}\big(\frac{1}{s-1} - \frac{1}{s}\big) = \frac{n^{-s}}{s(s-1)}$. Summing against $\Lambda(n)$, which is allowed by absolute convergence, gives the first identity, since $\sum_n \Lambda(n) n^{-s} = -\zeta'(s)/\zeta(s)$. The second follows from $\int_1^\infty x\log x \cdot x^{-s-1}\, dx = (s-1)^{-2}$ and $\int_1^\infty x\cdot x^{-s-1}\, dx = (s-1)^{-1}$. □

The right side of the second identity is meromorphic in $\operatorname{Re} s > 0$. Near $s = 1$ we have $-\zeta'/\zeta(s) = (s-1)^{-1} - \gamma + O(s-1)$ and $1/(s(s-1)) = (s-1)^{-1} - 1 + O(s-1)$, so the first term is $(s-1)^{-2} - (1+\gamma)(s-1)^{-1} + O(1)$. This is exactly what is subtracted: the right side is regular at $s = 1$. Its only poles in $\operatorname{Re} s > 0$ are at the zeros ρ of ζ, each with residue $-\operatorname{mult}(\rho)/(\rho(\rho-1)) \neq 0$, where $\operatorname{mult}(\rho)$ is the multiplicity of the zero. This explains the choice of G_ζ. The terms $x\log x - (1+\gamma)x$ remove the pole at $s = 1$, and the kernel $1/(s(s-1))$ makes the sum over zeros in the explicit formula converge absolutely. Under the Riemann Hypothesis, $G_\zeta(x) = O(\sqrt{x})$, by the explicit formula $G_\zeta(x) = -\sum_\rho x^\rho/(\rho(\rho-1)) + O(1)$ and $\sum_\rho |\rho|^{-2} < \infty$; see [**Ing32**] and Remark 4.2.3. The converse needs only a one-sided bound, by the following theorem of Landau.

THEOREM 1.3.3 (Landau [**Lan05**]; see [**Ing32**]). *Let $x_0 \geq 1$, and let $h\colon [x_0,\infty) \to [0,\infty)$ be locally integrable. Suppose $F(s) = \int_{x_0}^\infty h(x)\, x^{-s-1}\, dx$ converges for some real s and diverges for some real s, and let σ_c be the infimum of the real s for which it converges. Then F, which is analytic in $\operatorname{Re} s > \sigma_c$, has no analytic continuation to any neighbourhood of $s = \sigma_c$.*

The substitution $x = e^u$ turns this into the statement for Laplace transforms of nondecreasing functions, which is the form quoted as Theorem 4.3.2 in Chapter 4.

PROPOSITION 1.3.4 (One-sided criterion). *Let $0 < \theta < 1$ and $x_0 \geq 1$. Suppose $G_\zeta(x) \leq x^\theta$ for all $x \geq x_0$. Then $\zeta(s) \neq 0$ for $\operatorname{Re} s > \theta$.*

PROOF. Put $h(x) = x^\theta - G_\zeta(x) \geq 0$ on $[x_0,\infty)$. Since $G_\zeta(x) = O(x\log x)$, the integral $F(s)$ of Theorem 1.3.3 converges for $\operatorname{Re} s > 1$. By Proposition 1.3.2, F equals $x_0^{\theta-s}/(s-\theta)$ minus the right side of the second identity, plus an entire function (the integral over $[1, x_0]$). This expression is meromorphic in $\operatorname{Re} s > 0$, with poles only at $s = \theta$ and at the zeros of ζ.

The residue of the expression at $s = \theta$ is $1 + m/(\theta(\theta-1))$, where $m \geq 0$ is the order of vanishing of ζ at θ. It is not zero: if $m \geq 1$ then $m/(\theta(1-\theta)) \geq 4$. So the expression has a pole at $s = \theta$. At a real zero $\beta \in (0,1)$ of ζ of multiplicity m, other than θ, its residue is

$$\frac{m}{\beta(\beta-1)} \;=\; -\frac{m}{\beta(1-\beta)} \;<\; 0.$$

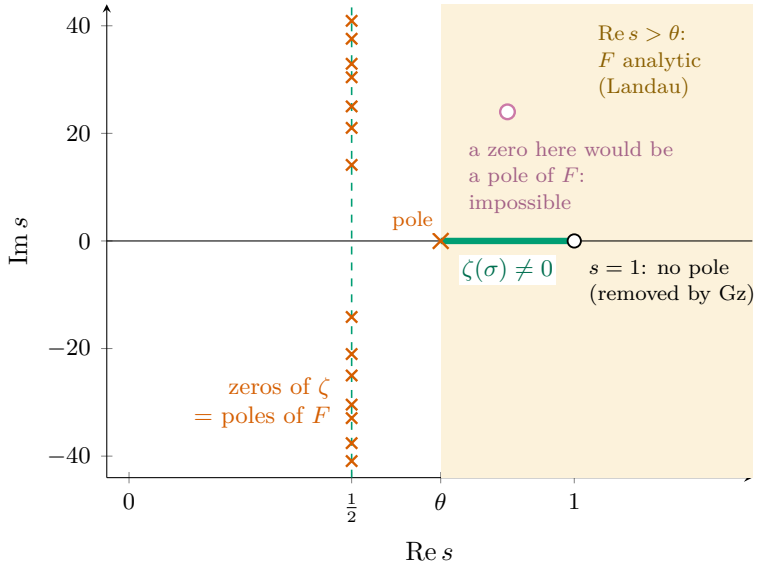


FIGURE 1.5. The proof of Proposition 1.3.4. In $\operatorname{Re} s > 0$ the Mellin transform F of $x^\theta - G_\zeta(x)$ has poles only at $s = \theta$ and at the zeros of ζ (crosses, at their true ordinates); $s = 1$ is not a pole. A real zero in $(\theta, 1)$ would give F a pole with negative residue, which $F \geq 0$ on the real axis forbids; so Landau's theorem makes F analytic in $\operatorname{Re} s > \theta$ (shaded), and a zero there (open circle) cannot exist. The value $\theta = 0.7$ and the open circle are illustrative.

Let σ_c be the infimum of the real s for which the integral converges. The integral cannot converge for every real s: then F would be analytic in $\operatorname{Re} s > 0$, and it has a pole at $s = \theta$. So Theorem 1.3.3 applies, and σ_c is a real pole of the meromorphic expression. Suppose $\sigma_c > \theta$. The point $s = 1$ is not a pole, so $\sigma_c = \beta$ is a zero of ζ in $(\theta, 1)$. For real $s > \beta$ we have $F(s) \geq 0$, because $h \geq 0$. But the residue of F at β is negative, so $F(s) \to -\infty$ as s decreases to β. This is a contradiction. Hence $\sigma_c \leq \theta$, so the integral converges and is analytic in $\operatorname{Re} s > \theta$. It follows that the meromorphic expression has no poles there, and so ζ has no zeros there. □

The proof does not use the classical fact that ζ has no zero on the real segment $(0, 1)$, which follows from $(1 - 2^{1-\sigma})\zeta(\sigma) = \sum_{n \geq 1} (-1)^{n-1} n^{-\sigma} > 0$ for $0 < \sigma < 1$. A real zero in $(\theta, 1)$ is excluded by the sign of its residue alone. The same mechanism serves for L-functions, whose real zeros are not excluded in advance (Remark 4.3.7), and in the function-field model (Theorem 2.6.9).

Figure 1.5 shows the argument in the s-plane.

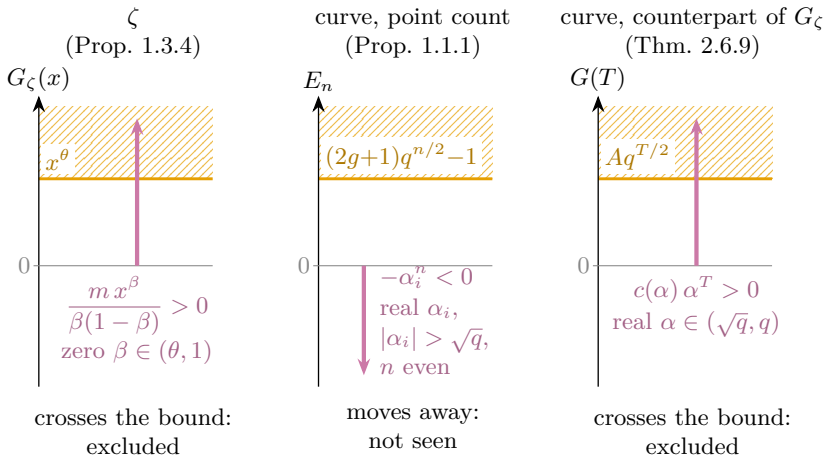


FIGURE 1.6. The direction of the inequality (schematic; only signs are drawn). Each panel shows a one-sided upper bound (orange line, the hatched values are forbidden) and the term that a bad real zero or reciprocal root adds to the bounded quantity. For ζ a zero $\beta \in (\theta, 1)$ of multiplicity m adds a positive term to the explicit formula for G_ζ; in the proof of Proposition 1.3.4 the same sign appears as the negative residue $m/(\beta(\beta-1))$ of the Mellin transform of $x^\theta - G_\zeta(x)$. For a curve, with $E_n = \#C(\mathbb{F}_{q^n}) - q^n - 1 = -\sum_i \alpha_i^n$, a real root of modulus above $\sqrt{q}$ pushes E_n down for even n, away from the bound of Proposition 1.1.1 over $\mathbb{F}_{q^n}$ (Remark 1.1.3). The counterpart $G(T)$ of G_ζ (Definition 2.6.7) receives $c(\alpha)\alpha^T$ with $c(\alpha) > 0$ (Proposition 2.6.8), so an upper bound for $G(T)$, which corresponds to a lower bound for point counts, excludes it (Remark 2.6.11); Figure 2.18 shows this reversal with data for $q = 9$.

Combining Proposition 1.3.4 with the functional equation: if for every $\varepsilon > 0$ one has $G_\zeta(x) \le x^{1/2+\varepsilon}$ for all large x, then every nontrivial zero of ζ lies on $\operatorname{Re} s = 1/2$. So the target is an upper bound $G_\zeta(x) \le x^{1/2+\varepsilon}$. In the dictionary of Section 1.2, G_ζ plays the role of the error $\#C(\mathbb{F}_{q^n}) - q^n - 1 = -\sum_i \alpha_i^n$ of a point count, with the sign reversed: its exact counterpart is a tail-smoothed sum of the numbers $q^n + 1 - \#C(\mathbb{F}_{q^n}) = \sum_i \alpha_i^n$, computed in Section 2.6 (Definition 2.6.7, Proposition 2.6.10). Because of the reversal, an upper bound for the counterpart of G_ζ is a *lower* bound for point counts, so the one-sided bound of this section is not the transcription of the Stepanov–Bombieri upper bound (Figure 1.6). Proposition 1.3.4 plays the role of Theorem 2.6.9, and for a bound in this direction no second bound is needed (Remark 1.1.3).

REMARK 1.3.5 (What the argument uses in place of this criterion). The criterion that the argument uses differs from Proposition 1.3.4 in two ways, and Chapter 4 develops both.

(1) The twisted sum of one Dirichlet character need not be real, and real zeros of $L(s,\chi)$ with χ real are not excluded in advance. The argument works with a *positive* combination of G_ζ and the real parts of twisted sums, over a finite set of primitive characters closed under inversion (Definition 4.4.1). With positive weights every real zero contributes a residue of one sign. The residue argument in the proof of Proposition 1.3.4 then applies, so a one-sided bound still suffices (Remark 4.3.7).
(2) The inequality is needed, and is proved, only on windows $[Y, Y^{1+\delta}]$ with Y arbitrarily large. The criterion in that form is Theorem 4.3.3. The classical form of Landau's theorem is not enough, because the bounds that the argument proves hold only on windows (Remark 4.3.8).

1.4. Divisibility is at most size

1.4.1. The brief version. One builds a lattice of algebraic numbers attached to a theta function on an abelian variety with complex multiplication, and maps it to another lattice. Each image is forced to be divisible by many primes, to a known order. Its archimedean size is also known. The arithmetic analogue of "zeros = poles" then gives

$$\underbrace{\textstyle\sum_p (\text{forced order at } p)\,\log p}_{\text{arithmetic side}} \;\le\; \underbrace{\log(\text{size})}_{\text{archimedean side}} .$$

Figure 1.7 maps the steps and their status. The arithmetic side is, up to error, $b\,S_\zeta(a/b)$ for parameters a, b. The archimedean side is computed from periods. The main terms of the two sides cancel, by a period formula of Chowla–Selberg type, and what remains is an upper bound for $G_\zeta(a/b)$.

1.4.2. What the divisibility says about the smoothed sum. The arithmetic side has the shape of the prime factorization of a factorial. By Legendre's formula [**Leg08**], for an integer $a \ge 1$,

$$\operatorname{ord}_p(a!) = \sum_{i\ge 1} \lfloor a/p^i \rfloor, \qquad \text{so} \qquad \log a! = \sum_{n\le a} \Lambda(n)\,\lfloor a/n \rfloor.$$

Drop the floors and subtract a constant $b > 0$ from each term, keeping only positive parts. The result is

$$\sum_{n\le a/b} \Lambda(n)\Big(\frac{a}{n} - b\Big) \;=\; b\,S_\zeta(a/b) \;=\; a\Big(\log\frac{a}{b} - 1 - \gamma\Big) + b\,G_\zeta(a/b), \tag{1.4.1}$$

where the second equality is Definition 1.3.1. So $b\,S_\zeta(a/b)$ is the logarithm of a "smoothed factorial". Figure 1.8 compares the factorial with its smoothing.

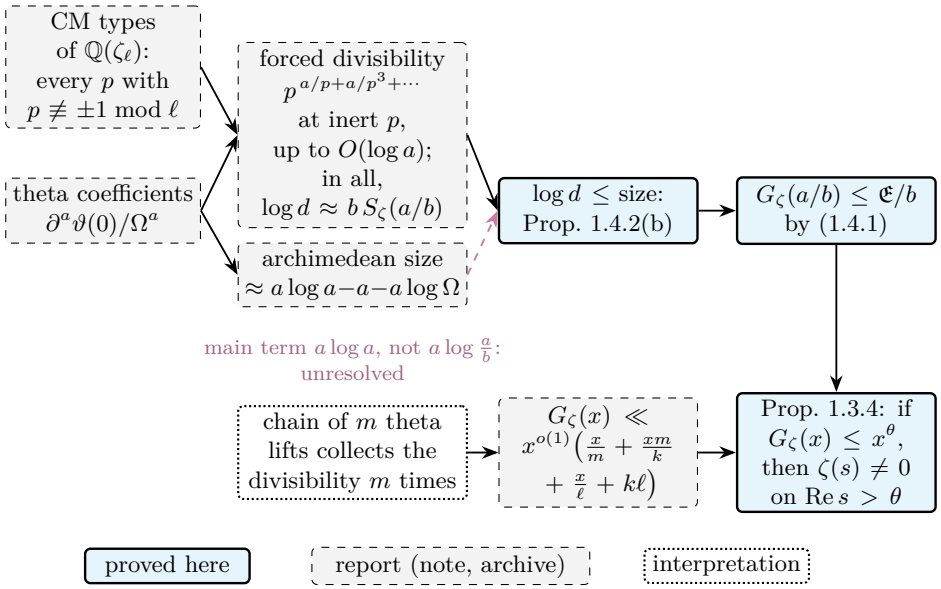


FIGURE 1.7. The outline of Sections 1.3–1.6, in the abelian form reported in Section 1.4, with the status the chapter gives each step (schematic). Solid boxes are proved in this chapter; dashed boxes report the note [**Alp26b**] or statements of the archive [**Alp26a**] that are not on its proof path; the dotted box is an interpretation. The inequality $G_\zeta(a/b) \le \mathfrak{E}/b$ holds if the divisibility at each p is at least $\sum_i (a/p^i - b)_+$ and the size is $a(\log(a/b) - 1 - \gamma) + \mathfrak{E}$; the readable statements do not show how b enters the size (dashed arrow). Balancing the four terms gives at best $x^{3/4}$, and the exponent $\frac{1}{2}$ needs $m \approx \sqrt{x}$. The book replaces $\mathbb{Q}(\zeta_\ell)$ by an imaginary quadratic field inside a host field; Table 1.1 says where each step is developed.

Suppose an algebraic number is known to be divisible, at each prime p, to order at least $\sum_i (a/p^i - b)_+$. Then (1.2.2) bounds $b\,S_\zeta(a/b)$ from above by the logarithm of its size. By (1.4.1), if the size is $a(\log(a/b) - 1 - \gamma) + \mathfrak{E}$, the inequality becomes

$$G_\zeta(a/b) \;\le\; \mathfrak{E}/b.$$

So the content of the bound is the error $\mathfrak{E}$. Chebyshev's binomial coefficients are the oldest example of such numbers, but their sizes have the wrong main term.

In the book this bookkeeping is done once and exactly. The balance identity, Proposition 6.1.4, relates a prime sum of the smoothed Legendre shape to an archimedean main term and a normalised Chebyshev function,

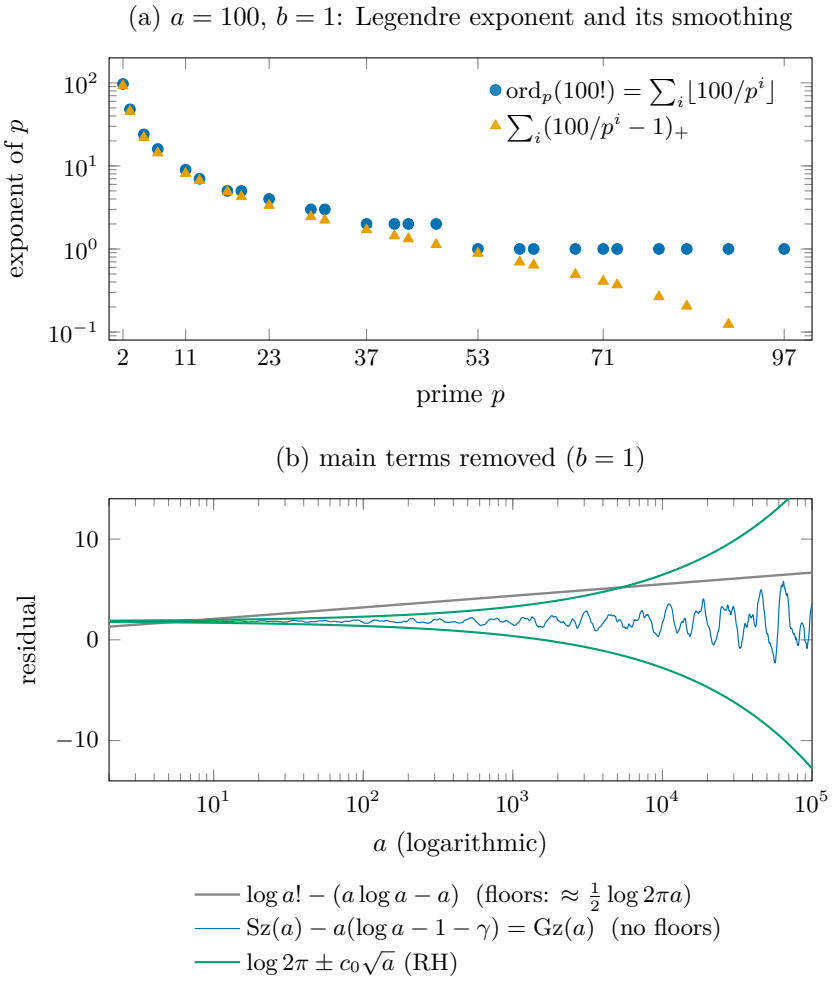


FIGURE 1.8. What the divisibility says about S_ζ. (a) For $a = 100$ and $b = 1$, the exponent $\mathrm{ord}_p(100!)$ (dots) and its smoothing $\sum_i (100/p^i - 1)_+$ (triangles); weighted by $\log p$ and summed over p they give $\log 100!$ and $S_\zeta(100)$. (b) With the main terms removed ($b = 1$), $\log a!$ leaves about $\frac{1}{2} \log 2\pi a$ (grey), while $S_\zeta(a)$ leaves $G_\zeta(a)$ (blue), which fluctuates on the scale $\sqrt{a}$ of the band $\log 2\pi \pm c_0 \sqrt{a}$. The arithmetic content is in this error.

for the characters of the Galois group of a host field; Remark 6.1.5 explains the relation to Legendre's formula.

1.4.3. The numbers used. The following description is a report of the note [**Alp26b**, p. 3, step 2] and of statements in the archive that are readable but not on its proof path [**Alp26a**, `Formalism/StatementCM.lean`,

`Formalism/ProductFormula/StatementOCCM.lean`]. It is not proved here. It describes the first, abelian form of the construction. The end of the subsection compares it with the form that Chapters 6 and 7 develop.

Let X be an elliptic curve with complex multiplication by an imaginary quadratic field K, let ϑ be a theta function on X, let ∂ be the invariant derivation (the derivation dual to an invariant differential), and let Ω be a period. The numbers are the normalized Taylor coefficients $\partial^a\vartheta(0)/\Omega^a$ at the origin. They are algebraic, and their size is about $a\log a - a - a\log\Omega$ in logarithmic terms. At a prime p inert in K at which X has good reduction, the reduction is supersingular, and $\partial^p \equiv 0$ modulo p. Each block of p derivatives therefore contributes a factor p. The note states the resulting divisibility as

$$p^{\,a/p+a/p^3+\cdots-O(\log a)} \;\Big|\; \frac{\partial^a\vartheta(0)}{\Omega^a},$$

an exponent of Legendre type in which only the odd powers of p occur; compare Remark 6.1.5. A single CM curve sees only the inert primes. The archive's abelian statements use abelian varieties with complex multiplication by $\mathbb{Q}(\zeta_\ell)$ for a prime ℓ. A CM type of $\mathbb{Q}(\zeta_\ell)$ is a set of embeddings containing one from each complex-conjugate pair. It is encoded there as a subset $\Phi \subset (\mathbb{Z}/\ell)^\times$ containing exactly one of g and $-g$ for each g, together with a dual type. Which primes p gain is decided by the class of p^i modulo ℓ relative to Φ, and Φ together with its dual type sees every prime except those congruent to ±1, up to an error of about a/ℓ. This is how the full sum $b\,S_\zeta(a/b)$ appears.

The archimedean main term involves $\log\Omega$. The Chowla–Selberg formula [**SC67**] expresses the periods of CM elliptic curves through values of the Gamma function, equivalently through $L'/L(0,\chi)$ for the quadratic character χ of K. Colmez conjectured a generalization to CM abelian varieties, in terms of logarithmic derivatives at $s=0$ of Artin L-functions, and proved it for abelian CM fields up to a rational multiple of $\log 2$ [**Col93**]; Obus removed that ambiguity [**Obu13**]. By the functional equation these are combinations of γ, $\log\pi$ and $L'/L(1,\chi)$, the constants in the main terms of G_ζ and of its twisted analogues. This is the mechanism by which the main terms can cancel.

One point of these notes is unresolved. The readable statements do not show how the parameter b enters the archimedean side: the size quoted above has main term $a\log a$, while (1.4.1) has $a\log(a/b)$. The archive's abelian estimate takes the needed relation between the periods as a hypothesis, $\gamma + \log(\pi/v) - \ell' \le (\log\Omega_1 + \log\Omega_2)/2$ for auxiliary parameters v and ℓ' [**Alp26a**, `Formalism/ProductFormula/StatementOCCM.lean`].

The argument as developed in this book keeps the mechanism and changes the cast. The CM field is an imaginary quadratic field $K = \mathbb{Q}(\sqrt{-d})$ inside a *host field* E, chosen with auxiliary primes. The primes that gain are those inert in K. Fourier analysis on $\mathrm{Gal}(E/\mathbb{Q})$ reaches the other primes (Chapters 5 and 6). The report above suggests a sum of

$L'/L(1,\chi)$ over many characters of Colmez type in the main term, but only $L'(1,\chi_{-d})/L(1,\chi_{-d})$ enters it, through a period of Chowla–Selberg type (Remarks 7.5.2 and 7.5.3).

1.4.4. Bost's slope inequality. To use many numbers at once one replaces (1.2.2) by an inequality for lattices.

Definition 1.4.1. A *Hermitian lattice* is a pair $\overline{L} = (L, \|\cdot\|)$, where L is a free $\mathbb{Z}$-module of finite rank and $\|\cdot\|$ is a Euclidean norm on $L \otimes \mathbb{R}$. Let $\overline{L}$ be a Hermitian lattice with $L \neq 0$.

(i) The *Arakelov degree* is $\widehat{\deg}\,\overline{L} = -\log \operatorname{covol}(L)$, where $\operatorname{covol}(L)$ is the covolume of L in $L \otimes \mathbb{R}$.
(ii) The *slope* is $\mu(\overline{L}) = \widehat{\deg}\,\overline{L}/\operatorname{rk} L$.
(iii) The *maximal slope* $\mu_{\max}(\overline{L})$ is the maximum of $\mu(\overline{L'})$ over nonzero subgroups $L' \subseteq L$ with the restricted norm.

The maximum exists by reduction theory; see [**Bos96**]. For $L = \mathbb{Z}$ with norm $|n| \cdot c$ the degree is $-\log c$. A nonzero integer Δ generates a sublattice of $\mathbb{Z}$ (standard norm) of degree $-\log|\Delta|$.

Proposition 1.4.2 (Slope inequality; cf. [**Bos96**]). *Let $\overline{L}$ and $\overline{M}$ be Hermitian lattices with $L \neq 0$. Let $\varphi\colon L \to M$ be an injective homomorphism, and let $\kappa > 0$ satisfy $\|\varphi(x)\| \leq \kappa\|x\|$ for all $x \in L \otimes \mathbb{R}$.*

(a) Then $\mu(\overline{L}) \leq \mu_{\max}(\overline{M}) + \log\kappa$.
(b) Let $d \geq 1$ be an integer with $\varphi(L) \subseteq dM$. Then $\log d \leq \mu_{\max}(\overline{M}) - \mu(\overline{L}) + \log\kappa$.

Proof of (a). Let $r = \operatorname{rk} L$, and give $\varphi(L)$ the norm restricted from M. The map $\varphi \otimes \mathbb{R}$ multiplies r-dimensional volumes by the product of its singular values, each at most κ. So $\operatorname{covol}\varphi(L) \leq \kappa^r \operatorname{covol}(L)$, that is, $\widehat{\deg}\,\overline{\varphi(L)} \geq \widehat{\deg}\,\overline{L} - r\log\kappa$. By definition of $\mu_{\max}$, $\widehat{\deg}\,\overline{\varphi(L)} \leq r\,\mu_{\max}(\overline{M})$. Combine the two inequalities and divide by r. □

Proof of (b). Give dM the norm restricted from M. Multiplication by d scales covolumes of rank-k sublattices by d^k, so $\mu_{\max}(\overline{dM}) = \mu_{\max}(\overline{M}) - \log d$. Apply (a) to φ viewed as a map $L \to dM$. □

Part (b) is (1.2.2) for lattices. The forced divisibility of the theta coefficients enters as d, with $\log d$ about $b\,S_\zeta(a/b)$, and the sizes enter through $\mu_{\max}(\overline{M})$ and $\log\kappa$. The archive states the inequality of part (a) in [**Alp26a**, `Formalism/LatticeSlopes/Statement.lean`]. Bost formulated the slope method as the Arakelov-geometric form of the auxiliary-function method of transcendence theory [**Bos96**]. That is the place to look for an arithmetic version of "a nonzero function has as many zeros as poles". Figure 1.9 works part (b) out in rank two.

Chapter 7 sets up Hermitian lattices over the ring of integers of a number field (Definition 7.2.1) and quotes Bost's inequality in that generality (Theorem 7.2.5). The slope input of the argument is a comparison of an *average*

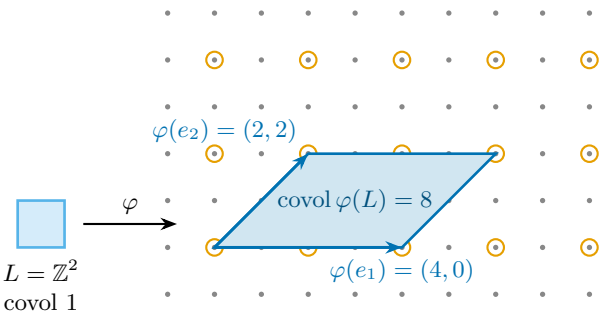


FIGURE 1.9. Proposition 1.4.2(b) in rank two. Here $L = M = \mathbb{Z}^2$ with the standard norm and $\varphi = 2\left(\begin{smallmatrix} 2 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, so $\varphi(L) \subseteq dM$ with $d = 2$. The covolume grows from 1 to $8 \le \kappa^2 \approx 20.9$, where $\kappa = \|\varphi\| = 2\sqrt{3+\sqrt{5}}$, and $\log 2 \le \mu_{\max}(\overline{M}) - \mu(\overline{L}) + \log \kappa = \log \kappa \approx 1.52$.

slope with a slope, the frame theorem (Theorem 7.4.2). Bost's inequality is one step of its proof, and is not applied to each period map. Remark 7.4.3 spells out the difference. The argument does not compute the size of the target lattice. A covolume bound for lattices of integral forms bounds it, and this is where Arakelov geometry enters (Theorem 8.2.3 in Chapter 8).

1.5. Multiplicity is the bottleneck

The following is a report of statements in the archive that are readable but not on its proof path [**Alp26a**, `Formalism/ProductFormula/StatementOCCM.lean`, `OCCMEstimate`, `ZetaQuarterPowerCorollary`]. It is not proved here. After the main terms cancel, the abelian construction gives a bound of the shape

$$G_\zeta(x) \ \ll \ x^{o(1)}\Big(\frac{x}{m} + \frac{xm}{k} + \frac{x}{\ell} + k\ell\Big)$$

plus terms that are negligible for suitable parameters. Here m is the number of times the forced divisibility is collected, k bounds the order of the jets, and ℓ is the prime with CM field $\mathbb{Q}(\zeta_\ell)$. The four terms are read here as follows.

(i) x/m: the fixed loss in the divisibility at each prime is shared among m copies. This is the term q/m of Section 1.1.
(ii) xm/k: each copy costs archimedean size. This is the constraint that m cannot exceed about $\sqrt{q}$ in Section 1.1.
(iii) x/ℓ: a CM type sees a prime only through its class modulo ℓ, and prime powers near ℓ are miscounted.

(iv) $k\ell$: a larger CM field costs size.

Balancing with $m \approx x^{1/4}$, $k \approx x^{1/2}$ and $\ell \approx x^{1/4}$ makes all four terms $x^{3/4}$, and no choice does better, since the product of the four terms is x^3. By Proposition 1.3.4 such a bound would give $\zeta(s) \neq 0$ for $\operatorname{Re} s > 3/4$. The archive's statement of this consequence is the definition of a proposition, not a theorem: its hypothesis is that suitable parameters and CM data exist for all large x, and it is off the proof path. Its conclusion alone, a zero-free half-plane $\operatorname{Re} s > 3/4$, would be far beyond any known zero-free region, so it should be read as a description of the method, not as an established result.

To reach $1/2$, the first two terms alone require $m \approx \sqrt{x}$ and $k \approx x$. The coupling $k\ell$ between the order of the jets and the size of the CM field must then disappear. A further off-path statement [**Alp26a**, `Formalism/HalfLinks/Statement.lean`] includes, among others, the constraints $x^{1-\theta} \leq m$ and $xm \ll x^{\theta}k$. At $\theta = 1/2$ they say $m \geq \sqrt{x}$. This is the chain length named in the note [**Alp26b**, p. 1]. Figure 1.10 carries out this balancing exactly.

In the argument as developed in this book the first two terms survive and the last two are gone. Proposition 9.3.1 bounds the normalised error by terms of sizes x/m, xm/k and m under the constraints $xm \leq k$ and $x^{1/2-\varepsilon} \leq m \leq x^{1/2+\varepsilon/2}$, and Remark 9.4.3 explains why their balance at $m \approx \sqrt{x}$ is the source of the exponent $1/2$. The inequality that results is the Chebyshev inequality on windows, Theorem 9.4.1, which is reached through the gap bound, Theorem 9.2.1. Both are in Chapter 9.

1.6. Theta lifts instead of differentiating integers

1.6.1. Where multiplicity came from over a finite field. In the sketch of Proposition 1.1.1, multiplicity came from one identity: $f = (\sum_i s_i v_i^{q/m})^m$ is an m-th power, because m is a power of the characteristic. In particular, in a local parameter all Hasse derivatives of f of orders $1, \ldots, m-1$ vanish. There is no such operation on $\mathbb{Z}$. One would want a derivation "in the direction of p" whose repeated vanishing at Δ means divisibility of Δ by a high power of p. Buium's p-derivations, such as $\delta_p(n) = (n - n^p)/p$, are one formal answer. Alpöge's argument does not use them. The Lean archive contains δ-ring structures only in p-adic Hodge theory, far from this step.

Stepanov's original argument did use derivatives, and that form of the method, with a chain of about $\sqrt{p}$ derivative conditions, is the closer model for what follows. It is proved in the simplest case in Section 2.7, and the comparison is made there.

1.6.2. What a theta lift is. For theta correspondences see [**How89, Ral84**], and for algebraic modular forms see [**Gro99**]. Let K be a CM field. Let V be a Hermitian space over K that is *totally definite*, meaning positive definite at every infinite place. Then the unitary group $U(V)$ is compact at

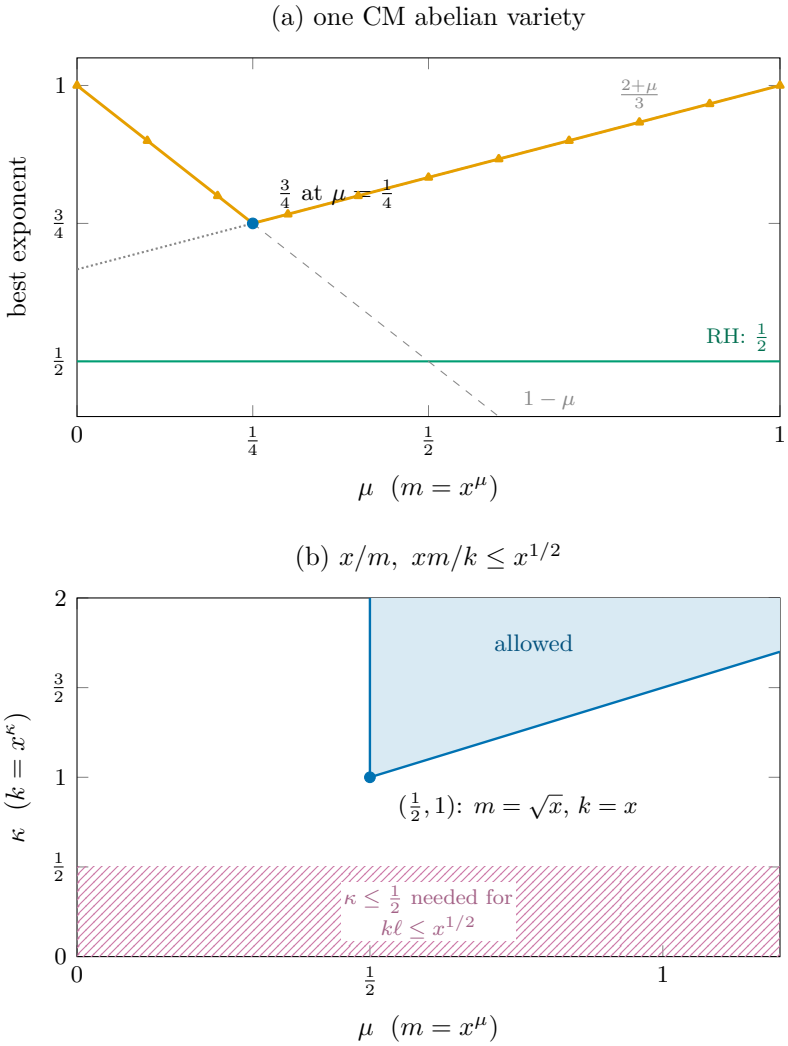


FIGURE 1.10. The exponent budget of Section 1.5, with $m = x^{\mu}$, $k = x^{\kappa}$ and $\ell = x^{\lambda}$. (a) The smallest possible largest exponent of the four terms is $\max(1-\mu, (2+\mu)/3)$, computed by exact linear programming; its minimum is $3/4$, at $\mu = 1/4$. (b) Making x/m and xm/k at most $x^{1/2}$ forces $\mu \ge \frac{1}{2}$ and $\kappa \ge \mu + \frac{1}{2} \ge 1$, which is incompatible with $k\ell \le x^{1/2}$ (hatched). The four terms are a reading of an off-path statement of the archive.

infinity, and an automorphic form on $U(V)$ of a given level is a function on a *finite* set: the classes of lattices in a genus. These are Gross's algebraic modular forms, the higher-dimensional analogue of functions on the finite set of ideal classes of a definite quaternion algebra. A theta correspondence

pairs $U(V)$ with $U(W)$ for a skew-Hermitian space W (a dual pair inside a symplectic group), and sends automorphic forms on one group to the other by integration against a theta kernel.

This section does not claim that the global theta lift between two definite groups is a matrix of representation numbers. That is not the standard description. What follows is the archive's finite-form version, read from [**Alp26a**, `Formalism/ThetaLiftIndependence/Statement.lean`].

(i) The lift sends a function f on class representatives L to $\sum_L w(L)\, f(L)\, \theta_L(\tau)$, a weighted sum of theta series of the lattices, with weights $w(L) = 1/n_L$ for positive integers n_L and a common denominator.

(ii) A chain composes m kernel operators $f \mapsto \big(j \mapsto \sum_i w_i f(i) K(i,j)\big)$ between finite class sets. If the scaled weights and the kernel entries are integral, the composite times the product of the denominators is integral.

(iii) There is an adjunction identity of seesaw type for such kernels.

(iv) If a linear map θ satisfies $\langle \theta x, \theta y\rangle = c \cdot \langle x, y\rangle$ on a subspace S stable under $\theta^*\theta$, with c a product of an L-value and local factors as in the Rallis inner product formula, then $\theta^*\theta = c$ on S. So the operator norm of θ on S is $\sqrt{|c|}$, and a norm identity for individual forms becomes an operator-norm bound.

How the kernels K of [**Alp26a**] are built from theta correspondences is not described here.

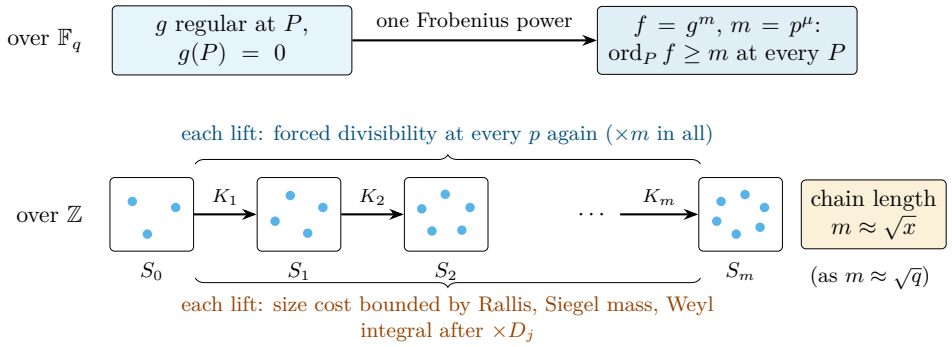
1.6.3. Why lifts substitute for derivatives. The following summarizes the note [**Alp26b**, pp. 3–4, steps 8–12] together with the statements above. The connecting argument is an interpretation; it is not proved here.

(i) *Integral and linear.* After clearing denominators, each kernel operator in the chain is an integral linear map of Hermitian lattices. So Proposition 1.4.2 applies to each lift and to their composite, as it applied to the jets of a theta function.

(ii) *Each lift collects the divisibility again.* Iterate lifts $U(X_0) \to U(X_1) \to \cdots \to U(X_m)$. According to the note, by seesaw the period of the iterated lift is a Bessel period independent of the Hecke characters, a period of the same CM kind as in Section 1.4. So the forced Legendre-type divisibility at each prime is collected once per lift, m times in all. This is the analogue of vanishing to order m.

(iii) *Each lift has a computable cost.* The Rallis formula, the Siegel mass formula (the total weight of a class set) and Weyl's dimension formula control the norms and ranks along the chain. These give the archimedean side of Proposition 1.4.2. Unlike m-fold differentiation, the cost does not couple to the size of the CM field.

(iv) *The length of the chain.* Balancing x/m against xm/k as in Section 1.5 gives $m \approx \sqrt{x}$, as $m \approx \sqrt{q}$ in Section 1.1.



schematic; the note's steps 8–12, with the interpretation of Section 1.6.3

FIGURE 1.11. Where multiplicity comes from (schematic). Over $\mathbb{F}_q$ one Frobenius power gives order m at every point. Over $\mathbb{Z}$, a chain of m integral kernel operators between finite class sets collects the forced divisibility once per lift, at a cost controlled by the Rallis formula, the Siegel mass and Weyl's dimension formula. Balancing gives $m \approx \sqrt{x}$. The picture follows the note [**Alp26b**] and the finite-form statements of [**Alp26a**], with the interpretation of Section 1.6.3; class-set sizes are illustrative.

Figure 1.11 draws the chain.

The analogy is summarized in the following table.

Curve over $\mathbb{F}_q$	Integers
$f = g^m$ with $m = p^\mu$	composite of m theta lifts
derivations of order $< m$ kill f	divisibility gained at each lift
$\deg f \leq mA + qB$	Rallis, Siegel mass and Weyl bounds
$m \approx \sqrt{q}$	chain length $m \approx \sqrt{x}$

Chapter 7 replaces this interpretation by definitions. The chain is the theta tower of Definition 7.1.2 (Figure 7.1). The periods are measured in the Hermitian lattices of Definitions 7.2.1 and 7.2.2. The conditions that the lattices along the chain must satisfy, the divisibility at inert primes and the archimedean costs among them, are collected in a slope record (Definitions 7.3.1 and 7.3.2). Remark 7.6.2 states the comparison with Stepanov's method in the terms of that chapter: the role of vanishing to high order is played by p-adic divisibility of lifted periods at all inert p at once.

1.7. What the outline omits, and where to read on

The outline above leaves out everything that makes the bounds uniform, and most of the book is about those costs. The note lists them [**Alp26b**, steps 11, 14–19]. Twisting over a 2-power cyclotomic tower kills unwanted contributions. A weakened form of Vandiver's conjecture bounds the classes

in $H^1(\mathcal{O}_F[1/\ell], \mathbb{Q}_\ell/\mathbb{Z}_\ell(-e))$, for certain CM fields F, that twisting cannot kill. The archive claims a proof of the latter, from Galois representations attached to automorphic forms. In the book these become the following chapters.

(1) *Hosts.* To isolate a given primitive character ψ one places it among the characters of the Galois group of a host field, and one needs such fields, with the right splitting behaviour, on windows. This is Chapter 5. The Fourier analysis that extracts a positive combination containing ψ is Chapter 6.
(2) *Congruence defects.* The lattices of the chain differ from the ideal ones by congruence defects, which are bounded through Selmer groups by Tate–Shafarevich groups of Tate twists. The bound is uniform only on windows (Theorem 10.5.1 in Chapter 10; Remark 10.4.10 discusses the relation to Vandiver's conjecture). This is the reason the criterion of Section 1.3 is needed on windows.
(3) *Galois representations.* The cases of the Tate–Shafarevich bound that need automorphy use Galois representations and a truncated local–global compatibility statement (Chapter 11).
(4) *Existence.* The families of arithmetic data to which the Chebyshev inequality applies have to be shown to exist (Chapter 12).

Chapter 3 states the main theorem, Theorem 3.1.1, and gives its whole proof with each of these tools as a black box. Figure 3.1 there displays the dependencies. Before that, Chapter 2 carries out the finite-field model completely, since everything in this chapter was phrased by analogy with it.

Remark 1.7.1 (Status). This chapter does not verify Alpöge's argument, and the Lean archive was read as text and not built. Whether it proves the Riemann Hypothesis is not established here. The readable statements cited in Sections 1.4–1.6 are not on the archive's proof path. They show what the authors had in mind, and the on-path argument may differ. Where the later chapters of this book differ from them, this chapter says so. The account of the error terms and of the role of the theta lifts is an interpretation.

CHAPTER 2

The function-field model: Stepanov–Bombieri and the sign-only proof

This chapter treats the two arguments for the Riemann hypothesis for curves over finite fields that Alpöge's note sets side by side [**Alp26b**]. The first is a complete proof. It is the method of Stepanov [**Ste69**] in the form given by Bombieri [**Bom74**]: an auxiliary function, forced to vanish to high order at the points to be counted, and a comparison of zeros with poles. It uses the Riemann–Roch theorem and nothing deeper, and it is given here in full, from the definition of the zeta function to the Riemann hypothesis (Sections 2.1–2.5). The second is the "sign-only" proof, which the note records in one footnote as the identity $qN = M^k + R$. It uses no auxiliary function and no multiplicity; it uses elementary Fourier analysis and the parity of a cohomological degree. It is reconstructed here as a complete deduction from two cohomological hypotheses, which are stated and not proved (Sections 2.8–2.10). It is therefore a conditional derivation, and not a second proof.

Nothing in this chapter is a logical input to the later chapters. Its output is the analogy on which they lean. Chapter 1 was phrased by analogy with the first proof, and Chapter 3 assembles an argument whose parts were found by transporting the two arguments to the integers. The dictionary of Section 2.13 matches each step proved here with the part of Chapters 3–12 that plays its role, and says where the match is exact, where it is loose, and where it fails.

Four results of the chapter are negative.

(1) The upper bound for point counts does not by itself give the Riemann hypothesis through a positivity argument of Landau–Pringsheim type. The missing input is identified exactly, and it cannot be supplied formally (Section 2.6).
(2) The one-sided bound on a smoothed prime sum that the number-field argument proves corresponds, over a function field, to a *lower* bound on point counts. It does not correspond to the Stepanov–Bombieri upper bound (Proposition 2.6.10).
(3) The method forces the square root in the Stepanov–Bombieri error. A different choice of parameters does not remove it (Proposition 2.3.10).

(4) The second hypothesis of the sign-only proof fails for the quadric as the note writes it, for some values of k in the simplest cases (Remark 2.10.7). So the sign-only proof is conditional, and the first proof is not.

2.1. Setting

Let p be a prime and $q = p^e$. Sections 2.1 to 2.7 are written in the language of function fields and valuations, in which the proof needs no geometry beyond the Riemann–Roch theorem; the translation to curves follows the definition.

Definition 2.1.1.

(1) A *function field* over $\mathbb{F}_q$ is a finitely generated extension $K/\mathbb{F}_q$ of transcendence degree one in which $\mathbb{F}_q$ is algebraically closed. We also say that K has *full constant field* $\mathbb{F}_q$.
(2) A *place* of K is the maximal ideal P of a discrete valuation ring $\mathcal{O}_P$ with $\mathbb{F}_q \subset \mathcal{O}_P \subset K$ and fraction field K; v_P is its valuation, $\kappa(P) = \mathcal{O}_P/P$, $\deg P = [\kappa(P) : \mathbb{F}_q]$, and $z(P)$ is the class of $z \in \mathcal{O}_P$ in $\kappa(P)$.
(3) A *divisor* is an element of the free abelian group on the places; (z) is the divisor of $z \in K^\times$, and $(z)_0$, $(z)_\infty$ are its divisors of zeros and of poles. For a divisor D, $\mathcal{L}(D) = \{z \in K^\times : (z) + D \geq 0\} \cup \{0\}$ and $\ell(D) = \dim_{\mathbb{F}_q} \mathcal{L}(D)$.
(4) B_d is the number of places of degree d, and $N_n = \sum_{d|n} d B_d$.

Geometrically $K = \mathbb{F}_q(C)$ for a smooth projective geometrically irreducible curve $C/\mathbb{F}_q$ of genus g, the places are the closed points of C, $\mathcal{L}(D) = H^0(C, \mathcal{O}_C(D))$ is the space of sections of a line bundle, and $N_n = \#C(\mathbb{F}_{q^n})$; see [**Vak25**, Theorem 16.3.3] for the equivalence between curves and function fields and [**Sti09**, App. B] for the comparison of the two languages. Figure 2.1 illustrates the definition. The notation $\mathcal{L}(D)$, $\ell(D)$ is that of Chapter 1.

The theorem to be proved is the following.

Theorem 2.1.2 (Weil). *Let $K/\mathbb{F}_q$ be a function field of genus g. There are complex numbers $\alpha_1, \dots, \alpha_{2g}$ such that*

(1) $N_n = q^n + 1 - \sum_{i=1}^{2g} \alpha_i^n$ for all $n \geq 1$;
(2) $\alpha \mapsto q/\alpha$ permutes the multiset $\{\alpha_i\}$;
(3) $|\alpha_i| = \sqrt{q}$ for all i.

In particular $|N_n - q^n - 1| \leq 2g\, q^{n/2}$ for all $n \geq 1$.

Parts (1) and (2) are Theorem 2.2.6; part (3) and the bound are Theorem 2.5.3. Figure 2.2 is a map of the proof and of the two routes from the upper bound to the Riemann hypothesis, only one of which arrives.

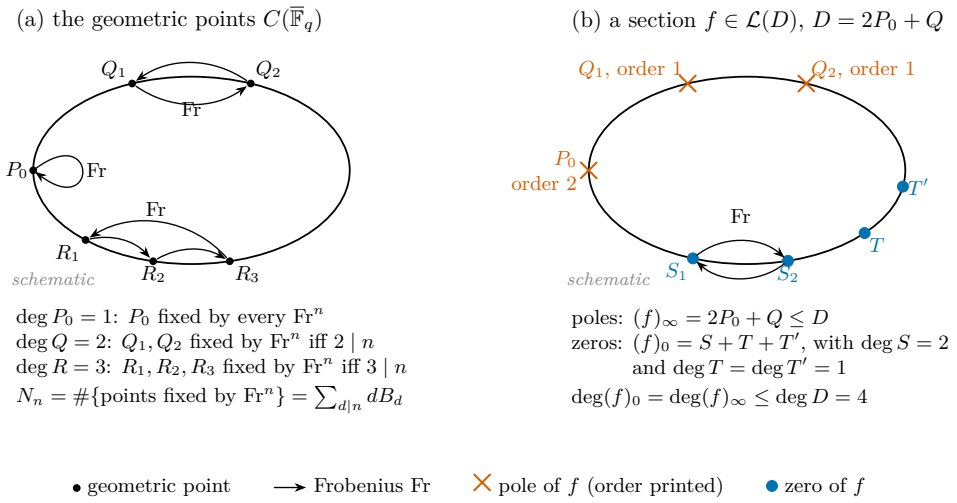


FIGURE 2.1. Closed points, a divisor and a section on a curve C over $\mathbb{F}_q$ (Definition 2.1.1). (a) The geometric points $C(\overline{\mathbb{F}}_q)$, drawn as a loop; Frobenius Fr (black arrows) fixes P_0, swaps Q_1, Q_2 and cycles R_1, R_2, R_3, so P_0, Q, R are places of degree 1, 2, 3. A place of degree d contributes its d geometric points to $N_n = \#C(\mathbb{F}_{q^n})$ exactly when $d \mid n$; this is $N_n = \sum_{d|n} dB_d$. (b) A nonzero $f \in \mathcal{L}(D)$ for $D = 2P_0 + Q$: its poles (vermillion crosses, orders printed) are bounded by D, and its zeros (blue dots; here a place S of degree 2 and two places T, T' of degree 1) have total degree $\deg(f)_0 = \deg(f)_\infty \le \deg D = 4$ (fact S2). Schematic: the curve is drawn with one real dimension and no genus, and the positions are arbitrary.

2.1.1. Standard facts. The proof is complete relative to the following facts. Those that are algebraic geometry of curves are in [**Vak25**], with locators in the numbering of its public version of October 21, 2025; the locators into [**Sti09**], where the same facts are proved in the language of valuations, are at the level of sections.

(S1) Weak approximation: for distinct places $P_1, \dots, P_s$, elements $z_1, \dots, z_s \in K$ and an integer n there is $z \in K$ with $v_{P_j}(z - z_j) \ge n$ for all j [**Sti09**, §1.3].

(S2) $\deg(z)_0 = \deg(z)_\infty = [K : \mathbb{F}_q(z)]$ for $z \notin \mathbb{F}_q$ [**Vak25**, Exercises 16.3.E and 18.4.E], [**Sti09**, §1.4]; so a principal divisor has degree 0, an effective divisor has nonnegative degree, a function without zeros or poles is constant, and $\mathcal{L}(D) = 0$ if $\deg D < 0$.

(S3) Riemann–Roch: $\ell(D) = \deg D + 1 - g + \ell(W - D)$ for a canonical divisor W, with $\deg W = 2g - 2$; so $\ell(D) \ge \deg D + 1 - g$, with equality if $\deg D \ge 2g - 1$ [**Vak25**, Exercise 18.4.B, §18.5.3, Exercise 18.5.A], [**Sti09**, §1.5].

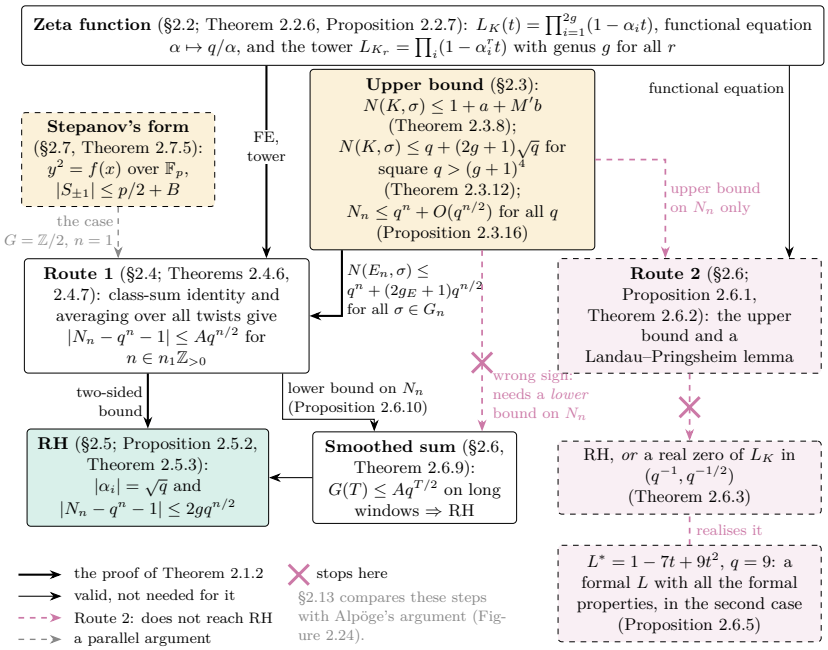


FIGURE 2.2. The structure of the proof of Theorem 2.1.2. Thick arrows: the proof given (Sections 2.2, 2.3, 2.4, 2.5); thin arrows: valid implications the proof does not need. Dashed purple: Route 2 (Section 2.6), which from the upper bound alone reaches only "the Riemann hypothesis or a real zero of L_K in $(q^{-1}, q^{-1/2})$" (Theorem 2.6.3); Proposition 2.6.5 realises the second case for a formal L-polynomial. The smoothed-sum criterion (Theorem 2.6.9) is valid, but its hypothesis is a lower bound on point counts (Proposition 2.6.10), which Route 1 provides and Section 2.3 does not. Dashed grey: Section 2.7 bounds $|S_{\pm 1}|$ from above for $y^2 = f(x)$ over $\mathbb{F}_p$, which gives a two-sided bound as Route 1 does for $G = \mathbb{Z}/2$. Box fills: orange, the bounds of the Stepanov–Bombieri method; green, the Riemann hypothesis; purple, the failure of Route 2 and the counterexample. Arrow labels say what each step passes on. Schematic.

(S4) For $r \ge 1$ the compositum $K_r = K\mathbb{F}_{q^r}$ is a function field with full constant field $\mathbb{F}_{q^r}$, and over a place of K of degree m lie $\gcd(m,r)$ places of K_r, each of degree $m/\gcd(m,r)$ over $\mathbb{F}_{q^r}$; so K_r has N_r places of degree one [**Sti09**, §3.6, §5.1].

(S5) $\mathbb{F}_q(x)$ has $q+1$ places of degree one and finitely many of each degree [**Sti09**, §1.2].

(S6) $K/\mathbb{F}_q(x)$ is finite and separable for some $x \in K$ [**Vak25**, Theorem 10.5.17], [**Sti09**, §3.10].

(S7) In a finite extension E/F of function fields each place of F has between 1 and $[E : F]$ places above it, and $\mathrm{Gal}(E/F)$ permutes them transitively if E/F is Galois [**Sti09**, §3.1, §3.7].
(S8) $\mathrm{Gal}(\mathbb{F}_{q^f}/\mathbb{F}_q)$ is cyclic of order f, generated by $\mathrm{Fr}_q\colon c \mapsto c^q$; the group $\mathbb{F}_{q^f}^\times$ is cyclic; and $\mathbb{F}_q = \{c : c^q = c\}$.
(S9) If E/F is a finite Galois extension of fields and $F' \supseteq F$, then EF'/F' is Galois and restriction embeds $\mathrm{Gal}(EF'/F')$ in $\mathrm{Gal}(E/F)$.

Facts S2, S3, S5, S6 and the first half of S7 are statements about curves, line bundles and finite morphisms of the kind assumed throughout this book. Weak approximation, the count of places in a constant field extension, the transitivity in S7, and S8, S9 are commutative algebra and field theory, which is why the references to [**Sti09**] are kept.

One more fact about valuation rings is used, and it has a three-line proof. If $\mathcal{O} \subseteq \mathcal{O}' \subsetneq K$ are valuation rings of K containing $\mathbb{F}_q$ and $\mathcal{O}$ is discrete, then $\mathcal{O} = \mathcal{O}'$. Indeed, let v be the valuation of $\mathcal{O}$ and t a uniformizer. If some $z \in \mathcal{O}'$ had $v(z) = -m < 0$, then zt^m is a unit of $\mathcal{O}$, so $t^{-m} = z(zt^m)^{-1} \in \mathcal{O}'$; hence $t^{-1} = t^{m-1}t^{-m}$ lies in $\mathcal{O}'$, and then $\mathcal{O}' \supseteq \mathcal{O}[t^{-1}] = K$.

2.1.2. Frobenius and automorphisms. Characteristic p enters the whole proof through the next proposition.

PROPOSITION 2.1.3. *Let $K/\mathbb{F}_q$ be a function field, $j \geq 0$ an integer and $\pi = p^j$.*

(1) $z \mapsto z^\pi$ is an injective ring endomorphism of K.
(2) $v_P(z^\pi) = \pi\, v_P(z)$ for every place P and $z \in K^\times$.
(3) If $\deg P = 1$ and $z \in \mathcal{O}_P$, then $z^q - z \in P$.

PROOF. Proof of (1): the map is multiplicative, and it is additive because $\binom{p}{i} \equiv 0 \pmod p$ for $0 < i < p$; a ring endomorphism of a field is injective. Proof of (2): valuations are multiplicative. Proof of (3): $z(P) \in \kappa(P) = \mathbb{F}_q$, and $c^q = c$ for $c \in \mathbb{F}_q$ by S8. □

Let $\mathrm{Aut}(K/\mathbb{F}_q)$ be the group of field automorphisms of K that fix $\mathbb{F}_q$. For $\sigma \in \mathrm{Aut}(K/\mathbb{F}_q)$ and a place P, the image $\sigma(P)$ is a place, with $\mathcal{O}_{\sigma(P)} = \sigma(\mathcal{O}_P)$, $v_{\sigma(P)}(z) = v_P(\sigma^{-1}z)$ and $\deg \sigma(P) = \deg P$; and $\sigma(\mathcal{L}(D)) = \mathcal{L}(\sigma D)$ for every divisor D. All of this is transport of structure along σ.

2.2. The zeta function

This section proves parts (1) and (2) of Theorem 2.1.2. The argument is F. K. Schmidt's [**Sch31**]; see also [**Sti09**, §5.1]. All identities are between formal power series or rational functions of t.

PROPOSITION 2.2.1. *Let $K/\mathbb{F}_q$ be a function field of genus g. For each $n \geq 0$ the number A_n of effective divisors of degree n is finite, and the number h of divisor classes of degree 0 is finite.*

PROOF. Fix $x \in K \setminus \mathbb{F}_q$. A place P of K lies over a place of $\mathbb{F}_q(x)$ of degree at most $\deg P$, so $B_d < \infty$ by S5 and S7, and hence $A_n < \infty$. Let D_0 be a divisor of degree at least g. For D of degree 0, S3 gives $\ell(D + D_0) \geq 1$, so $D + D_0$ is equivalent to an effective divisor of degree $\deg D_0$. Distinct classes give distinct effective divisors, so $h \leq A_{\deg D_0}$. □

DEFINITION 2.2.2. Let $K/\mathbb{F}_q$ be a function field. Its *zeta function* is

$$Z(t) = Z_K(t) = \sum_{n \geq 0} A_n t^n \in \mathbb{Z}[[t]],$$

where A_n is the number of effective divisors of K of degree n; and $\delta \geq 1$ denotes the positive generator of the image of the degree map on divisors.

Unique factorisation of effective divisors into places gives the Euler product $Z(t) = \prod_P (1 - t^{\deg P})^{-1}$, and applying $t\frac{d}{dt}\log$,

$$t\,\frac{Z'(t)}{Z(t)} = \sum_P \deg P \sum_{j \geq 1} t^{j \deg P} = \sum_{n \geq 1} N_n t^n. \tag{2.2.1}$$

So the point counts N_n are the coefficients of the logarithmic derivative of Z, exactly as $\Lambda(n)$ gives the coefficients of $-\zeta'/\zeta$.

PROPOSITION 2.2.3. *Let $K/\mathbb{F}_q$ be a function field and c a divisor class. Then c contains exactly $(q^{\ell(c)} - 1)/(q - 1)$ effective divisors.*

PROOF. Fix $D \in c$. The map $f \mapsto D + (f)$ from $\mathcal{L}(D) \setminus \{0\}$ to the effective divisors equivalent to D is surjective, and $(f) = (f')$ exactly when $f/f' \in \mathbb{F}_q^\times$, by S2. □

PROPOSITION 2.2.4. *Let $K/\mathbb{F}_q$ be a function field of genus g with h divisor classes of degree 0, let δ be the positive generator of the image of the degree map on divisors, let n_0 be the least multiple of δ with $n_0 \geq \max(0, 2g-1)$, and let $F(t) = \sum_{0 \leq \deg c \leq 2g-2} q^{\ell(c)} t^{\deg c}$, the sum running over divisor classes c of K. Then*

$$(q-1)Z(t) = F(t) + h\,q^{1-g}\frac{(qt)^{n_0}}{1-(qt)^\delta} - \frac{h}{1-t^\delta}. \tag{2.2.2}$$

In particular Z is a rational function of t^δ with a simple pole at $t = 1$.

PROOF. There are h classes of each degree in $\delta\mathbb{Z}$ and none of other degrees. By Proposition 2.2.3, $(q-1)Z(t) = \sum_{\deg c \geq 0}(q^{\ell(c)} - 1)t^{\deg c}$. For $\deg c \geq 2g - 1$, S3 gives $\ell(c) = \deg c + 1 - g$. Split the sum at degree $2g - 2$ and sum the two geometric series that remain: the terms $q^{\ell(c)}t^{\deg c}$ with $\deg c \geq n_0$ give the middle term of (2.2.2), and the terms $-t^{\deg c}$ over all classes of degree at least 0 give the last. The polynomial F is a polynomial in t^δ; the middle term is regular at $t = 1$ because $q^\delta \neq 1$; the last term has a simple pole there. □

PROPOSITION 2.2.5 (F. K. Schmidt). *Let $K/\mathbb{F}_q$ be a function field. Then $\delta = 1$: there is a divisor of degree one.*

PROOF. The plan is to compare the order of the pole at $t = 1$ of both sides of a product formula. Let $r \geq 1$ and let Z_r be the zeta function of $K_r/\mathbb{F}_{q^r}$ (fact S4). For a place P of K of degree m put $d = \gcd(m, r)$. By S4 the factors of $Z_r(t^r)^{-1}$ at the places above P multiply to $(1 - t^{rm/d})^d$. Also $\prod_{\zeta^r=1}(1 - (\zeta t)^m) = (1 - t^{mr/d})^d$, because ζ^m runs d times over the (r/d)-th roots of unity. Hence

$$Z_r(t^r) = \prod_{\zeta^r=1} Z(\zeta t). \tag{2.2.3}$$

Take $r = \delta$. As Z is a function of t^δ, the right side is $Z(t)^\delta$, which has a pole of order δ at $t = 1$. By Proposition 2.2.4 for K_δ, $Z_\delta(u)$ has a simple pole at $u = 1$, so $Z_\delta(t^\delta)$ has a simple pole at $t = 1$. Hence $\delta = 1$. □

THEOREM 2.2.6 (F. K. Schmidt: rationality and functional equation). *Let $K/\mathbb{F}_q$ be a function field of genus g with h divisor classes of degree 0.*

(1) $Z(t) = L(t)/((1 - t)(1 - qt))$ with $L = L_K \in \mathbb{Z}[t]$ of degree $2g$, $L(0) = 1$, leading coefficient q^g and $L(1) = h$.
(2) $Z(1/(qt)) = q^{1-g}t^{2-2g}Z(t)$, equivalently $L(t) = q^g t^{2g} L(1/(qt))$.
(3) The map $\alpha \mapsto q/\alpha$ permutes the multiset $\{\alpha_1, \ldots, \alpha_{2g}\}$ of reciprocals of the roots of L, no α_i equals 1 or q, and $N_n = q^n + 1 - \sum_i \alpha_i^n$ for all $n \geq 1$.

The numbers α_i, for which $L(t) = \prod_{i=1}^{2g}(1 - \alpha_i t)$, are the *reciprocal roots* of K, and L_K is its *L-polynomial*.

PROOF. Since $\delta = 1$ (Proposition 2.2.5), Proposition 2.2.4 gives (2.2.2) with $n_0 = \max(0, 2g - 1)$ and $F(t) = \sum_{0 \leq \deg c \leq 2g-2} q^{\ell(c)} t^{\deg c}$. If $g = 0$, then $F = 0$, $n_0 = 0$, and $(q - 1)Z = \frac{hq}{1-qt} - \frac{h}{1-t} = \frac{h(q-1)}{(1-t)(1-qt)}$; since $Z(0) = 1$, $h = 1$ and $L = 1$.

Let $g \geq 1$, so $n_0 = 2g - 1$. Write $(q - 1)Z = F + R$ with $R(t) = hq^{1-g}\frac{(qt)^{2g-1}}{1-qt} - \frac{h}{1-t}$.

Proof of (1). We have

$$(q-1)(1-t)(1-qt)Z(t) = (1-t)(1-qt)F(t) + hq^g t^{2g-1}(1-t) - h(1-qt),$$

a polynomial of degree at most $2g$ because $\deg F \leq 2g - 2$. So $L := (1 - t)(1 - qt)Z$ is a polynomial of degree at most $2g$ lying in $\mathbb{Z}[[t]]$, hence in $\mathbb{Z}[t]$, and $L(0) = Z(0) = 1$. At $t = 1$ the identity gives $(q - 1)L(1) = h(q - 1)$.

Proof of (2). The map $c \mapsto W - c$ is an involution of the classes of degree in $[0, 2g - 2]$, and $\ell(W - c) = \ell(c) - \deg c - 1 + g$ by S3. Hence

$$\begin{aligned} F\Big(\frac{1}{qt}\Big) &= \sum_c q^{\ell(W-c)}(qt)^{-(2g-2-\deg c)} \\ &= \sum_c q^{\ell(c)-\deg c-1+g}(qt)^{\deg c-2g+2} = q^{1-g}t^{2-2g}F(t). \end{aligned}$$

For R a direct computation gives

$$R\Big(\frac{1}{qt}\Big) = -\frac{hq^{1-g}t^{2-2g}}{1-t} + \frac{hqt}{1-qt} = q^{1-g}t^{2-2g}R(t).$$

This proves (2) for Z. For L, $L(1/(qt)) = \frac{(qt-1)(t-1)}{qt^2}\,q^{1-g}t^{2-2g}Z(t) = q^{-g}t^{-2g}L(t)$.

Proof of (3). With $L = \sum_j a_j t^j$ the functional equation reads $a_{2g-j} = q^{g-j}a_j$. As $a_0 = 1$, we get $a_{2g} = q^g$; so $\deg L = 2g$ and $\prod_i \alpha_i = q^g$. Then

$$L(t) = q^g t^{2g}\prod_i\Big(1-\frac{\alpha_i}{qt}\Big) = q^g\prod_i\Big(-\frac{\alpha_i}{q}\Big)\prod_i\Big(1-\frac{q}{\alpha_i}t\Big) = \prod_i\Big(1-\frac{q}{\alpha_i}t\Big),$$

so $\{q/\alpha_i\} = \{\alpha_i\}$. Since $L(1) = h \geq 1$, no α_i equals 1, and hence none equals q. Finally (2.2.1) and (1) give $\sum_n N_n t^n = \frac{t}{1-t} + \frac{qt}{1-qt} - \sum_i \frac{\alpha_i t}{1-\alpha_i t}$; compare coefficients. □

Figure 2.3 shows the zeros and poles of Z and the symmetry of the functional equation.

2.2.1. The tower of constant field extensions. The tower $K_r = K\mathbb{F}_{q^r}$ is used below to pass from one point count to all of them. Fact S4 describes how places split in K_r, but it says nothing about the genus, and the Riemann–Roch theorem for K_r involves the genus of K_r, not that of K. Geometrically the genus is the dimension of $H^1(C, \mathcal{O}_C)$, which does not change under extension of the base field [**Vak25**, Exercise 18.2.H]. That argument uses cohomology, which this proof otherwise avoids. The next proposition derives the invariance from the zeta function instead.

PROPOSITION 2.2.7. *Let $K/\mathbb{F}_q$ be a function field of genus g with reciprocal roots $\alpha_1, \ldots, \alpha_{2g}$, and let $r \geq 1$. Then $L_{K_r}(t) = \prod_i(1-\alpha_i^r t)$. In particular K_r has genus g.*

PROOF. Let g_r be the genus of K_r. By (2.2.3) and Theorem 2.2.6(1) for K and K_r, $L_{K_r}(t^r) = \prod_{\zeta^r=1} L_K(\zeta t) = \prod_i(1-\alpha_i^r t^r)$, using $\prod_\zeta(1-\zeta t) = 1-t^r$ and $\prod_\zeta(1-q\zeta t) = 1-q^r t^r$. Comparing degrees in t gives $2g_r r = 2gr$. □

REMARK 2.2.8. The invariance of the genus under constant field extension need not be assumed alongside fact S4. Proposition 2.2.7 derives it from the zeta function. The degree of L_{K_r} is twice the genus of K_r by Theorem 2.2.6 applied to K_r, and the product formula (2.2.3) fixes that degree as $2g$. The proof uses only S4 and Riemann–Roch for each field separately.

2.3. The upper bound

This section proves the Stepanov–Bombieri upper bound for point counts. The proof has four named steps, shown in Figure 2.4: a Siegel lemma, a non-vanishing statement, forced multiplicity, and a slope inequality. The names are chosen so that the steps can be compared with the number-field argument in Section 2.13. The bound is proved in a form

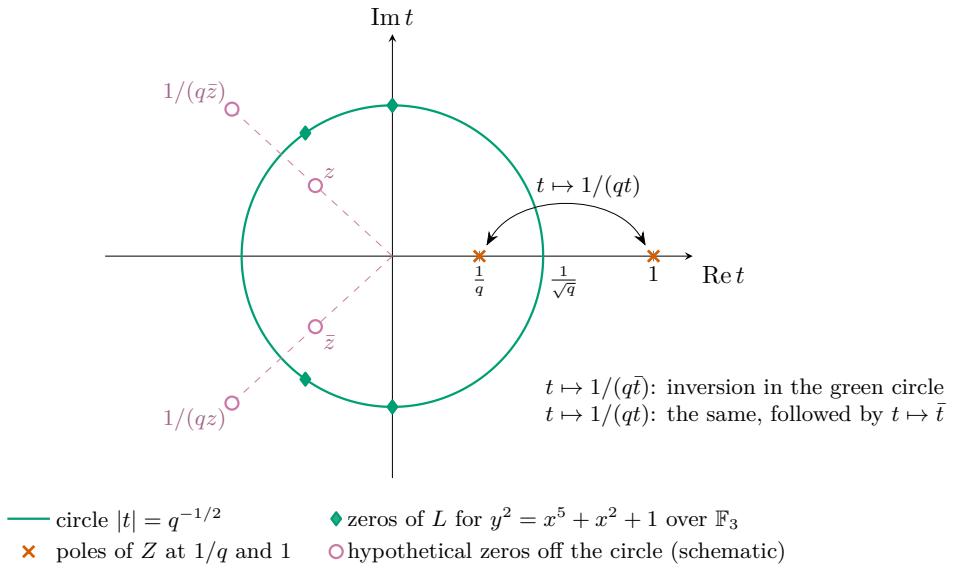


FIGURE 2.3. The t-plane for $q = 3$ and the functional equation of Theorem 2.2.6(2). The map $t \mapsto 1/(qt)$ is inversion in the circle $|t| = q^{-1/2}$ (green), $t \mapsto 1/(q\bar{t})$, followed by complex conjugation. The zeros of L_K are closed under conjugation, because L_K has real coefficients, so they are also closed under inversion in that circle. The map swaps the poles of Z at $1/q$ and 1 (vermillion crosses). Green diamonds: the four zeros of $L = 1 + 2t + 6t^2 + 6t^3 + 9t^4$ for the genus-2 curve $y^2 = x^5 + x^2 + 1$ over $\mathbb{F}_3$, computed numerically from its point counts. They lie on the circle, as Theorem 2.5.3 proves in general. Purple open circles: a hypothetical set of zeros z, $\bar{z}$, $1/(q\bar{z})$, $1/(qz)$ off the circle, which Theorem 2.2.6 allows and Theorem 2.5.3 excludes. Their position is schematic.

twisted by an automorphism σ, which costs nothing and is what Section 2.4 needs.

DEFINITION 2.3.1. Let $K/\mathbb{F}_q$ be a function field and $\sigma \in \mathrm{Aut}(K/\mathbb{F}_q)$.

(1) A place P is *σ-Frobenius* if $\sigma(z) - z^q \in P$ for all $z \in \mathcal{O}_P$.
(2) $N(K, \sigma) = \sum \deg P$, the sum over the σ-Frobenius places P.

By Proposition 2.1.3(3) and S8, the id-Frobenius places are those of degree one, so $N(K, \mathrm{id}) = N_1$. A σ-Frobenius place has $\sigma(\mathcal{O}_P) \subseteq \mathcal{O}_P$, hence $\sigma(\mathcal{O}_P) = \mathcal{O}_P$ by the maximality of discrete valuation rings proved in Section 2.1, and therefore $\sigma(P) = P$. Geometrically, the σ-Frobenius places are the points of C at which the q-Frobenius agrees with the automorphism induced by σ: the rational points of a twist of C.

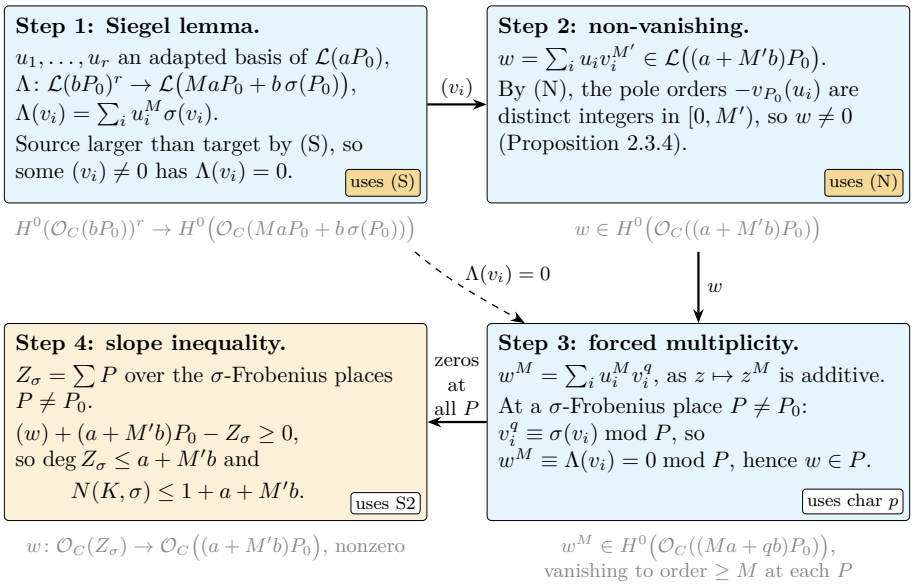


FIGURE 2.4. The four steps of the proof of Theorem 2.3.8, for a function field $K/\mathbb{F}_q$, a place P_0 of degree one, $\sigma \in \mathrm{Aut}(K/k)$ and $q = MM'$ with M a power of p. Step 1 uses the Siegel condition (S) to find a nonzero kernel element (v_i) of Λ; Step 2 uses the non-vanishing condition (N) and Proposition 2.3.4 to see that $w \neq 0$; Step 3 uses that $z \mapsto z^M$ is additive, the only use of the characteristic (Proposition 2.1.3), and the relation $\Lambda(v_i) = 0$ of Step 1 (dashed arrow), to make w vanish at every σ-Frobenius place $P \neq P_0$; Step 4 compares degrees (fact S2). The tag in each box names the hypothesis that step uses. Grey lines give the line bundles on C in which each step takes place. Sky blue: auxiliary data; orange: the bound produced and the conditions (S), (N). Schematic.

2.3.1. An adapted basis and independence. In the chain $\mathcal{L}(jP_0)$, $0 \leq j \leq a$, each step adds at most one dimension, so the pole orders at P_0 of the elements of $\mathcal{L}(aP_0)$ form a set of exactly $\ell(aP_0)$ integers. Some basis has pairwise distinct pole orders. Not every basis does (Remark 2.3.3).

PROPOSITION 2.3.2. *Let $K/\mathbb{F}_q$ be a function field, P_0 a place of degree one and $a \geq 0$. Then $\mathcal{L}(aP_0)$ has a basis $u_1, \dots, u_r$ with $0 \leq -v_{P_0}(u_1) < \cdots < -v_{P_0}(u_r) \leq a$.*

PROOF. For a uniformizer t at P_0 and $0 \leq j \leq a$, the map $f \mapsto (t^j f)(P_0)$ sends $\mathcal{L}(jP_0)$ to $\kappa(P_0) = \mathbb{F}_q$, with kernel $\mathcal{L}((j-1)P_0)$. Choose one element of pole order j for each j for which the quotient is nonzero. These elements are independent, because their pole orders are distinct, and their number is $\ell(aP_0) - \ell(-P_0) = \ell(aP_0)$ by S2. □

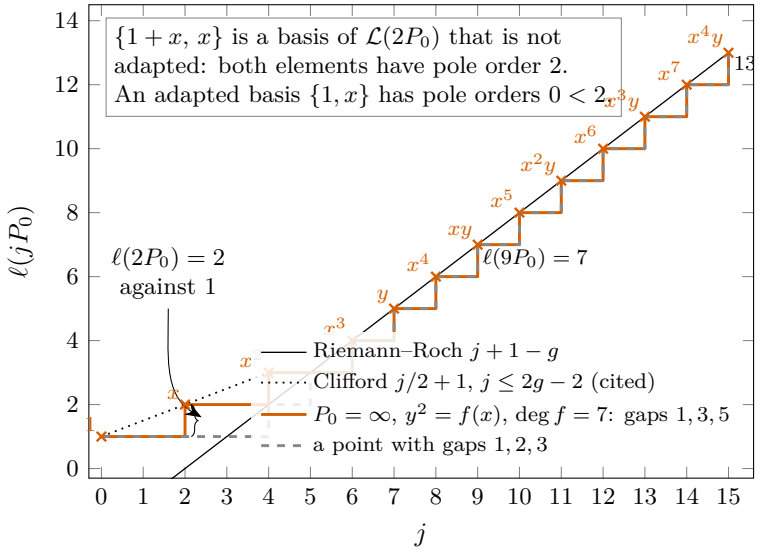


FIGURE 2.5. The staircase $j \mapsto \ell(jP_0)$, $0 \le j \le 15$, for $g = 3$ (Proposition 2.3.2). Vermillion: P_0 the place at infinity of $y^2 = f(x)$ with $\deg f = 7$, whose gaps are $1, 3, 5$; each jump ($\times$) is the pole order of the labelled element of an adapted basis (x has a pole of order 2 and y one of order 7). Grey dashed: a point with gaps $1, 2, 3$, where $\ell(jP_0) = \max(1, j + 1 - g)$; the two staircases agree for $j \ge 2g - 1 = 5$. Black: the Riemann–Roch value $j + 1 - g$, a lower bound that is exact for $j \ge 2g - 1$ (fact S3). Dotted: Clifford's bound $j/2 + 1$ for $j \le 2g - 2$, not proved here; the vermillion staircase meets it at $j = 0, 2, 4$. The box gives the non-adapted basis of Remark 2.3.3. At $j = 2$ the staircases differ; this is why $(a, b) = (2, 9)$ satisfies (S) at infinity, $2 \cdot 7 > 13$, although $q_0 = 3 \le (g + 1)^2$ (Remark 2.3.14).

We call such a basis *adapted*. Figure 2.5 shows the filtration by pole order.

REMARK 2.3.3. If $\ell(aP_0) \ge 2$, then some basis of $\mathcal{L}(aP_0)$ has two elements with the same pole order at P_0. If $-v_{P_0}(u_1) < -v_{P_0}(u_2)$, then $u_1 + u_2$ and u_2 span the same space and have the same pole order. So the argument below needs an adapted basis, though its conclusion does not depend on the basis (Remark 2.3.5).

PROPOSITION 2.3.4. *Let v be a discrete valuation on a field K, let $M' \ge 1$ be an integer, and let $u_1, \dots, u_r \in K^\times$ have values $v(u_i)$ pairwise incongruent modulo M'. If $v_1, \dots, v_r \in K$ are not all zero, then $\sum_i u_i v_i^{M'} \ne 0$.*

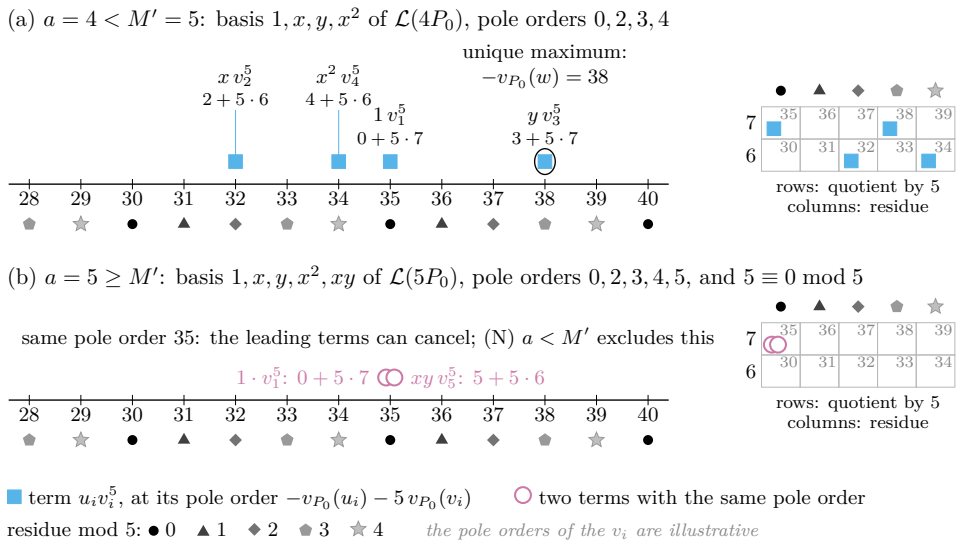


FIGURE 2.6. Why the section $w = \sum_i u_i v_i^{M'}$ is nonzero (Proposition 2.3.4, Step 2 of Theorem 2.3.8), for $q = 25$, $g = 1$, $M' = 5$. Pole orders $-v_{P_0}$ are drawn on a number line, with their residues modulo 5 shown by shape and grey shade; at the right the same terms are placed in a grid by quotient and residue modulo 5, as in Figure 2.19. (a) For $a = 4$, the adapted basis $1, x, y, x^2$ of $\mathcal{L}(4P_0)$ has pole orders $0, 2, 3, 4$. With pole orders $7, 6, 7, 6$ of $v_1, \ldots, v_4$, the terms $u_i v_i^5$ (sky-blue squares) have pole orders $35, 32, 38, 34$, one in each of four residue classes; the largest is attained once, so it is the pole order of w, and $w \neq 0$. A section w computed for one curve with these parameters has pole order 38. (b) If $a \geq M'$, here $a = 5$, two basis elements have pole orders congruent modulo M', so two terms can have the same pole order (purple open circles, one grid cell) and their leading terms can cancel; condition (N) excludes this. The pole orders of the v_i are illustrative, not those of the computed section: for $a < M'$ any choice gives distinct residues.

PROOF. For $v_i \neq 0$, $v(u_i v_i^{M'}) = v(u_i) + M' v(v_i) \equiv v(u_i) \pmod{M'}$. So the nonzero terms have pairwise distinct values, and the value of the sum is the least of them. □

Figure 2.6 pictures the proof. This is the only non-vanishing argument in the proof of the upper bound, and it is the step for which Section 2.13 finds no counterpart in the number-field argument.

REMARK 2.3.5. Let $K/\mathbb{F}_q$ be a function field, P_0 a place of degree one, $a \geq 0$ an integer, M' a power of p, and let $(u_i)_{i \leq r}$ and $(u'_j)_{j \leq r}$ be two bases

of $\mathcal{L}(aP_0)$, with $u_i = \sum_j c_{ij} u'_j$ and $c_{ij} \in \mathbb{F}_q$. Since $c \mapsto c^{M'}$ is a bijection of $\mathbb{F}_q$, for all $v_1, \dots, v_r \in K$

$$\sum_i u_i v_i^{M'} = \sum_j u'_j \Big(\sum_i c_{ij}^{1/M'} v_i \Big)^{M'},$$

and the matrix $(c_{ij}^{1/M'})$ is invertible. So the statement that $\sum_i u_i v_i^{M'} = 0$ forces all $v_i = 0$ holds for one basis if and only if it holds for all.

Remark 2.3.6. Bombieri proves the same injectivity with the roles of the two spaces exchanged [**Bom74**, pp. 236–237]. Let $K/\mathbb{F}_q$ be a function field, P_0 a place of degree one, $q = MM'$ with M a power of p, and $a, b \geq 0$ integers. Bombieri takes a basis $s_1, \dots, s_r$ of $\mathcal{L}(bP_0)$ with $v_{P_0}(s_1) < \cdots < v_{P_0}(s_r)$ and coefficients σ_i in $\mathcal{L}(aP_0)$. If $\sum_{i \geq \rho} \sigma_i^M s_i^q = 0$ with $\sigma_\rho \neq 0$, then, writing $v = v_{P_0}$,

$$Mv(\sigma_\rho) + qv(s_\rho) \geq \min_{i>\rho} \big(Mv(\sigma_i) + qv(s_i) \big) \geq -Ma + qv(s_{\rho+1}),$$

so $Mv(\sigma_\rho) \geq q - Ma > 0$ when $a < q/M$. Then σ_ρ has a zero at P_0 and no poles, so $\sigma_\rho = 0$ by S2, a contradiction. The hypothesis $a < q/M$ is condition (N) below.

2.3.2. The auxiliary section.

Definition 2.3.7. Let $K/\mathbb{F}_q$ be a function field, P_0 a place of degree one, $\sigma \in \mathrm{Aut}(K/\mathbb{F}_q)$, and $q = MM'$ with $M = p^\mu$ a power of p. A pair of integers $a, b \geq 0$ is *admissible* if it satisfies the two conditions

(N) $a < M'$;
(S) $\ell(aP_0)\,\ell(bP_0) > \ell\big(MaP_0 + b\,\sigma(P_0)\big)$.

Condition (N) will give non-vanishing and condition (S) is the hypothesis of a Siegel lemma, whence the letters. The integer M is the multiplicity; it corresponds to m in Section 1.1, and a, b to A, B there.

Theorem 2.3.8 (Stepanov–Bombieri inequality, twisted). *Let $K/\mathbb{F}_q$ be a function field, P_0 a place of degree one, $\sigma \in \mathrm{Aut}(K/\mathbb{F}_q)$, $q = MM'$ with $M = p^\mu$, and (a, b) an admissible pair. Then there is a nonzero $w \in \mathcal{L}((a + M'b)P_0)$ vanishing at every σ-Frobenius place $P \neq P_0$. Consequently*

$$N(K, \sigma) \leq 1 + a + M'b. \tag{2.3.1}$$

Geometrically, if $K = \mathbb{F}_q(C)$, then w is a nonzero section of $\mathcal{L}((a + M'b)P_0) = H^0(C, \mathcal{O}_C((a + M'b)P_0))$.

Proof. Fix an adapted basis $u_1, \dots, u_r$ of $\mathcal{L}(aP_0)$, $r = \ell(aP_0)$ (Proposition 2.3.2).

Step 1: a Siegel lemma, by pigeonhole. This step produces coefficients with a linear relation. Consider the map

$$\Lambda \colon \mathcal{L}(bP_0)^r \to \mathcal{L}\big(MaP_0 + b\,\sigma(P_0)\big), \qquad (v_i) \mapsto \sum_i u_i^M \sigma(v_i).$$

It is well defined, since $u_i^M \in \mathcal{L}(MaP_0)$ by Proposition 2.1.3(2) and $\sigma(v_i) \in \mathcal{L}(b\,\sigma(P_0))$ by transport of structure. It is $\mathbb{F}_q$-linear because σ fixes $\mathbb{F}_q$. The source has $q^{r\ell(bP_0)}$ elements and the target $q^{\ell(MaP_0+b\sigma(P_0))}$. By (S) two distinct tuples have the same image, and their difference is a nonzero tuple (v_i) with

$$\sum_i u_i^M \sigma(v_i) = 0. \tag{2.3.2}$$

No height estimate is needed: the size of a solution is its membership in $\mathcal{L}(bP_0)$.

Step 2: non-vanishing. This step defines the section. Put $w = \sum_i u_i v_i^{M'}$. By (N) the pole orders $-v_{P_0}(u_i)$ are distinct integers in $[0, M')$, so Proposition 2.3.4 gives $w \neq 0$. By Proposition 2.1.3(2), $w \in \mathcal{L}((a + M'b)P_0)$.

Step 3: forced multiplicity. This step shows that w vanishes at the points to be counted. Since M is a power of p and $MM' = q$, Proposition 2.1.3(1) gives

$$w^M = \sum_i u_i^M v_i^q. \tag{2.3.3}$$

Let $P \neq P_0$ be σ-Frobenius. All u_i, v_i lie in $\mathcal{O}_P$, because their only pole is at P_0, and $v_i^q - \sigma(v_i) \in P$ by the definition of a σ-Frobenius place. By (2.3.2) and (2.3.3),

$$w^M = \sum_i u_i^M (v_i^q - \sigma(v_i)) \in P.$$

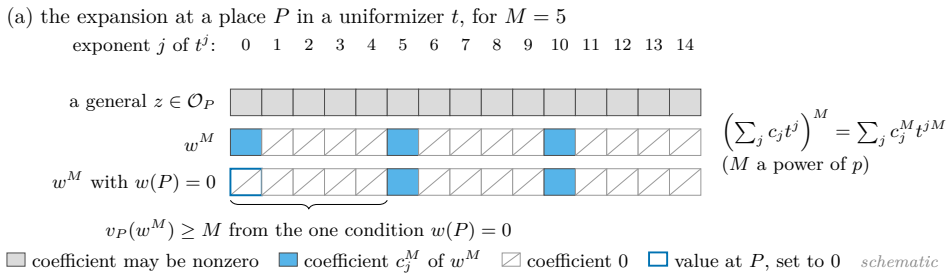
As P is prime, $w \in P$. Thus $v_P(w) \geq 1$, and the section w^M of $\mathcal{O}_C((Ma + qb)P_0)$ vanishes to order at least M at P. No derivative was computed: in a local parameter, $(\sum_{j\geq 1} c_j t^j)^M = \sum_j c_j^M t^{jM}$, so all jets of order less than M of an M-th power vanish once its value does (Figure 2.7).

Step 4: the slope inequality. Let Z_σ be the sum of the σ-Frobenius places $P \neq P_0$; it is finite because each is a zero of $w \neq 0$. Then $(w) + (a + M'b)P_0 - Z_\sigma \geq 0$. An effective divisor has nonnegative degree (S2), so $\deg Z_\sigma \leq a + M'b$. Equivalently, multiplication by w is a nonzero map of line bundles $\mathcal{O}_C(Z_\sigma) \to \mathcal{O}_C((a+M'b)P_0)$, and the degree of the source is at most that of the target. Finally $N(K, \sigma) \leq \deg P_0 + \deg Z_\sigma$. □

For $\sigma = \mathrm{id}$, (2.3.1) is Bombieri's inequality (7) [**Bom74**, p. 238]. Figure 2.8 draws the count of Step 4. The function f of Section 1.1 is w^M: the inequality $m(N-1) \leq mA + qB$ there is (2.3.1) multiplied by M.

2.3.3. Balancing. It remains to find admissible pairs for which the bound (2.3.1) is small.

Proposition 2.3.9. *Let $K/\mathbb{F}_q$ be a function field of genus g, P_0 a place of degree one and $\sigma \in \mathrm{Aut}(K/\mathbb{F}_q)$. Let $M, M' \geq 1$ be integers with $MM' = q$, M a power of p and $M' \geq g + 2$. Then the pair $a = M' - 1$, $b = M + g +$*



(b) conditions imposed, for $q = 121$, $g = 2$, $M = 11$, $(a, b) = (10, 15)$

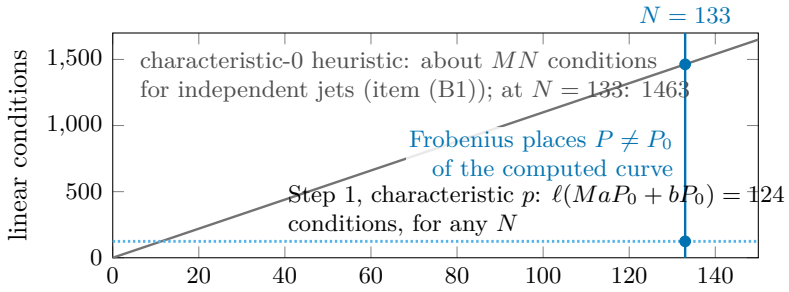


FIGURE 2.7. Multiplicity in characteristic p (Step 3 of Theorem 2.3.8). (a) The local expansion at a place P, in a uniformizer t, for $M = 5$: a general function may have every coefficient nonzero (grey); an M-th power has nonzero coefficients only in degrees divisible by M (sky blue), because $(\sum_j c_j t^j)^M = \sum_j c_j^M t^{jM}$ for M a power of p; if it also vanishes at P (blue outline), its first nonzero coefficient is in degree at least M. One condition gives order M. Schematic. (b) Linear conditions imposed to make w^M vanish to order M at N places, for $q = 121$, $g = 2$, $M = 11$, $(a, b) = (10, 15)$: in Step 1 the number is $\ell(MaP_0 + bP_0) = 124$, whatever N is (sky blue, dotted); for independent jets, as in characteristic 0, it would be about MN (grey; a heuristic count, item (B1) of Section 2.13, not a theorem). Blue: the 133 Frobenius places other than P_0 of the curve of Figure 2.8.

$\lfloor Mg/(M' - 1 - g) \rfloor$ is admissible, and

$$1 + a + M'b = MM' + (g+1)M' + M'\left\lfloor \frac{Mg}{M' - 1 - g} \right\rfloor. \tag{2.3.4}$$

If $M' \geq 2(g+1)$, the right side of (2.3.4) *is at most $MM' + (g+1)M' + 2gM$.*

PROOF. Condition (N) is clear. As $a \geq g+1$, $b \geq g$ and $Ma + b \geq 2g+2$, S3 gives $\ell(aP_0)\ell(bP_0) \geq (a + 1 - g)(b + 1 - g)$ and $\ell(MaP_0 + b\sigma(P_0)) =$

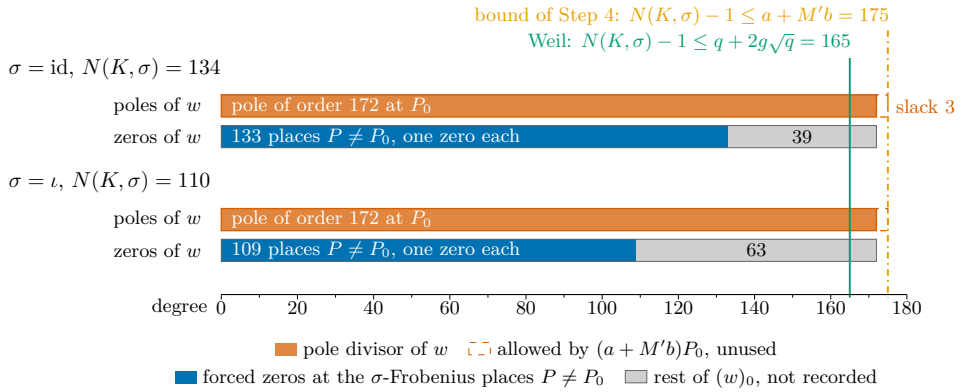


FIGURE 2.8. The degree count of Step 4 of Theorem 2.3.8 for the section computed on a curve of genus 2 over $\mathbb{F}_{121}$ with $M = M' = 11$, $(a,b) = (10,15)$, for $\sigma = \mathrm{id}$ and for $\sigma = \iota$ the hyperelliptic involution. Bars are in degree units. Vermillion: the pole divisor of w, of degree 172 at P_0 for both σ, inside the allowed $(a+M'b)P_0$ of degree 175 (dashed). Blue: the forced zeros, one at each of the $N(K,\sigma)-1$ σ-Frobenius places other than P_0 (133 and 109). Grey: the rest of the zero divisor, of degree 172 minus that number, since $\deg(w)_0 = \deg(w)_\infty = 172$; the computation does not record whether it consists of other zeros or higher multiplicities. Orange, dash-dot: the bound $N(K,\sigma)-1 \le a+M'b = 175$, that is $N(K,\sigma) \le 176$, of Theorem 2.3.8; green: the Weil bound $N(K,\sigma) \le q+1+2g\sqrt{q} = 166$.

$Ma+b+1-g$. Now

$$(2.3.5)\quad (a+1-g)(b+1-g)-(Ma+b+1-g) = (a-g)(b+1-g)-Ma.$$

With $a = M'-1$, $\frac{Ma}{a-g} = M + \frac{Mg}{M'-1-g} < b+1-g$, so (2.3.5) is positive, which is (S). The identity (2.3.4) is a substitution. Finally, writing x for $Mg/(M'-1-g)$, $M'\lfloor x\rfloor \le \frac{M'Mg}{M'-1-g} \le 2gM$ when $M' \le 2(M'-1-g)$, that is when $M' \ge 2(g+1)$. □

Figure 2.9 shows the admissible pairs in one case. So the error in (2.3.1) is at most $(g+1)q/M + 2gM$: the price of (N) falls with M, that of (S) rises. Bombieri takes $M = M' = \sqrt{q}$, assuming q a square. The two spaces then play symmetric roles. The square root in the error does not come from this choice (Proposition 2.3.10).

PROPOSITION 2.3.10. *Let $K/\mathbb{F}_q$ be a function field of genus $g \ge 1$, P_0 a place of degree one, $\sigma \in \mathrm{Aut}(K/\mathbb{F}_q)$, and $q = MM'$ with M a power of p. Let (a,b) be an admissible pair for which the three Riemann–Roch dimensions take their generic values: $\ell(aP_0) = a+1-g$, $\ell(bP_0) = b+1-g$*

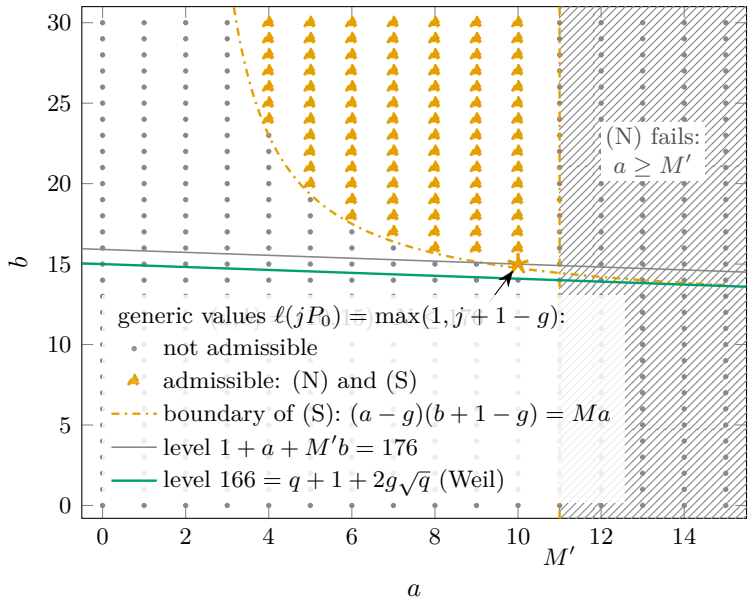


FIGURE 2.9. The admissible pairs of Definition 2.3.7 for $q = 121$, $g = 2$, $M = M' = 11$, $\sigma = \mathrm{id}$, with the generic Riemann–Roch values $\ell(jP_0) = \max(1, j+1-g)$. Condition (N) is $a < 11$; the region where it fails is hatched. Condition (S) holds strictly above the orange dash-dot curve $(a-g)(b+1-g) = Ma$ of (2.3.5). Orange triangles: admissible pairs; grey dots: the others. The level lines of the bound $1+a+M'b$ have slope $-1/M'$, so the best pair takes a as large as (N) allows and then b as small as (S) allows: $(10, 15)$ (star), with $N \le 176$. It is both the pair of Proposition 2.3.9 and Bombieri's pair (Theorem 2.3.12). Green: the level $166 = q + 1 + 2g\sqrt{q}$ of the Weil bound, below every admissible pair. Computed by enumerating the lattice points.

and $\ell(MaP_0 + b\sigma(P_0)) = Ma + b + 1 - g$. *Then*

(2.3.6)

$$1 + a + M'b - q > (g-1)\frac{q}{M} + gM \qquad \text{and} \qquad 1 + a + M'b - q > g + 2\sqrt{gq}.$$

PROOF. By (2.3.5), (S) reads $(a-g)(b+1-g) > Ma$. Here $a \ge g$ because $\ell(aP_0) \ge 1$, and $a = g$ is impossible because $Ma > 0$; so $a > g$ and (S) reads $b + 1 - g > M + \frac{Mg}{a-g}$. Multiplying by M',

$$M'b > q + (g-1)M' + \frac{qg}{a-g}.$$

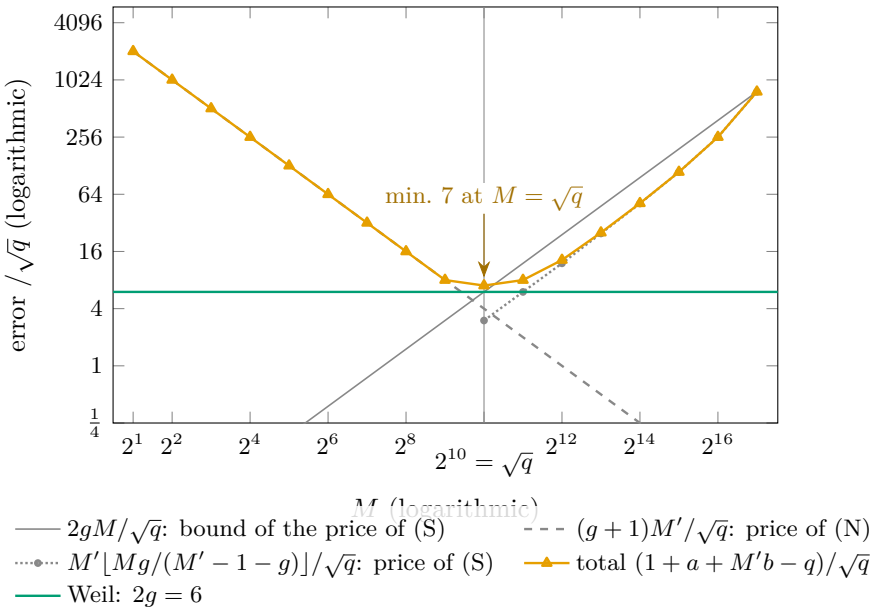


FIGURE 2.10. The error of Proposition 2.3.9, divided by $\sqrt{q}$, for $q = 2^{20}$, $g = 3$ and $M = 2^1, \dots, 2^{17}$, $M' = q/M$, on logarithmic axes. Orange triangles: the total $(1+a+M'b-q)/\sqrt{q}$ of (2.3.4). Grey dashed: its falling part $(g+1)M'/\sqrt{q}$, the price of (N), which caps a below M'. Grey dotted: its rising part $M'\lfloor Mg/(M'-1-g)\rfloor/\sqrt{q}$, the price of (S); it is 0 for $M < \sqrt{q}$, where the total equals the falling part. Thin grey: the bound $2gM/\sqrt{q}$ of the rising part from Proposition 2.3.9. The minimum, 7, is at $M = \sqrt{q}$; by Proposition 2.3.10, an error $O(\sqrt{q})$ requires $M \asymp_g \sqrt{q}$ when $g \geq 2$. Green: Weil's constant $2g = 6$. Data: the closed formula (2.3.4).

For the first inequality use $a - g < M'$, which holds by (N), so that $\frac{qg}{a-g} > gM$. For the second drop $(g-1)M' \geq 0$ and use $a + \frac{qg}{a-g} = g + (a-g) + \frac{qg}{a-g} \geq g + 2\sqrt{gq}$. □

REMARK 2.3.11. The square root does not enter the error of Theorem 2.3.8 through the choice $M = M' = \sqrt{q}$. By Proposition 2.3.10, no choice of M and (a, b) with generic Riemann–Roch values gives an error $o(\sqrt{q})$, and for $g \geq 2$ an error $O(\sqrt{q})$ requires $M \asymp_g \sqrt{q}$. The second inequality of (2.3.6) is the arithmetic–geometric mean inequality. The product of the two competing costs is gq whatever the split (Figure 2.10). Chapter 1 follows this balance to the integers. There the two costs are x/m and m, and the chain length is $m \approx \sqrt{x}$ (Section 1.5, Proposition 9.3.1).

2.3.4. Bombieri's theorem.

THEOREM 2.3.12 (Bombieri, twisted). *Let $g \geq 0$ be an integer and $q = q_0^2$ with q_0 a power of p and $q_0 > (g+1)^2$. Let $K/\mathbb{F}_q$ be a function field of genus g with a place of degree one. Then for every $\sigma \in \mathrm{Aut}(K/\mathbb{F}_q)$*

$$N(K,\sigma) \leq q + 2gq_0 + \left\lfloor \frac{gq_0}{g+1} \right\rfloor + g + 2 \leq q + (2g+1)q_0. \tag{2.3.7}$$

In particular every function field of genus g over $\mathbb{F}_q$ has $N_1 < q + 1 + (2g+1)\sqrt{q}$.

PROOF. Take $M = M' = q_0$, $b = q_0 + 2g$ and $a = \lfloor gq_0/(g+1) \rfloor + g + 1$ [**Bom74**, p. 239]. Condition (N): $a \leq q_0 - 1$ is equivalent to $\frac{gq_0}{g+1} < q_0 - g - 1$, that is to $q_0 > (g+1)^2$. Condition (S): as $a \geq g+1$, $b \geq g$ and $Ma+b \geq 2g-1$, fact S3 and (2.3.5) show that it suffices that $(a-g)(q_0+g+1) > q_0 a$, that is $a > \frac{gq_0}{g+1} + g$, which holds. Then $1 + a + q_0 b$ is the middle term of (2.3.7), and the last inequality of (2.3.7) is $a + 1 \leq q_0$, which is (N). If K has no place of degree one, $N_1 = 0$. □

This is Bombieri's Theorem 1 [**Bom74**, p. 236], and for $\sigma = \mathrm{id}$ it is Proposition 1.1.1.

REMARK 2.3.13 (The pair used in Chapter 1). The sketch in Section 1.1 uses, with $M = M' = q_0$, the pair $a = q_0 - 1$, $b = q_0 + 2g$ in place of Bombieri's. It is admissible under the same hypothesis: (N) is clear, and by (2.3.5) the margin in (S) is at least $(q_0 - 1 - g)(q_0 + g + 1) - q_0(q_0 - 1) = q_0 - (g+1)^2 > 0$. It gives $1 + a + q_0 b = q + (2g+1)q_0$, the last term of (2.3.7), where Bombieri's pair gives the slightly smaller middle term. Condition (1.1.1) of Chapter 1 is (S) with the Riemann–Roch lower bounds inserted.

Both (N) and (S) in the proof of Theorem 2.3.12 reduce to chains of equivalences ending in $q_0 > (g+1)^2$. The chain for (S) used the lower bound $\ell(aP_0) \geq a + 1 - g$, which is an equality only when $a \geq 2g - 1$. So the chain does not show that $q_0 > (g+1)^2$ is necessary.

REMARK 2.3.14. Let $b = q_0 + 2g$. For generic Riemann–Roch values an admissible a exists if and only if $q_0 > (g+1)^2$. At points with special gap sequences an admissible a can exist with $q_0 \leq (g+1)^2$. Let $q = 9$, $q_0 = 3$, K the function field of $y^2 = f(x)$ with $f \in \mathbb{F}_9[x]$ squarefree of degree 7, so $g = 3$, P_0 the place at infinity and $\sigma = \mathrm{id}$. Then $1, x \in \mathcal{L}(2P_0)$, so $\ell(2P_0) = 2$, while $\ell(9P_0) = 7$ and $\ell(15P_0) = 13$. So $(a,b) = (2,9)$ satisfies (N) and (S), since $2 \cdot 7 > 13$, although $q_0 = 3 \leq 16$, and it gives $N_1 \leq 30$. This cannot happen if $q_0 \geq g + 1$. Indeed, let (a,b) be admissible with $b = q_0 + 2g$ and $M = M' = q_0$. By S3, $\ell(bP_0) = q_0 + g + 1$ and the right side of (S) is $q_0 a + q_0 + g + 1$. If $a \geq 2g - 1$, then $\ell(aP_0) = a + 1 - g$ and the chain of equivalences in the proof of Theorem 2.3.12 is exact, so $q_0 > (g+1)^2$. If $a \leq 2g - 2$, then Clifford's theorem, which is not proved here, gives $\ell(aP_0) \leq a/2 + 1$, and (S) gives $(a/2)(q_0 + g + 1) > q_0 a$, so $a > 0$ and $q_0 < g + 1$.

The exposition [**Tao14**] carries out Bombieri's choice with the same parameters, and states it for every $q \geq (g+1)^4$.

REMARK 2.3.15. The parameter calculation of [**Tao14**] does not reach every $q \geq (g+1)^4$, for two reasons. At $q = (g+1)^4$ the parameter $\ell = \lfloor \frac{g}{g+1}\sqrt{q} \rfloor + g + 1$ of [**Tao14**] equals $\sqrt{q}$, against its condition $\ell < \sqrt{q}$. Its inequality (11) has right side $\ell\sqrt{q}+m+1$ instead of $\ell\sqrt{q}+m+1-g$, and then holds exactly when $g+1$ divides $g\sqrt{q}$; for $g = 1$, $q = 25$ it reads $28 > 28$. With the exact value it is correct (Theorem 2.3.12).

For $q = p^e$ with e odd no power of p equals $\sqrt{q}$, and the balanced choice is unavailable. The upper bound still has an error of order $\sqrt{q}$.

PROPOSITION 2.3.16. *Let $g \geq 0$ be an integer and $q = p^e$ with $\sqrt{q} \geq 2(g+1)$, and let $K/\mathbb{F}_q$ be a function field of genus g with a place of degree one. Then $N(K,\sigma) \leq q + ((g+1)\sqrt{p} + 2g)\sqrt{q}$ for every $\sigma \in \mathrm{Aut}(K/\mathbb{F}_q)$. Hence for every function field $K/\mathbb{F}_q$ of genus g and every n with $q^{n/2} \geq 2(g+1)$,*

$$N_n \leq q^n + \big((g+1)\sqrt{p} + 2g\big)q^{n/2}. \tag{2.3.8}$$

PROOF. Take $M = p^{\lfloor e/2 \rfloor}$ and $M' = p^{\lceil e/2 \rceil}$, so that $M \leq \sqrt{q} \leq M' \leq \sqrt{pq}$ and $M' \geq 2(g+1)$, and apply Proposition 2.3.9 and Theorem 2.3.8. For (2.3.8) apply this to K_n, which has full constant field $\mathbb{F}_{q^n}$, genus g (Proposition 2.2.7) and N_n places of degree one (S4); if $N_n = 0$ there is nothing to prove. □

REMARK 2.3.17. The upper bound does not need q to be a square to give an error of order $\sqrt{q}$. By Proposition 2.3.16, the square case only improves the constant, from $(g+1)\sqrt{p} + 2g$ to $2g+1$. Figure 2.11 compares the constants.

2.4. Route 1: the lower bound by Galois averaging

An upper bound for point counts is half of Theorem 2.1.2. Following [**Bom74**, p. 240], the other half comes from the same upper bound, applied to all the twists of a Galois cover of the projective line at once: the total of their point counts is known exactly, so if none is too large, none is too small (Figure 2.12). This is the reason for the automorphism σ in Theorem 2.3.8.

DEFINITION 2.4.1. Let E/F be a finite Galois extension of function fields with group G, both with full constant field $\mathbb{F}_q$, and let P be a place of E.

(1) $D(P) = \{\sigma \in G : \sigma(P) = P\}$ is the *decomposition group*; each $\sigma \in D(P)$ induces an automorphism $\bar{\sigma}$ of $\kappa(P)$ over $\kappa(P \cap F)$.
(2) $I(P)$, the *inertia group*, is the kernel of $\sigma \mapsto \bar{\sigma}$ on $D(P)$.
(3) $\Phi(P) = \{\sigma \in G : P \text{ is } \sigma\text{-Frobenius}\}$.

PROPOSITION 2.4.2. *Let E/F be a finite Galois extension of function fields with group G, both with full constant field $\mathbb{F}_q$, let P be a place of E, $Q = P \cap F$ and $\sigma \in G$. If P is σ-Frobenius, then* $\deg Q = 1$ *and* $\sigma \in D(P)$.

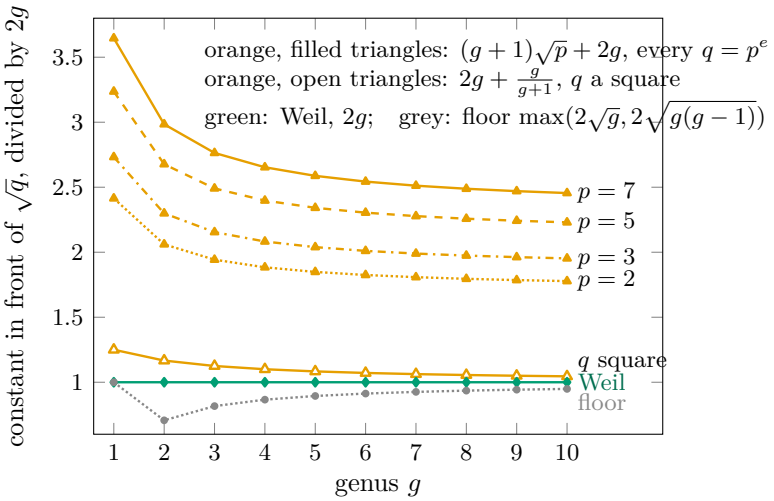


FIGURE 2.11. The constants in front of $\sqrt{q}$, divided by $2g$, for $1 \le g \le 10$. Orange, filled triangles: the constant $(g+1)\sqrt{p}+2g$ of Proposition 2.3.16, valid for every $q=p^e$ with $\sqrt{q} \ge 2(g+1)$, for $p = 7, 5, 3, 2$ (solid, dashed, dash-dot, dotted); it tends to $1+\sqrt{p}/2$. Orange, open triangles: $2g+\frac{g}{g+1}$ of Theorem 2.3.12 for q a square with $q > (g+1)^4$, which tends to 1. Green: Weil's sharp constant $2g$ (Theorem 2.5.3). Grey dotted: $\max(2\sqrt{g}, 2\sqrt{g(g-1)})$, which by Proposition 2.3.10 the method of Theorem 2.3.8 cannot reach when the Riemann–Roch values are generic. It equals $2g$ at $g=1$ and lies below it for $g \ge 2$. Closed formulas.

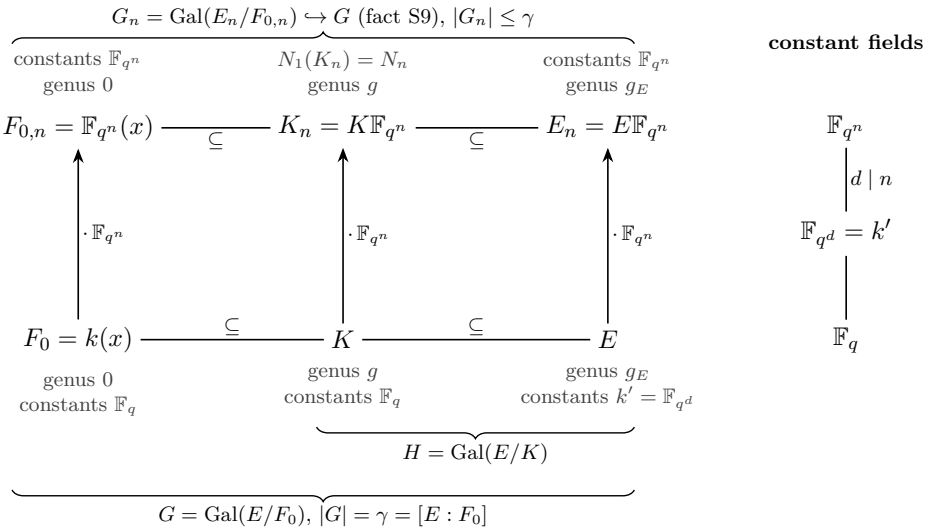
PROOF. For $z \in \mathcal{O}_Q$ we have $\sigma(z) = z$, so $z^q - z \in P \cap F = Q$. Every element of $\kappa(Q)$ is then a root of $X^q - X$, so $\kappa(Q) = \mathbb{F}_q$. That $\sigma(P) = P$ was shown after Definition 2.3.1. □

PROPOSITION 2.4.3. *Let E/F be a finite Galois extension of function fields with group G, both with full constant field $\mathbb{F}_q$, let Q be a place of F of degree one and P a place of E above Q. Then $\Phi(P)$ is a coset of $I(P)$ in $D(P)$, in particular nonempty, and*

$$|\Phi(P)| \cdot \deg P = |D(P)|. \tag{2.4.1}$$

PROOF. Here $\kappa(Q) = \mathbb{F}_q$, and $\mathrm{Gal}(\kappa(P)/\mathbb{F}_q)$ is cyclic of order $f = \deg P$, generated by Fr_q (S8). By Proposition 2.4.2 and the definition of a σ-Frobenius place, $\Phi(P) = \{\sigma \in D(P) : \bar{\sigma} = \mathrm{Fr}_q\}$, which is empty or a coset of $I(P)$.

We show first that $\Phi(P)$ is nonempty. Let $P = P_1, \dots, P_s$ be the places above Q (S7). Let $\bar{\theta}$ generate the group $\kappa(P)^\times$ (S8); so $\bar{\theta} \ne 0$ and $\kappa(P) = \mathbb{F}_q(\bar{\theta})$. By S1 there is $\theta \in E$ with $\theta \in \mathcal{O}_P$, $\theta(P) = \bar{\theta}$ and $v_{P_j}(\theta) \ge 1$ for



The tower. Take $n \in n_1\mathbb{Z}_{>0}$. Then q^n is a square, $q^{n/2} > (g_E+1)^2$ and E_n has a place of degree one. Theorem 2.3.12 for $E_n/\mathbb{F}_{q^n}$ bounds $N(E_n,\sigma)$ from above for every $\sigma \in G_n$, and Theorem 2.4.6 for $E_n \supseteq K_n \supseteq F_{0,n}$ turns these into $|N_n - q^n - 1| \le \max(1,\gamma-1)(2g_E+1)q^{n/2}$ (Theorem 2.4.7).

FIGURE 2.12. The fields of Section 2.4 and Subsection 2.4.1; inclusions run left to right, constant field extensions bottom to top. Bottom row: $K/\mathbb{F}_q$ is separable over $F_0 = k(x)$, $k = \mathbb{F}_q$, and E is its Galois closure, with $G = \mathrm{Gal}(E/F_0)$, $H = \mathrm{Gal}(E/K)$ and $\gamma = [E : F_0]$; the constant field of E is $k' = \mathbb{F}_{q^d}$, which may be larger than $\mathbb{F}_q$. Top row: the constant field extensions to $\mathbb{F}_{q^n}$, $d \mid n$; $E_n/F_{0,n}$ is Galois with group $G_n \hookrightarrow G$ (fact S9), all three fields have full constant field $\mathbb{F}_{q^n}$, and E_n, K_n have the genera g_E, g of E, K (Proposition 2.2.7). Theorem 2.4.7 applies Theorem 2.4.6 to the top row for $n \in n_1\mathbb{Z}_{>0}$, as in the box; there $N_1(K_n) = N_n$ by fact S4. Schematic.

$j \ge 2$. For $\tau \in G$, $v_P(\tau\theta) = v_{\tau^{-1}P}(\theta) \ge 0$. Put $\chi(X) = \prod_{\tau\in G}(X - \tau\theta)$. Its coefficients are G-invariant and lie in $\mathcal{O}_P$, hence in $F \cap \mathcal{O}_P = \mathcal{O}_Q$. Its reduction $\bar\chi = \prod_\tau(X - \overline{\tau\theta})$ lies in $\mathbb{F}_q[X]$ and has the root $\bar\theta$. So $\bar\chi(\bar\theta^q) = \bar\chi(\bar\theta)^q = 0$, and $\bar\theta^q = \overline{\tau\theta}$ for some τ. As $\bar\theta^q \ne 0$, $v_{\tau^{-1}P}(\theta) = 0$, which forces $\tau^{-1}P = P$. So $\tau \in D(P)$ and $\bar\tau(\bar\theta) = \bar\theta^q$; since $\bar\theta$ generates $\kappa(P)$, $\bar\tau = \mathrm{Fr}_q$.

It remains to count. The image of $D(P)$ in the cyclic group $\langle \mathrm{Fr}_q\rangle$ contains a generator, so it is everything. Hence $|D(P)| = |I(P)|f$ and $|\Phi(P)| = |I(P)|$. □

Figure 2.13 shows the coset $\Phi(P)$ at an unramified and at a ramified place.

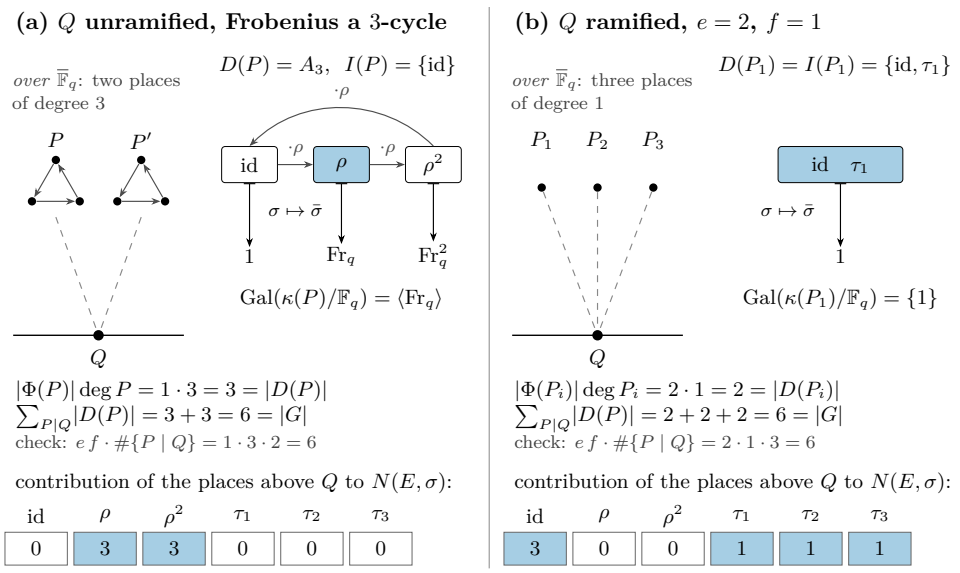


FIGURE 2.13. Frobenius cosets (Proposition 2.4.3) in a Galois extension E/F of function fields with group $G = S_3 = \{\mathrm{id}, \rho, \rho^2, \tau_1, \tau_2, \tau_3\}$ and full constant field $\mathbb{F}_q$, over a place Q of F of degree one; such places occur, for example, in the Galois closure of a cubic extension of $\mathbb{F}_q(x)$. Left in each panel: the geometric points above Q, grouped into places (Frobenius orbits, arrows). Right: $D(P)$ as cosets of $I(P)$ (outlined boxes), mapped onto $\mathrm{Gal}(\kappa(P)/\mathbb{F}_q)$ by $\sigma \mapsto \bar{\sigma}$. (a) Q unramified with Frobenius a 3-cycle: two places above Q, each of degree 3, with $D(P) = A_3$ and $I(P) = \{\mathrm{id}\}$; the coset over Fr_q is $\Phi(P) = \{\rho\}$ (blue), and $\Phi(P') = \{\rho^2\}$. (b) Q ramified with $D(P_i) = I(P_i) = \{\mathrm{id}, \tau_i\}$: three places of degree one, and $\Phi(P_i) = I(P_i)$ (blue). In both cases $|\Phi(P)| \deg P = |D(P)|$ and $\sum_{P|Q} |D(P)| = |G| = 6$. The bottom rows give, for each $\sigma \in G$, the degree of the σ-Frobenius places above Q; each row sums to $|G|$, so every place of degree one of F, ramified or not, contributes exactly $|G|$ to $\sum_\sigma N(E, \sigma)$ (Proposition 2.4.4). Blue: the σ for which a place above Q is σ-Frobenius. Schematic positions; the group theory is exact.

The sum of $N(E, \sigma)$ over $\sigma \in G$ counts each rational place of the base once for every way it can be a Frobenius fixed point upstairs. At an unramified place this is the familiar statement that Frobenius is a well-defined conjugacy class. At a branch point it is less clear what should be counted: the residue field extension no longer determines one element of G, and Bombieri's identity [**Bom74**, p. 240, eq. (9)] accordingly carries an error

term $O(1)$ for the branch points, which are finite in number. For the count of Definition 2.3.1 the identity is exact.

PROPOSITION 2.4.4 (class-sum identity). *Let E/F be a finite Galois extension of function fields with group G, both with full constant field $\mathbb{F}_q$, and let $N_1(F)$ be the number of places of F of degree one. Then*

$$\sum_{\sigma \in G} N(E,\sigma) = |G| \cdot N_1(F). \tag{2.4.2}$$

PROOF. The left side is $\sum_P |\Phi(P)| \deg P$, the sum over the places of E. By Propositions 2.4.2 and 2.4.3 only the P over places Q of degree one contribute, and each contributes $|D(P)|$. The places P over a given Q form one G-orbit (S7), with stabilisers $D(P)$, so together they contribute $|G|$. □

REMARK 2.4.5. For the count of Definition 2.3.1, the averaging identity over all twists of a Galois cover has no error term from the ramified places. A ramified place over a rational place is σ-Frobenius for $|I(P)|$ elements σ, and this compensates exactly. Bombieri counts unramified points only, and his identity has the error term for that reason.

THEOREM 2.4.6. *Let $F_0 = \mathbb{F}_q(x)$, let E/F_0 be a finite Galois extension with group G, and let F be an intermediate field, all three with full constant field $\mathbb{F}_q$, and let $N_1(F)$ be the number of places of F of degree one. Let $\Theta \geq 0$ and suppose $N(E,\sigma) \leq q+1+\Theta$ for all $\sigma \in G$. Then $N(E,\tau) \geq q+1-(|G|-1)\Theta$ for every $\tau \in G$, and*

$$q+1-(|G|-1)\Theta \leq N_1(F) \leq q+1+\Theta.$$

PROOF. By (2.4.2) for E/F_0 and S5, $\sum_\sigma N(E,\sigma) = |G|(q+1)$. Subtracting the upper bounds for the $\sigma \neq \tau$ gives the first bound. By (2.4.2) for E/F, $N_1(F)$ is the average of $N(E,\tau)$ over $\tau \in \mathrm{Gal}(E/F)$, and each of these lies between the two bounds. □

The base F_0 has trivial L-polynomial, so its point count $q+1$ is known exactly. Figure 2.14 shows the counts of the twists in an example. There is no base of this kind over the integers, which is one of the places where the analogy breaks (Section 2.13, item (B3)).

2.4.1. The Galois closure and the tower. Let $K/\mathbb{F}_q$ be a function field of genus g. To apply Theorem 2.4.6 to K we fix the following data.

(1) Choose $x \in K$ with $K/\mathbb{F}_q(x)$ finite separable (S6). Let E be the Galois closure of $K/\mathbb{F}_q(x)$ and $\gamma = [E : \mathbb{F}_q(x)]$.
(2) Let $\mathbb{F}_{q^d}$ be the algebraic closure of $\mathbb{F}_q$ in E. It is finite over $\mathbb{F}_q$, because $[\mathbb{F}_{q^d}(x) : \mathbb{F}_q(x)] = d \leq \gamma$. Then $E/\mathbb{F}_{q^d}$ is a function field with full constant field $\mathbb{F}_{q^d}$; let g_E be its genus. Choose a place P_1 of E and let d_1 be its degree over $\mathbb{F}_{q^d}$.
(3) For n divisible by d put $E_n = E\mathbb{F}_{q^n}$, $K_n = K\mathbb{F}_{q^n}$ and $F_{0,n} = \mathbb{F}_{q^n}(x)$. By S4 and Proposition 2.2.7 applied to $E/\mathbb{F}_{q^d}$, E_n has full constant field $\mathbb{F}_{q^n}$ and genus g_E; so K_n, lying between $F_{0,n}$ and E_n, has full

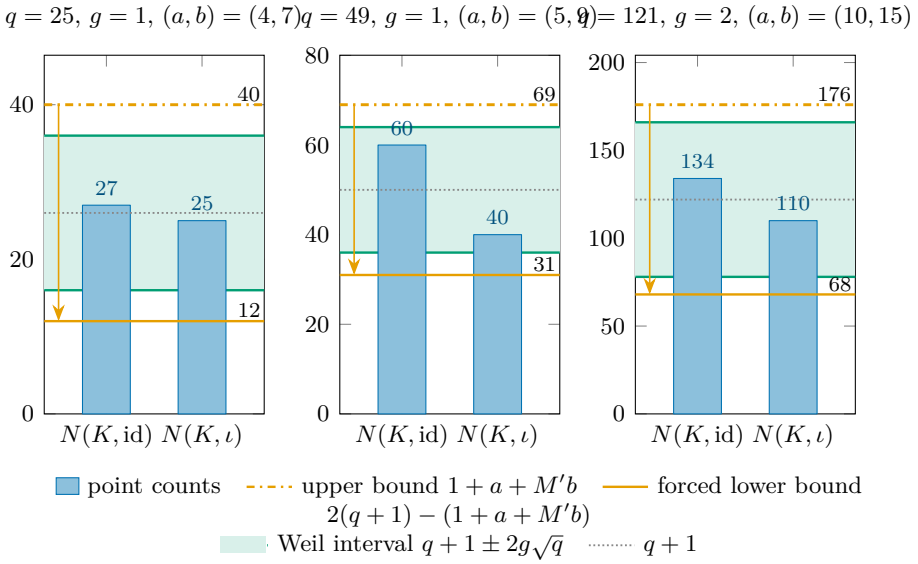


FIGURE 2.14. Route 1 in three computed cases with $G = \mathbb{Z}/2$: $y^2 = f(x)$ with $\deg f = 2g+1$ over $\mathbb{F}_{25}$ and $\mathbb{F}_{49}$ ($g = 1$) and over $\mathbb{F}_{121}$ ($g = 2$), with ι the hyperelliptic involution and $M = M' = \sqrt{q}$. Blue bars: the counts $N(K, \mathrm{id})$ and $N(K, \iota)$; their sum is exactly $2(q + 1)$ (Proposition 2.4.4). Orange dash-dot: the upper bound $1 + a + M'b$ of Theorem 2.3.8 for both counts. Orange solid: the lower bound $2(q + 1) - (1 + a + M'b)$ that it forces on each count (Theorem 2.4.6); it is the reflection of the upper bound in $q + 1$ (arrow). Green band: the Weil interval $q+1\pm 2g\sqrt{q}$; grey dotted: $q+1$. The computed $q = 27$ cases are omitted; with the split $(M, M') = (9, 3)$ the forced lower bound there is negative. The second computed $q = 49$ pair, $(a, b) = (6, 9)$, is also omitted.

constant field $\mathbb{F}_{q^n}$. By S9, $E_n/F_{0,n}$ is Galois with group G_n of order at most γ, and every element of G_n fixes $\mathbb{F}_{q^n}$.

(4) If also $dd_1 \mid n$, then E_n has a place of degree one, by S4 applied to P_1.

Let n_1 be the least even multiple of dd_1 with $q^{n_1/2} > (g_E + 1)^2$.

THEOREM 2.4.7. *Let $K/\mathbb{F}_q$ be a function field, let $x \in K$ be such that $K/\mathbb{F}_q(x)$ is finite and separable, let E be the Galois closure of $K/\mathbb{F}_q(x)$ and $\gamma = [E : \mathbb{F}_q(x)]$, let $\mathbb{F}_{q^d}$ be the algebraic closure of $\mathbb{F}_q$ in E and g_E the genus of $E/\mathbb{F}_{q^d}$, let d_1 be the degree over $\mathbb{F}_{q^d}$ of a place of E, and let n_1 be the least even multiple of dd_1 with $q^{n_1/2} > (g_E + 1)^2$. Then*

$$|N_n - q^n - 1| \le \max(1, \gamma - 1)(2g_E + 1)\, q^{n/2} \qquad \text{for all } n \in n_1\mathbb{Z}_{>0}. \tag{2.4.3}$$

PROOF. Let $n \in n_1\mathbb{Z}_{>0}$. Then q^n is a square, $q^{n/2} > (g_E+1)^2$, and E_n has a place of degree one. Theorem 2.3.12 for $E_n/\mathbb{F}_{q^n}$ gives $N(E_n,\sigma) \le q^n + (2g_E+1)q^{n/2}$ for all $\sigma \in G_n$. Theorem 2.4.6 for $E_n \supseteq K_n \supseteq F_{0,n}$ over $\mathbb{F}_{q^n}$, with $\Theta = (2g_E+1)q^{n/2}-1$, gives

$$q^n + 1 - (\gamma-1)(2g_E+1)q^{n/2} \le N_1(K_n) \le q^n + (2g_E+1)q^{n/2}.$$

By S4, $N_1(K_n) = N_n$. □

The next section uses only that γ and g_E do not depend on n. The constants in (2.4.3) are poor, and the bound holds only along one arithmetic progression; both defects disappear at the next step.

2.5. The Riemann hypothesis

PROPOSITION 2.5.1. *Let ϕ be a real number and I a finite nonempty interval of integers. If $\phi \notin 2\pi\mathbb{Z}$, then*

$$\Big|\frac{1}{|I|}\sum_{k\in I} e^{ik\phi}\Big| \le \frac{2}{|I|\,|1-e^{i\phi}|};$$

if $\phi \in 2\pi\mathbb{Z}$ the mean is 1.

PROOF. Sum the geometric series. □

A set W of positive integers *contains arbitrarily long intervals* if for every n it contains n consecutive integers. In such a set one can choose intervals $I_\nu \subseteq W$ with $|I_\nu| \to \infty$ and $\min I_\nu \to \infty$, for instance the upper halves of longer and longer intervals in W.

PROPOSITION 2.5.2. *Let $\beta_1,\dots,\beta_r \in \mathbb{C}$, let $R, A > 0$, and let W be a set of positive integers containing arbitrarily long intervals. If $|\sum_i \beta_i^k| \le AR^k$ for all $k \in W$, then $|\beta_i| \le R$ for all i.*

PROOF. Suppose $\rho = \max_i|\beta_i| > R$. Let $\rho e^{i\theta_1},\dots,\rho e^{i\theta_s}$ be the distinct values of modulus ρ among the β_i, with multiplicities $\mu_j \ge 1$. Then $T_k := \rho^{-k}\sum_i \beta_i^k = \sum_j \mu_j e^{ik\theta_j} + \epsilon_k$ with $\epsilon_k \to 0$, and $|T_k| \le A(R/\rho)^k \to 0$ on W. Choose intervals $I_\nu \subseteq W$ with $|I_\nu| \to \infty$ and $\min I_\nu \to \infty$. By Proposition 2.5.1 the mean of $T_k e^{-ik\theta_1}$ over I_ν tends to $\mu_1 \ge 1$. It also tends to 0, a contradiction. □

No sign is used here: the power sums are bounded in absolute value, and no cancellation among the β_i of largest modulus can hide them all. Compare Proposition 2.10.1, the same statement for weighted sums and all k, which is where the sign-only proof needs its sign.

THEOREM 2.5.3 (the Riemann hypothesis for K). *Let $K/\mathbb{F}_q$ be a function field of genus g with reciprocal roots $\alpha_1,\dots,\alpha_{2g}$. Then $|\alpha_i| = \sqrt{q}$ for all i, and $|N_n - q^n - 1| \le 2g\,q^{n/2}$ for all $n \ge 1$.*

PROOF. By Theorem 2.2.6(3) and (2.4.3),

$$\Big|\sum_i (\alpha_i^{n_1})^k\Big| = |q^{n_1 k} + 1 - N_{n_1 k}| \le A\,(q^{n_1/2})^k \qquad \text{for all } k \ge 1,$$

with $A = \max(1, \gamma - 1)(2g_E + 1)$. Proposition 2.5.2 gives $|\alpha_i|^{n_1} \le q^{n_1/2}$, so $|\alpha_i| \le \sqrt{q}$. By Theorem 2.2.6(3), q/α_i is also a reciprocal root, so $q/|\alpha_i| \le \sqrt{q}$. Hence $|\alpha_i| = \sqrt{q}$, and $|N_n - q^n - 1| = |\sum_i \alpha_i^n| \le 2gq^{n/2}$. □

This proves Theorem 2.1.2. The bounds were needed only for n in one arithmetic progression, with constants independent of n; the sharp constant $2g$ and all n come out at the end, from the tensor-power step of Proposition 2.5.2. Figure 2.15 shows the reciprocal roots of some curves.

2.6. Route 2: a one-sided bound and a Landau–Pringsheim lemma

This section asks whether the upper bound of Section 2.3, without the averaging of Section 2.4, gives the Riemann hypothesis through a positivity lemma of Landau–Pringsheim type. That is how the analytic step of the number-field argument works: a one-sided bound and Landau's theorem (Section 1.3, Theorem 4.3.3). Over a function field it does not work with the bound that Section 2.3 provides, and we identify the missing input exactly. The section then finds the function-field statement to which the number-field criterion does correspond.

2.6.1. The lemma.

PROPOSITION 2.6.1 (Landau–Pringsheim for exponential sums). *Let $\beta_1, \dots, \beta_s \in \mathbb{C}^\times$ be distinct, let $c_1, \dots, c_s \in \mathbb{C}^\times$, and let W be a set of positive integers containing arbitrarily long intervals. Suppose that $\sum_j c_j \beta_j^k$ is real and nonnegative for all $k \in W$. Then the largest of the moduli $|\beta_j|$ is itself one of the β_j; the coefficient of this β_j is real and positive; and it is at least $|c_i|$ for every i for which $|\beta_i|$ is largest.*

PROOF. Put $a_k = \sum_j c_j \beta_j^k$ and $\rho = \max_j |\beta_j|$. Let $J = \{j : |\beta_j| = \rho\}$ and $\beta_j = \rho e^{i\theta_j}$ for $j \in J$, the θ_j distinct modulo 2π. Put $b_k = \rho^{-k} a_k = \sum_{j \in J} c_j e^{ik\theta_j} + \epsilon_k$ with $\epsilon_k \to 0$. Choose intervals $I_\nu \subseteq W$ with $|I_\nu| \to \infty$ and $\min I_\nu \to \infty$. For real ψ let $m(\psi) = \lim_\nu \frac{1}{|I_\nu|} \sum_{k \in I_\nu} b_k e^{-ik\psi}$. By Proposition 2.5.1, $m(\theta_j) = c_j$ for $j \in J$, and $m(\psi) = 0$ if ψ is congruent to no θ_j. Fix $j_1 \in J$ and a complex number ϵ of modulus one. Since b_k is real and nonnegative on W and $1 + \mathrm{Re}(\epsilon e^{-ik\theta_{j_1}}) \ge 0$,

$$0 \le \lim_\nu \frac{1}{|I_\nu|} \sum_{k \in I_\nu} b_k \big(1 + \mathrm{Re}(\epsilon e^{-ik\theta_{j_1}})\big) = m(0) + \mathrm{Re}\big(\epsilon\, c_{j_1}\big).$$

Choosing $\epsilon = -\overline{c_{j_1}}/|c_{j_1}|$ gives $m(0) \ge |c_{j_1}| > 0$. So some $\theta_{j_0} \equiv 0$, that is $\beta_{j_0} = \rho$, and $c_{j_0} = m(0)$ is real, positive and at least $|c_{j_1}|$. □

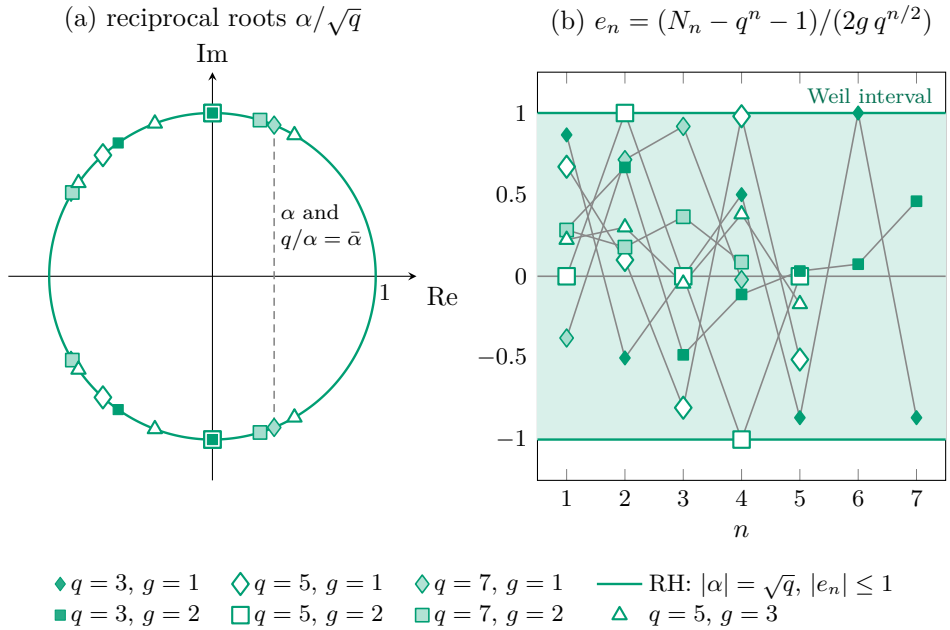


FIGURE 2.15. The Riemann hypothesis (Theorem 2.5.3) for seven curves $y^2 = f(x)$ of genus 1, 2, 3 over $\mathbb{F}_3$, $\mathbb{F}_5$, $\mathbb{F}_7$. Marker shape gives the genus (diamond $g = 1$, square $g = 2$, triangle $g = 3$) and marker fill gives q (filled $q = 3$, open $q = 5$, light $q = 7$); all data are green, the colour of genuine curves. (a) The reciprocal roots α of each L, divided by $\sqrt{q}$, lie on the unit circle (green). By Theorem 2.2.6(3), $\alpha \mapsto q/\alpha$ permutes them, and on the circle $q/\alpha = \bar{\alpha}$, so they come in conjugate pairs (dashed: one pair). The curve over $\mathbb{F}_5$ of genus 2 has double roots $\pm i\sqrt{5}$; the curve over $\mathbb{F}_3$ of genus 2 also has roots at $\pm i\sqrt{3}$. (b) The normalised errors e_n from brute-force point counts lie in the Weil interval $[-1, 1]$ (green band), which the curve over $\mathbb{F}_5$ of genus 2 reaches at $n = 2$ and $n = 4$; grey lines join the values of one curve. The roots were computed numerically from the L-polynomials determined by the counts.

The generating function $\sum_k a_k t^k = \sum_j \frac{c_j \beta_j t}{1-\beta_j t}$ has its poles at $t = 1/\beta_j$. The proposition says that if the coefficients are nonnegative, the pole nearest to 0 lies on the positive real axis and its coefficient is positive (Figure 2.16): Pringsheim's theorem for this class, of which Landau's theorem is the Dirichlet-series form [**Tit39**]. Restricting to a set W of windows costs nothing here, because there are finitely many frequencies; over the integers the windowed form is a theorem with a proof of its own (Theorem 4.3.3).

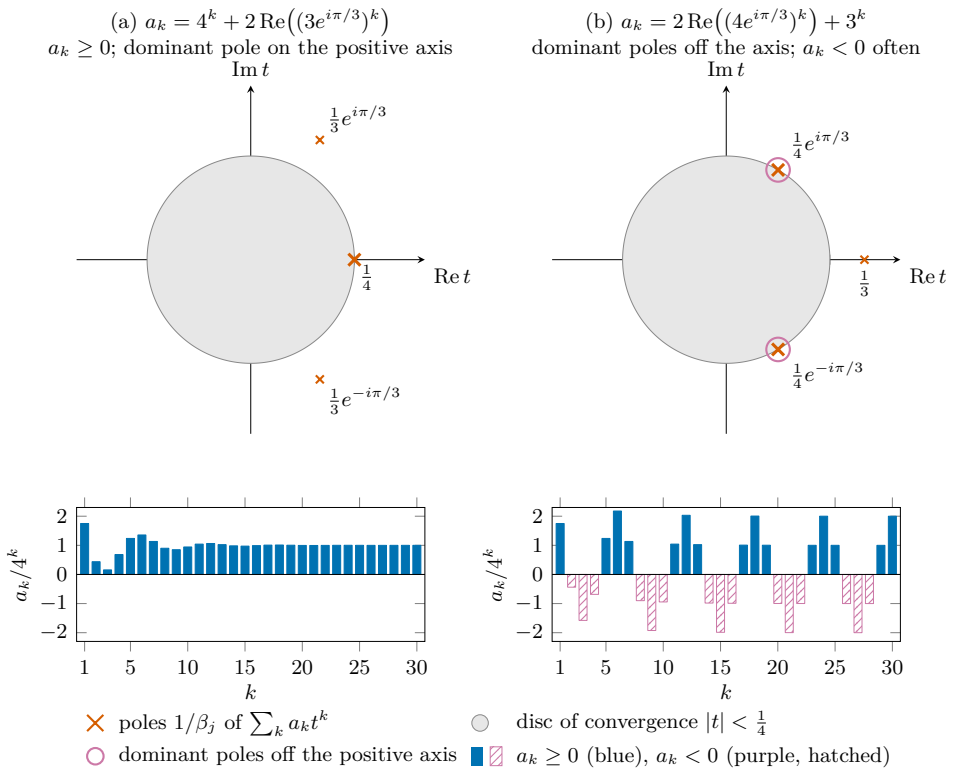


FIGURE 2.16. Proposition 2.6.1 in the t-plane. Each column shows a sequence $a_k = \sum_j c_j \beta_j^k$, the poles $t = 1/\beta_j$ of $\sum_k a_k t^k = \sum_j c_j \beta_j t/(1-\beta_j t)$ (vermillion crosses, the dominant ones larger), the disc of convergence $|t| < 1/4$ (grey), and below it the values $a_k/4^k$ for $k = 1, \ldots, 30$ (blue bar: $a_k \geq 0$; purple hatched bar: $a_k < 0$). (a) $a_k = 4^k + 2\,\mathrm{Re}\big((3e^{i\pi/3})^k\big)$ is nonnegative for every $k \geq 1$, and the pole on the boundary of the disc is on the positive real axis, at $1/4$, with positive coefficient, as the proposition requires. (b) $a_k = 2\,\mathrm{Re}\big((4e^{i\pi/3})^k\big) + 3^k$ has its poles on the boundary of the disc at $e^{\pm i\pi/3}/4$, off the real axis (purple circles), and $a_k < 0$ whenever $k \equiv 2, 3, 4 \pmod 6$, so on every window of three consecutive k. The proposition excludes this configuration for a sequence that is nonnegative on arbitrarily long intervals. Values from the closed formulas.

2.6.2. What a one-sided bound gives. Section 2.3 gives $N_n \leq q^n + 1 + Aq^{n/2}$, so the numbers $Aq^{n/2} + \sum_i \alpha_i^n$ are nonnegative. Positivity confines the first singularity to the real axis, the functional equation pairs roots outside $|\alpha| = \sqrt{q}$ with roots inside, and the bound holds along the whole tower. These inputs do not prove the Riemann hypothesis (Remark 2.6.6).

THEOREM 2.6.2 (dichotomy). *Let $K/\mathbb{F}_q$ be a function field of genus $g \geq 1$ with reciprocal roots α_i, let $n_1 \geq 1$ and $A > 0$, let W be a set of positive integers containing arbitrarily long intervals, and suppose $N_n \leq q^n + 1 + Aq^{n/2}$ for all $n \in n_1 W$. Then exactly one of the following holds.*

(1) $|\alpha_i| = \sqrt{q}$ for all i.
(2) The largest modulus of an α_i exceeds $\sqrt{q}$; for some α_i of largest modulus the power $\alpha_i^{n_1}$ is real and positive; and this value has the largest multiplicity among the $\alpha_j^{n_1}$ with α_j of largest modulus.

PROOF. Let $\rho = \max_i |\alpha_i|$; in case (2) the power in question is ρ^{n_1}. The two cases exclude each other. For $k \in W$ put

$$a_k := Aq^{n_1 k/2} + \sum_i \alpha_i^{n_1 k} = Aq^{n_1 k/2} + q^{n_1 k} + 1 - N_{n_1 k} \geq 0.$$

Collecting equal values of $\alpha_i^{n_1}$, a_k is an exponential sum whose frequencies are the distinct $\beta = \alpha_i^{n_1}$, with coefficients the multiplicities $\mu_\beta \geq 1$, together with the frequency $q^{n_1/2}$ with a positive coefficient. If $\rho^{n_1} > q^{n_1/2}$, the frequencies of largest modulus are among the β, and Proposition 2.6.1 gives (2). If $\rho \leq \sqrt{q}$, the functional equation (Theorem 2.2.6(3)) gives $|\alpha_i| \geq \sqrt{q}$ for all i, which is (1). □

Theorem 2.6.2 leaves one alternative to the Riemann hypothesis, and it has the shape of a failure of positivity: a real reciprocal root of maximal modulus that the upper bound for N_n does not see. The next theorem identifies the missing input exactly.

THEOREM 2.6.3 (the exact input Route 2 needs). *Let $K/\mathbb{F}_q$ be a function field with reciprocal roots α_i and L-polynomial L_K. The following are equivalent.*

(1) $|\alpha_i| = \sqrt{q}$ for all i.
(2) $L_K(t) \neq 0$ for real $t \in (q^{-1}, q^{-1/2})$.

The proof that (2) implies (1) uses only Sections 2.2 and 2.3 and Proposition 2.6.1, not Section 2.4.

PROOF. That (1) implies (2) is clear. Assume (2); we may assume $g \geq 1$. By (2.3.8), $N_n \leq q^n + 1 + Aq^{n/2}$ for all $n \geq n_0$, with $A = (g+1)\sqrt{p} + 2g$ and n_0 least with $q^{n_0/2} \geq 2(g+1)$. Apply Theorem 2.6.2 with $n_1 = 1$ and $W = \{n \geq n_0\}$. In case (2), $\rho > \sqrt{q}$ is itself a reciprocal root. Also $\rho \leq q$: otherwise apply Proposition 2.6.1 to $N_n = q^n + 1 - \sum_i \alpha_i^n \geq 0$, whose frequencies of largest modulus would be reciprocal roots, with negative coefficients. And $\rho \neq q$ by Theorem 2.2.6(3). So $t = 1/\rho$ is a real zero of L_K in $(q^{-1}, q^{-1/2})$, contradicting (2). □

REMARK 2.6.4. By Theorem 2.6.3, what Route 2 lacks is the absence of real zeros of L_K in $(q^{-1}, q^{-1/2})$. This is a statement about one real interval, and with the upper bound it is equivalent to the Riemann hypothesis. So the one-sided bound excludes every violation of the Riemann hypothesis except

a real one. The proof of Proposition 1.3.4 has no step that uses $\zeta(\sigma) \neq 0$ for $0 < \sigma < 1$. There the bound is in the other direction, and the sign of its residue excludes a real zero. The counterpart of Proposition 1.3.4 is Theorem 2.6.9 below, which needs no input on real zeros. Theorem 2.6.3 is the counterpart of Landau's argument applied to an upper bound for $\mathrm{Ch}(x) - x$. For that bound the residue at a real zero β of ζ, which is m/β for a zero of multiplicity m, has the sign that the bound permits. So the absence of real zeros would have to be supplied separately. Over a function field the corresponding statement is the whole difficulty. The next proposition shows that formal properties do not give it.

PROPOSITION 2.6.5 (a formal counterexample). *Let $q = 9$ and $L^*(t) = 1 - 7t + 9t^2 = (1 - \rho t)(1 - \rho' t)$, with $\rho = (7 + \sqrt{13})/2$ and $\rho' = 9/\rho$, and put $Z^* = L^*/((1-t)(1-9t))$ and $N_n^* = 9^n + 1 - \rho^n - \rho'^n$. Then:*

(1) $Z^* = \prod_d (1 - t^d)^{-B_d^*}$ with integers B_d^*, and $N_n^* = \sum_{d|n} d B_d^*$;
(2) L^* satisfies the functional equation of genus one, and $L^*(1) = 3 = N_1^*$;
(3) $N_n^* \geq 0$ and $B_d^* \geq 0$ for all $n, d \geq 1$;
(4) $N_n^* \leq q^n + 1$ for all $n \geq 1$;
(5) the reciprocal roots $\rho = 5.30\ldots$, $\rho' = 1.69\ldots$ are real, not of modulus 3, and $L^*(1/\rho) = 0$ with $1/\rho \in (1/9, 1/3)$.

PROOF. Proof of (1): $L^* \in \mathbb{Z}[t]$ with $L^*(0) = 1$, and $(1-t)^{-1}(1-9t)^{-1} \in 1 + t\mathbb{Z}[[t]]$; an element of $1 + t\mathbb{Z}[[t]]$ has a unique expansion $\prod_d (1 - t^d)^{-B_d}$ with $B_d \in \mathbb{Z}$, and taking $t\frac{d}{dt}\log$ gives the formula for N_n^*. Proof of (2): the coefficients $(1, -7, 9)$ satisfy $a_2 = q\, a_0$, and $L^*(1) = 3$, $N_1^* = 10 - 7 = 3$. Proof of (3): $N_1^* = 3$, and for $n \geq 2$, $N_n^* \geq 9^n - 2\rho^n > 0$ since $\rho < 5.31$; also $B_1^* = N_1^* = 3$. Let $d \geq 2$ and assume $B_e^* \geq 0$ for $e < d$, so that $eB_e^* \leq N_e^*$. Then

$$dB_d^* = N_d^* - \sum_{e|d,\, e<d} eB_e^* \geq N_d^* - \sum_{e \leq d/2} N_e^* \geq 9^d - 2(5.31)^d - 2 \cdot 3^d > 0,$$

using $N_e^* \leq 9^e + 1$; this is an induction on d. Proof of (4): $\rho, \rho' > 0$. Part (5) is arithmetic. □

No curve has this zeta function: by Theorem 2.5.3 a curve of genus one over $\mathbb{F}_9$ has at least 4 points. The proposition shows that no argument using only rationality, the functional equation, integrality, nonnegativity and one-sided upper bounds for point counts can prove the Riemann hypothesis (Figure 2.17). Bombieri's example $\omega_1 = q$, $\omega_2 = 1$, with all point counts 0 [**Bom74**, p. 239], makes the same point; it has class number $L(1) = 0$, which Proposition 2.6.5 avoids.

REMARK 2.6.6. The one-sided bound $N_n \leq q^n + 1 + Aq^{n/2}$, Proposition 2.6.1, the functional equation and the tower of constant field extensions do not together prove the Riemann hypothesis without the averaging of Section 2.4. These four inputs prove only "the Riemann hypothesis or a real

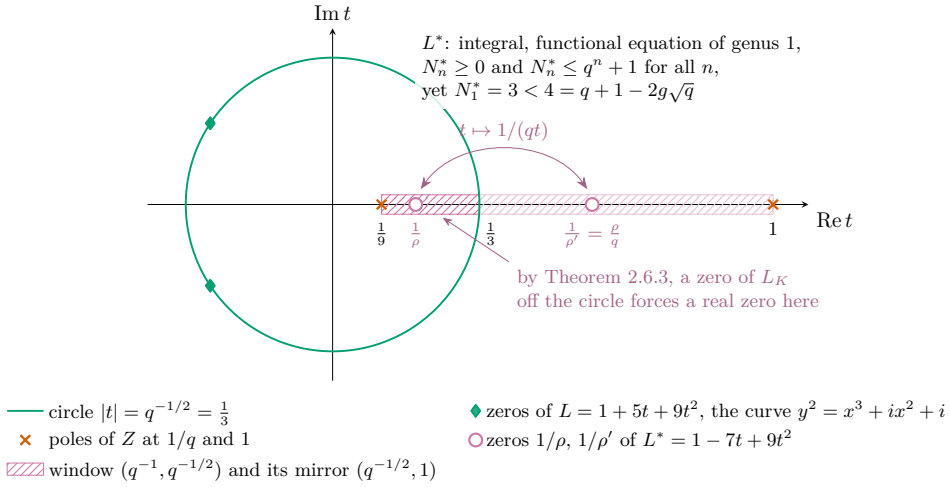


FIGURE 2.17. The window of Theorems 2.6.2 and 2.6.3 for $q = 9$, and the zeros of L^*. Green: the circle $|t| = q^{-1/2} = 1/3$ and the two zeros of $L = 1 + 5t + 9t^2$ for the curve $y^2 = x^3 + ix^2 + i$ over $\mathbb{F}_9 = \mathbb{F}_3[i]$, which has $N_1 = 15$ points (counted by brute force). Vermillion crosses: the poles of Z at $1/9$ and 1. Purple hatching: the interval $(q^{-1}, q^{-1/2})$, and lighter its image $(q^{-1/2}, 1)$ under $t \mapsto 1/(qt)$. By Theorem 2.6.3, L_K has all its zeros on the circle if and only if it has no real zero in $(q^{-1}, q^{-1/2})$; with the upper bound of Section 2.3, a zero off the circle forces one there (Theorem 2.6.2). Purple open circles: the zeros $1/\rho = 0.188\ldots$ and $1/\rho' = 0.589\ldots$ of $L^* = 1 - 7t + 9t^2$ (Proposition 2.6.5), one in each interval, swapped by $t \mapsto 1/(qt)$. L^* has every formal property used in Section 2.6, but $N_1^* = 3 < 4 = q + 1 - 2g\sqrt{q}$, so no curve has it as its L-polynomial.

reciprocal root in $(\sqrt{q}, q)$" (Theorems 2.6.2 and 2.6.3). The polynomial L^* of Proposition 2.6.5 has all of the formal properties those theorems use, and it realises the second case. The averaging of Section 2.4, or some other input specific to curves, is needed.

2.6.3. The centred smoothed sum. The number-field argument does not bound the Chebyshev function. It bounds the normalised smoothed sum $G_\zeta(x) = \sum_{n \leq x} \Lambda(n)(x/n - 1) - x \log x + (1 + \gamma)x$ of Definition 1.3.1 from above, and feeds the bound into a Landau theorem. Here is the exact

function-field counterpart, with $x = q^T$ and with N_n in the role of the sum of Λ over the elements of degree n.

DEFINITION 2.6.7. Let $K/\mathbb{F}_q$ be a function field of genus g with reciprocal roots α_i. For integers $n, T \geq 1$ and a complex number $\alpha \neq 1, q$ put

(1) $E_n = N_n - q^n - 1$ and $\mathcal{E}_n = \sum_{j \leq n} E_j$;
(2) $S(T) = \sum_{n=1}^{T} N_n(q^{T-n} - 1)$;
(3) $\gamma_K = \sum_i \frac{1}{\alpha_i - 1}$ and $\delta_K = \sum_i \frac{\alpha_i}{\alpha_i - 1}$;
(4) $c(\alpha) = \frac{(q-1)\alpha}{(q-\alpha)(\alpha-1)}$;
(5) $G(T) = S(T) - (T-1)(q^T - 1) + \gamma_K q^T + \delta_K$, the *centred smoothed sum* of K.

Here $\delta_K = \gamma_K + 2g = L_K'(1)/L_K(1)$, the number γ_K plays the role of Euler's constant, and $(T-1)(q^T-1)$ plays the role of $x \log x - x$. The numbers γ_K and δ_K are real, because the multiset of reciprocal roots is stable under complex conjugation.

PROPOSITION 2.6.8 (explicit formula). *Let $K/\mathbb{F}_q$ be a function field of genus g with reciprocal roots α_i and centred smoothed sum $G(T)$, and for complex $\alpha \neq 1, q$ let $c(\alpha) = \frac{(q-1)\alpha}{(q-\alpha)(\alpha-1)}$. For all $T \geq 1$,*

$$G(T) = \sum_{i=1}^{2g} c(\alpha_i)\, \alpha_i^T. \tag{2.6.1}$$

Also $c(\alpha) = c(q/\alpha)$, $c(\alpha) > 0$ for real $1 < \alpha < q$, and $c(\alpha) = (q-1)/|1-\alpha|^2$ if $|\alpha| = \sqrt{q}$.

PROOF. Sum geometric series: $\sum_{n=1}^{T}(q^n+1)(q^{T-n}-1) = (T-1)(q^T-1)$ and

$$\sum_{n=1}^{T} \alpha^n (q^{T-n} - 1) = \frac{\alpha}{q-\alpha}\, q^T + \frac{\alpha}{\alpha - 1} - c(\alpha)\, \alpha^T.$$

Sum the second identity over the α_i and use $\sum_i \frac{\alpha_i}{q-\alpha_i} = \gamma_K$, which holds because $\alpha \mapsto q/\alpha$ permutes the roots. The properties of c follow from $c(\alpha) = (q-1)/(q+1-\alpha-q/\alpha)$. □

THEOREM 2.6.9 (a one-sided bound on the smoothed sum suffices). *Let $K/\mathbb{F}_q$ be a function field with reciprocal roots α_i and centred smoothed sum $G(T)$, let $A > 0$, and let W be a set of positive integers containing arbitrarily long intervals. If $G(T) \leq Aq^{T/2}$ for all $T \in W$, then $|\alpha_i| = \sqrt{q}$ for all i.*

PROOF. Suppose $\rho = \max_i |\alpha_i| > \sqrt{q}$. Then $a_T = Aq^{T/2} - \sum_i c(\alpha_i)\alpha_i^T \geq 0$ on W. The frequencies of modulus ρ are reciprocal roots β, with coefficients $-\mu_\beta c(\beta) \neq 0$. If $\sqrt{q}$ is a reciprocal root, the coefficient $A - \mu_{\sqrt{q}}\, c(\sqrt{q})$ of the frequency $\sqrt{q}$ may vanish; in that case the frequency is omitted, which does not affect the frequencies of modulus $\rho > \sqrt{q}$. By Proposition 2.6.1, ρ is a reciprocal root and $-\mu_\rho c(\rho) > 0$, so $c(\rho) < 0$. Also $\rho \leq q$: otherwise

apply Proposition 2.6.1 to $N_n = q^n + 1 - \sum_i \alpha_i^n \geq 0$, whose frequencies of largest modulus would be reciprocal roots, with negative coefficients. And $\rho \neq q$ by Theorem 2.2.6(3). So $\sqrt{q} < \rho < q$, and $c(\rho) > 0$ by Proposition 2.6.8, a contradiction. Hence $\rho \leq \sqrt{q}$, and the functional equation gives equality. □

Applied to K_{n_1}, whose reciprocal roots are the $\alpha_i^{n_1}$ (Proposition 2.2.7), the theorem works along the tower. So the analytic step has an exact and valid function-field form: for the centred smoothed sum, real reciprocal roots count with a positive sign, and an upper bound alone suffices. It remains to say what that upper bound is in terms of point counts.

Section 2.3 also proves a one-sided upper bound. The sum $S(T)$ is a combination of the N_n with nonnegative weights $q^{T-n} - 1$, so an upper bound for each N_n gives an upper bound for $S(T)$, and $G(T)$ differs from $S(T)$ by terms that do not depend on the point counts beyond the constants γ_K, δ_K. The two upper bounds are nevertheless different statements, with opposite signs (Remark 2.6.11).

PROPOSITION 2.6.10 (which bound is asked for). *Let $K/\mathbb{F}_q$ be a function field with reciprocal roots α_i and centred smoothed sum $G(T)$, and put $E_n = N_n - q^n - 1$, $\mathcal{E}_n = \sum_{j \leq n} E_j$ and $\delta_K = \sum_i \frac{\alpha_i}{\alpha_i - 1}$. Assume $|\alpha_i| < q$ for all i. Then for all $T \geq 1$*

$$G(T) = \delta_K - \sum_{n \leq T} E_n - \sum_{n > T} E_n \, q^{T-n} = \delta_K - (1 - q^{-1}) \sum_{n \geq T} \mathcal{E}_n \, q^{T-n}. \tag{2.6.2}$$

The hypothesis $|\alpha_i| < q$ holds for every function field by Theorem 2.5.3, and it holds for the reciprocal roots ρ, ρ' of the polynomial L^* of Proposition 2.6.5.

PROOF. For $|\alpha| < q$,

$$\sum_{n \leq T} \alpha^n + \sum_{n > T} \alpha^n q^{T-n} = \frac{\alpha(\alpha^T - 1)}{\alpha - 1} + \frac{\alpha^{T+1}}{q - \alpha} = c(\alpha)\alpha^T - \frac{\alpha}{\alpha - 1}.$$

Summing over i and using $E_n = -\sum_i \alpha_i^n$ and (2.6.1) gives the first form. For the second, Abel summation gives $\sum_{n>T} (\mathcal{E}_n - \mathcal{E}_{n-1}) q^{T-n} = -q^{-1}\mathcal{E}_T + (1 - q^{-1}) \sum_{n>T} \mathcal{E}_n q^{T-n}$; adding $\sum_{n \leq T} E_n = \mathcal{E}_T$ gives $(1 - q^{-1})\mathcal{E}_T + (1 - q^{-1}) \sum_{n>T} \mathcal{E}_n q^{T-n}$. □

REMARK 2.6.11. A one-sided upper bound for $G(T)$ is the counterpart of the bound $G_\zeta(x) \leq x^{1/2+\varepsilon}$ of Chapter 1. It corresponds to a lower bound for the point counts N_n, and the Stepanov–Bombieri upper bound has the opposite sign. The upper bound for $S(T)$ obtained from upper bounds for the N_n is of order q^T. If $|\alpha_i| < q$ for all reciprocal roots α_i, then $\gamma_K = -\sum_{n \geq 1} E_n q^{-n}$, so the centring term $\gamma_K q^T$ cancels the whole series $\sum_{n \geq 1} E_n q^{T-n}$ that carries this order. What is left enters with the sign of $-E_n$, by the first form of (2.6.2). Let $A > 0$ and put $C_q = A\frac{\sqrt{q}+1}{\sqrt{q}-1}$. Then the following hold.

(1) If $E_n \le Aq^{n/2}$ for all n, which is the direction Section 2.3 proves, then $G(T) \ge \delta_K - C_q q^{T/2}$. This is a lower bound on G, and Proposition 2.6.1 cannot use it.
(2) If $E_n \ge -Aq^{n/2}$ for all n, a *lower* bound on point counts, then $G(T) \le \delta_K + C_q q^{T/2}$. This is the hypothesis of Theorem 2.6.9.
(3) A real reciprocal root $\alpha \in (\sqrt{q}, q)$ contributes $-\alpha^n < 0$ to E_n and $c(\alpha)\alpha^T > 0$ to $G(T)$. An upper bound on E_n does not see it; an upper bound on $G(T)$ does. For the polynomial L^* of Proposition 2.6.5, the right side of (2.6.1) with the reciprocal roots ρ, ρ' is $G^*(T) = \frac{8}{3}(\rho^T + \rho'^T)$.

Figure 2.18 shows the reversal.

REMARK 2.6.12. The same holds over $\mathbb{Q}$. Partial summation gives, with $\mathrm{Ch}(u) = \sum_{n \le u} \Lambda(n)$,

$$\sum_{n \le x} \Lambda(n)\Big(\frac{x}{n} - 1\Big) = x \int_1^x \frac{\mathrm{Ch}(u)}{u^2}\, du,$$

so that

$$G_\zeta(x) = x\Big(\int_1^x \frac{\mathrm{Ch}(u) - u}{u^2}\, du + 1 + \gamma\Big).$$

The classical estimate $\sum_{n \le x} \Lambda(n)/n = \log x - \gamma + o(1)$ gives $\int_1^\infty (\mathrm{Ch}(u) - u)u^{-2}\, du = -1 - \gamma$, hence

$$G_\zeta(x) = -x \int_x^\infty \frac{\mathrm{Ch}(u) - u}{u^2}\, du.$$

An upper bound on G_ζ is therefore a tail-averaged lower bound on $\mathrm{Ch}(u) - u$.

2.6.4. Summary of Route 2. With the bound that Section 2.3 provides, an upper bound on N_n, Route 2 proves "the Riemann hypothesis or a real zero of L_K in $(q^{-1}, q^{-1/2})$" and no more (Theorems 2.6.2 and 2.6.3, Proposition 2.6.5). With an upper bound on the centred smoothed sum it proves the Riemann hypothesis (Theorem 2.6.9), but that bound is a lower bound for point counts (Proposition 2.6.10). Route 1 supplies the lower bound, and with it both the condition of Theorem 2.6.3(2) and the hypothesis of Theorem 2.6.9. No Stepanov-type argument giving either directly is known to us. This leaves two questions.

QUESTION 2.6.13. Let $K/\mathbb{F}_q$ be a function field. Is there a Stepanov-type proof, without a Galois closure, that L_K has no real zero in $(q^{-1}, q^{-1/2})$? It would complete Route 2.

QUESTION 2.6.14. Let $K/\mathbb{F}_q$ be a function field with reciprocal roots α_i and centred smoothed sum $G(T)$, and let $\gamma_K = \sum_i \frac{1}{\alpha_i - 1}$. Is there a Stepanov-type upper bound for $G(T)$ with γ_K built in? It would be a function-field model of the number-field argument.

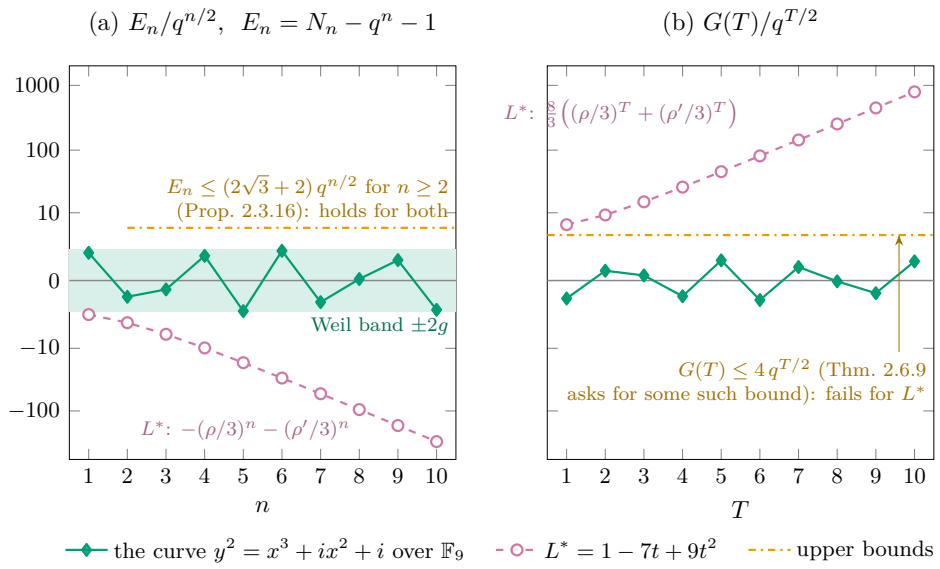


FIGURE 2.18. The sign reversal of Proposition 2.6.10 for $q = 9$, on a symmetric logarithmic scale ($v \mapsto \operatorname{sign}(v)\log_{10}(1 + |v|)$, ticks labelled by v; both panels share it). Green diamonds: the curve of Figure 2.17, of genus $g = 1$; purple dashed: L^* of Proposition 2.6.5, with $\rho = (7 + \sqrt{13})/2$ and $\rho' = 9/\rho$. (a) $E_n/q^{n/2}$: the curve stays in the Weil band $\pm 2g$ (green), and L^* falls like $-(\rho/3)^n$. The upper bound $E_n \le (2\sqrt{3} + 2)q^{n/2}$ of Proposition 2.3.16 for $n \ge 2$ (orange) holds for both, so it cannot tell them apart. (b) $G(T)/q^{T/2}$ (Definition 2.6.7): the curve stays bounded, and $G^*(T) = \frac{8}{3}(\rho^T + \rho'^T)$ grows. Theorem 2.6.9 asks for an upper bound $G(T) \le Aq^{T/2}$; every such bound fails for L^*. The orange level $A = 4 = 2g(\sqrt{q}+1)/(\sqrt{q}-1)$ is the one that the lower bound $E_n \ge -2gq^{n/2}$ gives through Proposition 2.6.10, up to the term $\delta_K q^{-T/2}$. Values computed from the explicit formula (2.6.1), which was checked against Definition 2.6.7 for the curve.

2.7. Stepanov's derivative form over a prime field

Bombieri notes that the constructions of Stepanov and Schmidt obtain vanishing of order at least m through derivatives or hyperderivatives of order up to $m - 1$, with m about $q^{1/2}$, whereas his own argument does not use derivations [**Bom74**, p. 235]. The derivative form is closer in structure to the chain of theta lifts of Chapter 7, so we prove it in the simplest case. The argument is a reconstruction of the standard one; agreement in detail with [**Ste69**] is not claimed.

In Bombieri's argument the multiplicity comes from one application of Frobenius, Step 3 of Theorem 2.3.8. The part of the number-field argument that produces multiplicity is the forced p-adic divisibility of periods along a chain of lifts (Section 1.6). Both make high-order jets vanish without paying for each jet separately. In structure, the derivative form of this section matches the chain more closely (Remark 2.7.7).

Throughout this section p is an odd prime, $f \in \mathbb{F}_p[X]$ is a polynomial of odd degree d with $3 \le d < p$, and

$$c = \frac{p-1}{2}, \qquad D = \frac{p-d}{2}, \qquad \text{so that} \qquad cd = \frac{d-1}{2}\,p + D.$$

In this section the letter ℓ denotes a positive integer, the multiplicity, and not the dimension of a Riemann–Roch space.

Definition 2.7.1. Let p be an odd prime, $f \in \mathbb{F}_p[X]$ and $c = (p-1)/2$.

(1) for $a = \pm 1$, $S_a = \{x \in \mathbb{F}_p : f(x)^c = a\}$, and $S_0 = \{x \in \mathbb{F}_p : f(x) = 0\}$;
(2) N^{aff} is the number of points of the affine curve $y^2 = f(x)$ over $\mathbb{F}_p$.

By Euler's criterion $N^{\mathrm{aff}} = 2|S_1| + |S_0|$, and $|S_1| + |S_{-1}| + |S_0| = p$.

Definition 2.7.2. Let p be an odd prime, $f \in \mathbb{F}_p[X]$ of odd degree d with $3 \le d < p$, $c = (p-1)/2$ and $D = (p-d)/2$, and let $J, \ell \ge 1$ be integers. An *auxiliary polynomial* is

$$r(X) = f(X)^{\ell} \sum_{j=0}^{J-1} \bigl(r_j(X) + s_j(X) f(X)^c\bigr) X^{jp}, \tag{2.7.1}$$

with $r_j, s_j \in \mathbb{F}_p[X]$ of degree at most $D-1$. Its *jet coefficients* are defined by $r_j^{[0]} = r_j$, $s_j^{[0]} = s_j$ and

$$r_j^{[k+1]} = (\ell - k) f' r_j^{[k]} + f\,(r_j^{[k]})', \qquad s_j^{[k+1]} = (\ell - k + c) f' s_j^{[k]} + f\,(s_j^{[k]})'.$$

Proposition 2.7.3. *Let $J, \ell \ge 1$ be integers and let r be an auxiliary polynomial* (2.7.1) *for f, J and ℓ, with coefficients r_j, s_j. If the r_j, s_j are not all zero, then $r \ne 0$.*

Proof. A nonzero term $r_j X^{jp}$ has degree $jp + \deg r_j$, with residue modulo p in $[0, D-1]$ and quotient j. A nonzero term $s_j f^c X^{jp}$ has degree $(j + \frac{d-1}{2})p + D + \deg s_j$, with residue in $[D, 2D-1]$ (note $2D - 1 < p$) and quotient $j + \frac{d-1}{2}$. Two distinct nonzero terms of the same kind differ in the quotient, and two of different kinds differ in the residue. So all nonzero terms have distinct degrees, their sum is nonzero, and so is its product with f^{ℓ}. □

This is the counterpart of Proposition 2.3.4: degrees, that is pole orders at infinity, are separated modulo p (Figure 2.19).

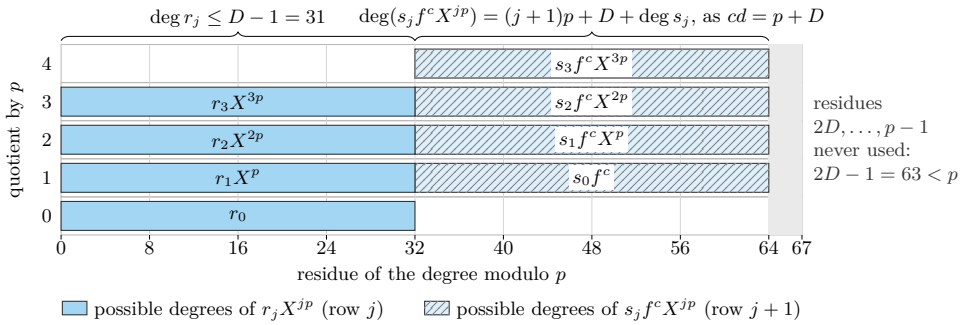


FIGURE 2.19. Why the auxiliary polynomial r of Definition 2.7.2 is nonzero (Proposition 2.7.3), for $p = 67$, $d = 3$, so $D = 32$, $c = 33$, $J = 4$. The degree of each nonzero term of $\sum_j (r_j + s_j f^c) X^{jp}$ is placed by its quotient (row) and residue (column) modulo p; it lies in the block shown. Terms $r_j X^{jp}$ (solid sky blue) lie in row j, residues 0 to $D-1$; terms $s_j f^c X^{jp}$ (hatched) lie in row $j+1$, residues D to $2D-1 < p$; residues $2D$ to $p-1$ (grey) are never used. Two terms of the same kind differ in the row, two of different kinds in the column, so distinct nonzero terms have distinct degrees and cannot cancel. Compare Figure 2.6: degrees at infinity play the role of pole orders at P_0.

PROPOSITION 2.7.4. *Let $J, \ell \ge 1$ be integers and let r be an auxiliary polynomial* (2.7.1) *for f, J and ℓ, with coefficients r_j, s_j and jet coefficients $r_j^{[k]}, s_j^{[k]}$. For $0 \le k \le \ell$ the k-th derivative of r is*

$$r^{(k)} = f^{\ell-k} \sum_j (r_j^{[k]} + s_j^{[k]} f^c) X^{jp}.$$

The polynomials $r_j^{[k]}, s_j^{[k]}$ have degree at most $D - 1 + k(d-1)$ and depend linearly on $(r_j, s_j)_j$.

PROOF. Differentiate $f^{\ell-k}(u + v f^c) X^{jp}$ and use $(X^{jp})' = 0$: the derivative of $f^{\ell-k} u$ is $f^{\ell-k-1}((\ell-k) f' u + f u')$, and that of $f^{\ell-k+c} v$ is $f^{\ell-k-1} f^c((\ell - k + c) f' v + f v')$. Each step raises the degree of a jet coefficient by at most $d - 1$. □

THEOREM 2.7.5. *Let p be an odd prime, let $f \in \mathbb{F}_p[X]$ be a polynomial of odd degree d with $3 \le d < p$, and put $c = (p-1)/2$ and $D = (p-d)/2$. Let ℓ be an even integer with $2 \le \ell \le p - d - 1$, and put*

$$t = \Big\lfloor \frac{\ell^2 d}{2(p - d - \ell)} \Big\rfloor + 1, \qquad B = d + \frac{(d+1)p - 2d - 2}{2\ell} + \frac{(t-1)p}{\ell}.$$

For $a = \pm 1$ let $S_a = \{x \in \mathbb{F}_p : f(x)^c = a\}$, and let N^{aff} be the number of points of the affine curve $y^2 = f(x)$ over $\mathbb{F}_p$. Then for both values of a

$$|S_a| \le \frac{p}{2} + B \le \frac{p}{2} + d + \frac{(d+1)p}{2\ell} + \frac{\ell d}{2} \cdot \frac{p}{p-d-\ell}, \tag{2.7.2}$$

and $|N^{\mathrm{aff}} - p| \le 2B + d$.

The sets S_a and the count N^{aff} are those of Definition 2.7.1.

PROOF. *Step 1: a Siegel lemma.* This step chooses the coefficients r_j, s_j of an auxiliary polynomial r for f, $J = \ell/2 + t$ and ℓ. Put $\sigma_k(X) = \sum_j (r_j^{[k]}(X) + a\, s_j^{[k]}(X))X^j$, of degree at most $D + J - 2 + k(d-1)$. Require $\sigma_k = 0$ as a polynomial for $0 \le k < \ell$. These are at most $\ell(D + J - 1) + \frac{\ell(\ell-1)}{2}(d-1)$ linear equations in $2JD$ unknowns. With $J = \ell/2 + t$ the number of unknowns minus the number of equations is

$$t(p - d - \ell) - \frac{\ell^2 d - \ell d - \ell}{2} > t(p - d - \ell) - \frac{\ell^2 d}{2} > 0$$

by the choice of t. So a nonzero solution exists.

Step 2: non-vanishing. $r \neq 0$ by Proposition 2.7.3.

Step 3: multiplicity. This step shows that r vanishes to order ℓ on S_a. Let $x \in S_a$. Then $x^{jp} = x^j$ and $f(x)^c = a$, so by Proposition 2.7.4, $r^{(k)}(x) = f(x)^{\ell-k}\sigma_k(x) = 0$ for $0 \le k < \ell$. Since $\ell < p$, x is a zero of r of order at least ℓ: if $(X - x)^e$ divided r exactly with $e < \ell$, then $r^{(e)}(x) = e! \cdot h(x) \neq 0$ for the cofactor h.

Step 4: degree count. We get $\ell|S_a| \le \deg r \le \ell d + (D-1) + cd + (J-1)p$. Substituting D, c and J gives the first inequality of (2.7.2); the second uses $t - 1 \le \frac{\ell^2 d}{2(p-d-\ell)}$.

For the point count, $N^{\mathrm{aff}} = 2|S_1| + |S_0| \le p + 2B + d$, and $N^{\mathrm{aff}} = 2p - 2|S_{-1}| - |S_0| \ge p - 2B - d$. □

The lower bound for N^{aff} came from the upper bound for $|S_{-1}|$, and $2|S_{-1}| + |S_0|$ is the affine point count of the quadratic twist $\epsilon y^2 = f(x)$, with ϵ a non-square. Together with the exact identity $|S_1| + |S_{-1}| + |S_0| = p$, this is Theorem 2.4.6 with $G = \mathbb{Z}/2$.

REMARK 2.7.6. In Step 4 the degree of r splits into three parts:

(1) an overhead $D - 1 + cd \approx \frac{(d+1)p}{2}$, independent of the chain length ℓ;
(2) a main term $(J-1)p \approx \frac{\ell}{2}p$, proportional to ℓ, which divided by ℓ is the true main term $p/2$;
(3) a correction $(t-1)p$, forced by the quadratic cost $\frac{\ell(\ell-1)}{2}(d-1)$ of the chain of ℓ jet conditions in Step 1; it needs $t \gtrsim \ell^2 d/(2p)$.

Per unit of multiplicity the error is $\frac{(d+1)p}{2\ell} + \frac{\ell d}{2}(1 + o(1))$, balanced at $\ell \asymp \sqrt{p}$, which gives $|S_a| \le p/2 + O(d\sqrt{p})$ (Figure 2.20). The constant is not Weil's. Minimising B over even ℓ, the ratio of $2B + d$ to Weil's $(d-1)\sqrt{p}$ is 2.78 for

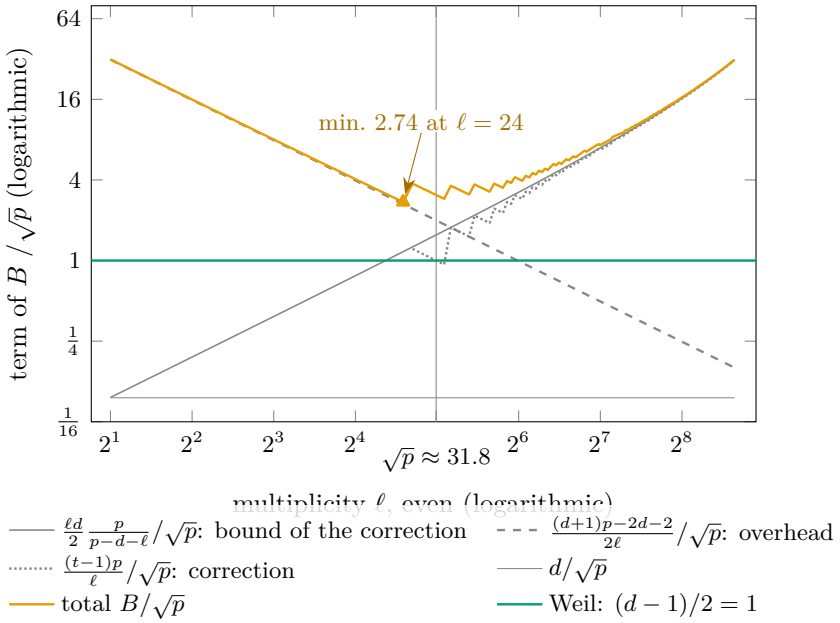


FIGURE 2.20. The balance of Remark 2.7.6 for $p = 1009$, $d = 3$: the terms of B in Theorem 2.7.5, $|S_a| \le p/2 + B$, divided by $\sqrt{p}$, against the even multiplicity $\ell \le 400$, on logarithmic axes. Grey dashed: the overhead $((d+1)p - 2d - 2)/(2\ell)$, the part of $\deg r$ that does not grow with ℓ, divided by ℓ; it falls. Grey dotted: the correction $(t-1)p/\ell$, forced by the quadratic number of jet conditions; it is 0 while $t = 1$ ($\ell \le 24$), jumps each time the integer t grows, and rises on average. Thin grey: its bound $\frac{\ell d}{2}\frac{p}{p-d-\ell}$ from (2.7.2), and the constant d. Orange: the total B, least, $2.74\sqrt{p}$, at $\ell = 24$ (triangle), near $\sqrt{p} \approx 31.8$. Green: $(d-1)\sqrt{p}/2$, half of Weil's bound $(d-1)\sqrt{p}$ for $|N^{\mathrm{aff}} - p|$, to compare with B, since $|N^{\mathrm{aff}} - p| \le 2B + d$. The layout matches Figure 2.10. Data: the closed formulas of (2.7.2).

$(p, d) = (1009, 3)$, at $\ell = 24$; 2.45 for $(1009, 5)$; 2.54 for $(10007, 3)$; 2.21 for $(10007, 7)$; and 2.26 for $(100003, 5)$.

REMARK 2.7.7. Bombieri's q_0-th power step, a single application of Frobenius, is not the structural counterpart of the multiplicity chain of Alpöge's argument. Theorem 2.7.5 matches that chain better. It has a chain of $\ell \approx \sqrt{p}$ jet conditions, an overhead independent of ℓ, a main term proportional to ℓ and a quadratic cost. In the number-field argument the errors of the first kind do not grow with the number m of lifts, the main term is a sum of m copies of an archimedean term, and the balance falls at $m \approx \sqrt{x}$ (Remark 9.4.3). Bombieri's argument has no chain length to

balance. This remark compares shapes and is not a theorem. Over general $\mathbb{F}_q$ the derivative form needs Hasse derivatives [**Sch76**].

2.8. The sign-only identity $qN = M^k + R$

We now turn to the second argument. Alpöge's note mentions a proof of the Weil and Deligne bounds over finite fields that "uses only a sign", written as $qN = M^k + R$ [**Alp26b**], and says that the plan for the integers was to combine "two quite different ideas present in a toy setting". We read the two ideas as Stepanov's multiplicity, which is Sections 2.3–2.7, and the sign-only proof. The note calls the second one "new and shocking" and its ideas "Fourier-analytic".

2.8.1. Where the identity appears. The identity appears twice in the note. In the introduction it is called the sign-only proof of Weil and Deligne. In the timeline, item 3 reads, in part, "$qN = M^k + R$, using $x_1 \cdot x_2 + \cdots + x_{k-1} \cdot x_k$". A footnote, attached to the phrase "the definite sign of the left-hand side", sketches the proof in one paragraph:

> sum over h first for the left-hand side, with e.g. quadric coefficients too Bertini gives a definite sign from middle cohomology, one off the one on M^k, which is the zero-mode on the right-hand side with R square-root-quality up to an inch or so as Katz would say by iterating Parseval k times, opposite signs meaning no cancellation of inverse roots, QED; vectorize to go up dimensions.

That paragraph is the source for Sections 2.8–2.10. The note defines none of t, M, N, and item 3 of the timeline has not been located in the Lean archive [**Alp26a**]. So the statements and proofs below are a reconstruction: the definitions are a reading of the footnote, the proofs are supplied here, and neither has been confirmed by the note's author. Where the reconstruction leaves something unproved, this is said (Section 2.10.5).

2.8.2. Two ways to bound a point count. Both arguments aim at the Riemann hypothesis for curves over $\mathbb{F}_q$, in the form $\left|\#C(\mathbb{F}_q) - q - 1\right| \leq 2g\sqrt{q}$, or equivalently that every eigenvalue α of Frobenius on $H^1(C)$ has $|\alpha| = \sqrt{q}$. The first proves it. The second derives the corresponding bound for character sums from two hypotheses, of which the second is not proved (Theorem 2.10.2). They proceed by different methods.

Stepanov's method is elementary: it uses the Riemann–Roch theorem and nothing else. Write m for the multiplicity, as in Chapter 1; it is the integer called M in Theorem 2.3.8. The steps are these.

(1) Choose m and build, by linear algebra in Riemann–Roch spaces, a nonzero function that vanishes to order at least m at every rational point. The construction uses the Frobenius $z \mapsto z^q$, which is the identity on residue fields of degree one, and the additivity of $z \mapsto$

z^{p^j}. Vanishing to order m comes from an m-th power, in place of m derivatives.

(2) A nonzero function has as many zeros as poles. Hence m times the number of rational points is at most the degree of the poles.

(3) The degree of the poles grows with m, but slower than m times q, as long as m is at most about $\sqrt{q}$. Dividing by m and taking $m \approx \sqrt{q}$ gives $\#C(\mathbb{F}_q) \le q + O(g\sqrt{q})$.

The lower bound follows by applying the upper bound to twists (Section 2.4), and the bounds over all $\mathbb{F}_{q^n}$ give $|\alpha| \le \sqrt{q}$ (Section 2.5). The method rests on *multiplicity*: one function is forced to vanish to high order at many points. More multiplicity gives a better bound, but costs degree. The balance point is $m \approx \sqrt{q}$.

The sign-only proof uses no auxiliary function and no multiplicity. It uses two inputs.

(a) An identity $qN = M^k + R$, where M is the sum to be bounded, N is a weighted point count on a quadric, and R is bounded by elementary Fourier analysis (Section 2.9).

(b) A *sign*: the Frobenius eigenvalues in qN and in M^k enter the identity with opposite signs, so they cannot cancel against each other (Section 2.10).

The sign is a statement about cohomology, which is why the proof is not elementary in the way Stepanov's is.

2.8.3. Notation. From here to the end of Section 2.12 the letters M, N, R, k and t have the meaning they have in the note's identity. They no longer denote the multiplicity, a point count of a function field, a radius and an exponent as in Sections 2.1–2.7; in particular M is now a complex number, and k is the number of variables of a quadratic form.

Let $q = p^f$. For $y \in \mathbb{F}_q$ set

$$e_q(y) \;=\; \exp\Big(\frac{2\pi i}{p}\,\operatorname{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(y)\Big),$$

where the trace $\operatorname{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(y) \in \mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ is lifted to $\{0,\dots,p-1\}$. Then $e_q\colon \mathbb{F}_q \to \mathbb{C}^\times$ is a nontrivial homomorphism from the additive group, and every additive character of $\mathbb{F}_q$ is $y \mapsto e_q(ay)$ for a unique $a \in \mathbb{F}_q$. The only fact used about e_q is orthogonality:

$$\sum_{h\in\mathbb{F}_q} e_q(hy) \;=\; \begin{cases} q, & y = 0,\\ 0, & y \neq 0. \end{cases} \tag{2.8.1}$$

For $q = p$ this is $e_p(y) = e^{2\pi i y/p}$, the classical case.

For $g\colon \mathbb{F}_q \to \mathbb{C}$ define the Fourier transform $\widehat{g}(\xi) = \sum_{x\in\mathbb{F}_q} g(x)\, e_q(x\xi)$ and the norm $\|g\|_2^2 = \sum_x |g(x)|^2$. By (2.8.1),

$$\|\widehat{g}\|_2^2 \;=\; q\,\|g\|_2^2 \qquad \text{(Parseval)}. \tag{2.8.2}$$

2.8.4. The function t. The note never says what t is. The reading adopted here is that t is a *trace function*: the function $x \mapsto \mathrm{Tr}(\mathrm{Fr}_x \mid \mathcal{F}_{\bar{x}})$ of a lisse ℓ-adic sheaf $\mathcal{F}$ on an open subset of the affine line, extended by zero, where Fr_x is the geometric Frobenius at x. There are two reasons.

(i) The note says the proof gives "Weil and Deligne". Weil's bound is the case $t(x) = \chi(f(x))$ or $t(x) = e_q(f(x))$, for a multiplicative character χ and a polynomial f. Deligne's bound is the case of a general pure sheaf.

(ii) The footnote speaks of "inverse roots" and "middle cohomology". Both are about the cohomology of a sheaf whose trace function is t.

Then $M = \sum_x t(x)$ is the character sum to be bounded. For example, for $t(x) = \chi(f(x))$ with χ quadratic, M is $\#C(\mathbb{F}_q) - q$ up to the points at infinity, where C is the curve $y^2 = f(x)$; with the sets $S_{\pm 1}$ and the count N^{aff} of Definition 2.7.1, $M = |S_1| - |S_{-1}| = N^{\mathrm{aff}} - p$ when $q = p$. In what follows only these properties of t are used: $|t| \le r$ for a constant r (the rank of $\mathcal{F}$), and t is the trace function of a sheaf to which the trace formula applies.

2.8.5. The identity.

Definition 2.8.1. Let $k \ge 2$ be an integer. The *path form* is

$$Q(x) \;=\; x_1x_2 + x_2x_3 + \cdots + x_{k-1}x_k,$$

the quadratic form of the path graph with vertices $1, \dots, k$ and edges $\{i, i+1\}$, and $Z \subset \mathbb{A}^k$ is the affine quadric $Q = 0$. For a function $t\colon \mathbb{F}_q \to \mathbb{C}$ put

$$S = \sum_{h \in \mathbb{F}_q} \sum_{x \in \mathbb{F}_q^k} t(x_1)\cdots t(x_k)\; e_q\big(h\,Q(x)\big), \qquad M = \sum_{x \in \mathbb{F}_q} t(x),$$

$$N = \sum_{x \in Z(\mathbb{F}_q)} t(x_1)\cdots t(x_k), \qquad R = \sum_{h \neq 0} \sum_{x \in \mathbb{F}_q^k} t(x_1)\cdots t(x_k)\, e_q\big(hQ(x)\big).$$

The note writes the form as $x_1 \cdot x_2 + \cdots + x_{k-1} \cdot x_k$, which can be read in two ways. The path form is the reading adopted here; the other reading is treated, and set aside, in Remark 2.9.2.

Proposition 2.8.2 (the identity). *Let $t\colon \mathbb{F}_q \to \mathbb{C}$ be a function, let $k \ge 2$, let Q be the path form in k variables, and let $Z \subset \mathbb{A}^k$ be the quadric $Q = 0$. Then*

$$\sum_{h \in \mathbb{F}_q} \sum_{x \in \mathbb{F}_q^k} t(x_1)\cdots t(x_k)\, e_q\big(hQ(x)\big) = q \sum_{x \in Z(\mathbb{F}_q)} t(x_1)\cdots t(x_k)$$

$$= \Big(\sum_{x \in \mathbb{F}_q} t(x)\Big)^k + \sum_{h \neq 0} \sum_{x \in \mathbb{F}_q^k} t(x_1)\cdots t(x_k)\, e_q\big(hQ(x)\big).$$

With the sums of Definition 2.8.1, the proposition reads $S = qN = M^k + R$.

PROOF. Summing over h first and using (2.8.1) with $y = Q(x)$ keeps only the x with $Q(x) = 0$, each with weight q. So $S = qN$. Splitting S by h, the term $h = 0$ is $\sum_x t(x_1)\cdots t(x_k) = M^k$ and the rest is R. □

So the five symbols are:

- q, the size of the field. It enters as the value of the h-sum in (2.8.1).
- N, the point count of the quadric Z, with each point weighted by $t(x_1)\cdots t(x_k)$.
- M, the sum whose bound is wanted. M^k is the *zero mode*: the frequency $h = 0$.
- k, the number of variables. It is a free parameter, and the proof sends it to infinity.
- R, the frequencies $h \neq 0$.

The identity itself is orthogonality of characters and nothing else; it holds for every function t. Its content is that the three terms can be controlled by three different means: N by the cohomology of a quadric, M^k by the cohomology of a product, and R by Fourier analysis.

2.9. Iterated Parseval: the bound on R

The note says that R is of "square-root quality" by "iterating Parseval k times". The following is one way to carry this out.

The Fourier transform is the one on the additive group of $\mathbb{F}_q$ defined in Section 2.8.3. It will be applied to a sequence of functions $g_1, g_2, \ldots$ on $\mathbb{F}_q$, built from t one variable at a time, starting with $g_1 = t$. This is possible because the phase $e_q(hQ(x))$ factors along the edges of the path:

$$e_q\big(hQ(x)\big) = e_q(hx_1x_2)\, e_q(hx_2x_3)\cdots e_q(hx_{k-1}x_k).$$

The variable x_1 occurs only in the first factor. So for fixed x_2,

$$\sum_{x_1} t(x_1)\, e_q(hx_1x_2) = \widehat{t}(hx_2),$$

which is the Fourier transform of t at the point hx_2.

2.9.1. The case $k = 3$. Here $Q = x_1x_2 + x_2x_3$, and x_1 and x_3 each meet only x_2. Summing over both of them,

$$\sum_{x_1,x_2,x_3} t(x_1)t(x_2)t(x_3)\, e_q\big(h(x_1x_2 + x_2x_3)\big) = \sum_{x_2} t(x_2)\, \widehat{t}(hx_2)^2.$$

If $|t| \le c$ and $h \neq 0$, then the triangle inequality and Parseval give

$$\Big|\sum_{x_2} t(x_2)\, \widehat{t}(hx_2)^2\Big| \le c \sum_{y} |\widehat{t}(y)|^2 = c\, q\, \|t\|_2^2 \le c^3q^2,$$

using that $y = hx_2$ runs over $\mathbb{F}_q$ once. The trivial bound is c^3q^3. So one power of q is saved, and no Gauss sum was evaluated. For general k one sums from one end and uses Parseval at each step.

2.9.2. The general case.

PROPOSITION 2.9.1 (iterated Parseval). *Let $t\colon \mathbb{F}_q \to \mathbb{C}$ be a function with $|t(x)| \le c$ for all x, let $k \ge 2$, let Q be the path form in k variables, and let $R = \sum_{h\neq 0} \sum_{x\in\mathbb{F}_q^k} t(x_1)\cdots t(x_k)\, e_q\big(hQ(x)\big)$. For each $h \in \mathbb{F}_q$ with $h \neq 0$,*

$$\Big|\sum_{x\in\mathbb{F}_q^k} t(x_1)\cdots t(x_k)\, e_q\big(hQ(x)\big)\Big| \;\le\; c^k\, q^{(k+1)/2}.$$

Hence $|R| \le c^k q^{(k+3)/2}$.

PROOF. Because Q is a path, the variables can be summed one at a time from one end. Put $g_1 = t$ and, for $2 \le j \le k$,

$$g_j(y) \;=\; t(y) \sum_{x\in\mathbb{F}_q} g_{j-1}(x)\, e_q(hxy) \;=\; t(y)\, \widehat{g_{j-1}}(hy).$$

Summing over x_1, then x_2, and so on up to x_{k-1} shows, by induction on j, that

$$\sum_{x_1,\dots,x_{j-1}} t(x_1)\cdots t(x_j)\, e_q\big(h(x_1x_2+\cdots+x_{j-1}x_j)\big) = g_j(x_j),$$

so the sum in the proposition is $\sum_y g_k(y)$. Since $h \neq 0$, the map $y \mapsto hy$ is a bijection of $\mathbb{F}_q$, so by (2.8.2)

$$\|g_j\|_2 \;\le\; c\,\|\widehat{g_{j-1}}\|_2 \;=\; c\sqrt{q}\,\|g_{j-1}\|_2.$$

Iterating from $\|g_1\|_2 \le c\sqrt{q}$ gives $\|g_{k-1}\|_2 \le c^{k-1}q^{(k-1)/2}$. Last, by Cauchy–Schwarz and Parseval once more,

$$\Big|\sum_y g_k(y)\Big| = \Big|\sum_y t(y)\,\widehat{g_{k-1}}(hy)\Big| \le \|t\|_2\,\|\widehat{g_{k-1}}\|_2$$
$$\le c\sqrt{q}\cdot\sqrt{q}\,c^{k-1}q^{(k-1)/2} = c^k q^{(k+1)/2}.$$

Summing over the $q-1$ values of h gives the bound for R. □

The trivial bound for R is $c^k q^{k+1}$, so the proposition saves a factor $q^{(k-1)/2}$. Compare it with M^k. If M has the square-root size $c\sqrt{q}$ that is to be proved, then $|M^k| \le c^k q^{k/2}$, and the bound on R exceeds this by the fixed factor $q^{3/2}$, whatever k is. In the reading adopted here, this fixed loss is the note's "up to an inch". The last step of the proof of Theorem 2.10.2 removes it by letting $k \to \infty$.

The argument depends on the shape of the path. For a general quadratic form one would diagonalise and evaluate Gauss sums. The note says "Gauss sum cancellation replaced by iterated Parseval": for the path, each variable meets only the next one, and Parseval suffices (Figure 2.21). The same proof works for the form $Q_a = \sum_i a_i x_i x_{i+1}$ with all $a_i \in \mathbb{F}_q^\times$, since it only needs $ha_i \neq 0$.

The note's form $x_1\cdot x_2 + \cdots + x_{k-1}\cdot x_k$ can also be read as the hyperbolic form $x_1x_2 + x_3x_4 + \cdots + x_{k-1}x_k$ with k even, in which every variable occurs once. That form does not give a remainder of square-root quality.

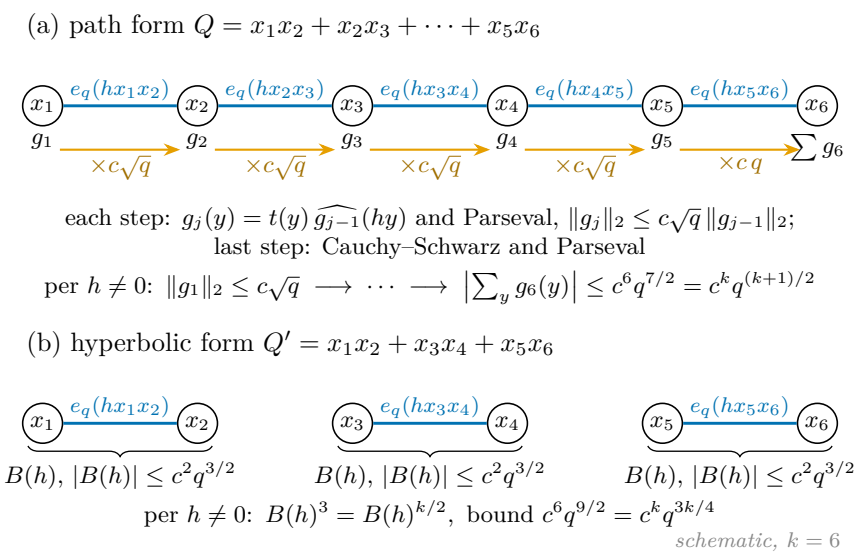


FIGURE 2.21. The two readings of the note's form for $k = 6$, drawn as graphs: each edge $\{i, j\}$ is a factor $e_q(hx_ix_j)$ of the phase $e_q(hQ(x))$. (a) The path form of Definition 2.8.1. Every vertex except the two ends lies on two edges, so the variables can be summed one at a time from one end; each summation is a Fourier transform, and Parseval costs a factor $c\sqrt{q}$ in $\|g_j\|_2$ (orange arrows). The last step is Cauchy–Schwarz and Parseval. This gives the bound $c^kq^{(k+1)/2}$ for each $h \neq 0$, and $|R| \le c^kq^{(k+3)/2}$ (Proposition 2.9.1). (b) The hyperbolic form of Remark 2.9.2. Every variable lies on one edge, the sum splits into $k/2$ equal factors $B(h)$, and in the bound $|R'| \le c^kq^{3k/4+1}$ the exponent cannot be improved. Against the size $c^kq^{k/2}$ of M^k if $|M| \le c\sqrt{q}$, the path loses the fixed factor $q^{3/2}$, the matching a factor of order $q^{k/4+1}$. Schematic.

REMARK 2.9.2 (the hyperbolic reading). Let $t\colon \mathbb{F}_q \to \mathbb{C}$ be a function, let k be even, let $Q' = x_1x_2 + x_3x_4 + \cdots + x_{k-1}x_k$, and let $R' = \sum_{h\neq 0}\sum_{x\in\mathbb{F}_q^k} t(x_1)\cdots t(x_k)\,e_q\big(hQ'(x)\big)$, the sum R of Definition 2.8.1 with Q' in place of Q. For $h \neq 0$ the sum over x is a product of $k/2$ equal factors:

$$\sum_{x\in\mathbb{F}_q^k} t(x_1)\cdots t(x_k)\,e_q\big(hQ'(x)\big) = B(h)^{k/2},$$

$$B(h) = \sum_{x,y\in\mathbb{F}_q} t(x)\,t(y)\,e_q(hxy) = \sum_{y\in\mathbb{F}_q} t(y)\,\widehat{t}(hy).$$

If $|t| \le c$, the Cauchy–Schwarz inequality and (2.8.2) give $|B(h)| \le \|t\|_2 \, \|\widehat{t}\|_2 \le c^2 q^{3/2}$, so

$$|R'| \;\le\; c^k \, q^{3k/4+1}.$$

The exponent cannot be improved. Let $t = \chi$ be a nontrivial multiplicative character of $\mathbb{F}_q$, extended by $\chi(0) = 0$, with Gauss sum $\tau(\chi) = \sum_x \chi(x) e_q(x)$, of absolute value $\sqrt{q}$. Then $\widehat{\chi}(\xi) = \overline{\chi}(\xi)\,\tau(\chi)$, so $B(h) = \overline{\chi}(h)\,\tau(\chi)\,(q-1)$. If $\chi^{k/2}$ is trivial, for instance if χ is quadratic and 4 divides k, then

$$R' = \tau(\chi)^{k/2}\,(q-1)^{k/2+1}, \qquad |R'| = q^{k/4}\,(q-1)^{k/2+1}.$$

So the remainder for Q' exceeds the size $c^k q^{k/2}$ of M^k by a factor of order $q^{k/4+1}$, which grows with k, where the path form loses the fixed factor $q^{3/2}$. If the proof of Theorem 2.10.2 is run with Q', under the hypothesis corresponding to (H2), its Step 2 has only the bound $r^k q^{n(3k/4+1)}$ over $\mathbb{F}_{q^n}$, and Step 5 gives $|\alpha_j| \le q^{3/4+1/k}$, in the limit only $|\alpha_j| \le q^{3/4}$. The path form is therefore the reading that agrees with the footnote's "square-root-quality" and "iterating Parseval k times": in it every variable except the two ends lies on two edges, and each summation is a Fourier transform.

2.10. The sign-only theorem

This section states the sign-only bound in full and proves it. The statement is a reading of the note's footnote, and the two hypotheses under which it is proved are assumptions, not results: the second of them is the content of the note's word "Bertini" and is not proved here.

2.10.1. Trace functions and Frobenius eigenvalues. Let $U \subseteq \mathbb{A}^1_{\mathbb{F}_q}$ be a nonempty open subset with $0 \notin U$, and let $j\colon U \to \mathbb{A}^1$ be the inclusion. Let $\mathcal{F}$ be a lisse $\overline{\mathbb{Q}}_\ell$-sheaf on U of rank r, for a prime $\ell \neq p$, which is pointwise pure of weight 0; and let t be its trace function, extended by zero to $\mathbb{F}_q$. So t is the trace function of $j_!\mathcal{F}$, and $|t(x)| \le r$ for all x, because the eigenvalues of Fr_x on the stalk have absolute value 1. For $n \ge 1$ let t_n be the trace function over $\mathbb{F}_{q^n}$, and let M_n, N_n, R_n be the sums of Definition 2.8.1 over $\mathbb{F}_{q^n}$, with t_n in place of t and e_{q^n} in place of e_q. Again $|t_n| \le r$.

Write $\overline{\mathbb{F}}_q$ for an algebraic closure of $\mathbb{F}_q$ and Fr for the geometric Frobenius of $\mathbb{F}_q$ acting on cohomology with compact supports. The sheaf $(j_!\mathcal{F})^{\boxtimes k}$ on $\mathbb{A}^k$ has trace function $x \mapsto t(x_1)\cdots t(x_k)$. By the Grothendieck–Lefschetz trace formula,

$$\begin{aligned} M_n &= \sum_i (-1)^i \,\mathrm{Tr}\big(\mathrm{Fr}^n \mid H^i_c(\mathbb{A}^1_{\overline{\mathbb{F}}_q}, j_!\mathcal{F})\big), \\ N_n &= \sum_i (-1)^i \,\mathrm{Tr}\big(\mathrm{Fr}^n \mid H^i_c(Z_{\overline{\mathbb{F}}_q}, (j_!\mathcal{F})^{\boxtimes k}|_Z)\big). \end{aligned} \tag{2.10.1}$$

The eigenvalues of Fr on these groups are the "inverse roots" of the note: the reciprocals of the zeros and poles of the corresponding L-functions. They

are the counterparts of the reciprocal roots α_i of Theorem 2.2.6. In the rest of this chapter $d = \dim H^1_c(\mathbb{A}^1_{\overline{\mathbb{F}}_q}, j_!\mathcal{F})$.

2.10.2. The two hypotheses. The proof assumes that the cohomology on each side lives in a single degree:

(H1) $H^i_c(\mathbb{A}^1_{\overline{\mathbb{F}}_q}, j_!\mathcal{F}) = 0$ for $i \neq 1$;

(H2) for infinitely many k, $H^i_c(Z_{\overline{\mathbb{F}}_q}, (j_!\mathcal{F})^{\boxtimes k}|_Z) = 0$ for $i \neq k-1$.

Hypothesis (H1) holds when $\mathcal{F}$ is geometrically irreducible and not geometrically constant: H^0_c vanishes because U is an affine curve and $\mathcal{F}$ is lisse, and H^2_c is, up to a twist, the coinvariants of the geometric monodromy. Hypothesis (H2) is the "definite sign from middle cohomology" of the note: Z has dimension $k-1$, and H^{k-1}_c is its middle degree. The note says that Bertini gives it, possibly after giving the quadric generic coefficients, $Q_a = \sum_i a_i x_i x_{i+1}$. Neither a proof of (H2) nor a reference for it is given here. Half of it is automatic (Remark 2.10.6), and the other half fails for some k in simple cases (Remark 2.10.7); see also Remark 2.10.8 and Section 2.10.5.

2.10.3. A growth lemma.

PROPOSITION 2.10.1. *Let $\gamma_1, \dots, \gamma_s \in \mathbb{C}$ be distinct, let $m_1, \dots, m_s \in \mathbb{C}$ be nonzero, and let $A, C > 0$. If $\left|\sum_i m_i \gamma_i^n\right| \leq A C^n$ for all $n \geq 1$, then $|\gamma_i| \leq C$ for all i.*

PROOF. For $|z|$ small,

$$\sum_{n\geq 1} \Big(\sum_i m_i \gamma_i^n\Big) z^n = \sum_i \frac{m_i \gamma_i z}{1 - \gamma_i z}.$$

By the hypothesis the left side converges for $|z| < 1/C$, so it defines an analytic function there, which agrees with the rational function on the right. The right side has a pole at each $1/\gamma_i$ with $\gamma_i \neq 0$, since the γ_i are distinct and $m_i \neq 0$. So every pole lies in $|z| \geq 1/C$, that is, $|\gamma_i| \leq C$. □

The proposition needs the m_i to be nonzero after the equal γ_i have been collected. A sum $\sum_i m_i \gamma_i^n$ in which large γ occur with coefficients of both signs can be small for every n, because the coefficients of one value of γ can add up to zero.

2.10.4. The theorem.

THEOREM 2.10.2 (Sign-only bound). *Let $U \subseteq \mathbb{A}^1_{\mathbb{F}_q}$ be a nonempty open subset with $0 \notin U$, with inclusion $j\colon U \to \mathbb{A}^1$. Let $\mathcal{F}$ be a lisse $\overline{\mathbb{Q}}_\ell$-sheaf on U of rank r, for a prime $\ell \neq p$, pointwise pure of weight 0, and let t be its trace function, extended by zero to $\mathbb{F}_q$. For $k \geq 2$ let Z be the affine quadric $x_1x_2 + \cdots + x_{k-1}x_k = 0$ in $\mathbb{A}^k$. Assume (H1) and (H2). Then every eigenvalue α of* Fr *on $H^1_c(\mathbb{A}^1_{\overline{\mathbb{F}}_q}, j_!\mathcal{F})$ satisfies $|\alpha| \leq \sqrt{q}$. Consequently*

$$\Big|\sum_{x\in\mathbb{F}_q} t(x)\Big| \;\leq\; \sqrt{q}\, \dim H^1_c(\mathbb{A}^1_{\overline{\mathbb{F}}_q}, j_!\mathcal{F}).$$

PROOF. Let $\alpha_1, \dots, \alpha_d$ be the eigenvalues of Fr on $H^1_c(\mathbb{A}^1_{\overline{\mathbb{F}}_q}, j_!\mathcal{F})$, and, for k allowed by (H2), let $\beta_1, \dots, \beta_e$ be those on $H^{k-1}_c(Z_{\overline{\mathbb{F}}_q}, (j_!\mathcal{F})^{\boxtimes k}|_Z)$, both with multiplicity. Fix such a k. The proof has five steps.

Step 1: the identity. Over $\mathbb{F}_{q^n}$, Proposition 2.8.2 gives $q^n N_n = M_n^k + R_n$, for every $n \geq 1$.

Step 2: the bound on R. Proposition 2.9.1 over $\mathbb{F}_{q^n}$, with $c = r$, gives $|R_n| \leq r^k q^{n(k+3)/2}$ for all $n \geq 1$.

Step 3: eigenvalues. By (2.10.1) and (H1), $M_n = -\sum_j \alpha_j^n$, so

$$M_n^k = (-1)^k \sum_I \alpha_I^n, \qquad \alpha_I = \alpha_{i_1} \cdots \alpha_{i_k},$$

the sum over all $I = (i_1, \dots, i_k) \in \{1, \dots, d\}^k$. By (2.10.1) and (H2), $N_n = (-1)^{k-1} \sum_j \beta_j^n$, so $q^n N_n = (-1)^{k-1} \sum_j (q\beta_j)^n$.

Step 4: the sign. The degrees k and $k-1$ differ by one. This is the note's "one off". Hence

$$R_n \;=\; q^n N_n - M_n^k \;=\; (-1)^{k-1}\Big(\sum_j (q\beta_j)^n + \sum_I \alpha_I^n\Big).$$

Every eigenvalue enters with the same sign. After grouping equal values, $R_n = (-1)^{k-1} \sum_\gamma m_\gamma \gamma^n$ with distinct complex numbers γ and integers $m_\gamma \geq 1$, the same for all n; and every α_I is one of the γ. No term cancels: this is the note's "opposite signs meaning no cancellation of inverse roots".

Step 5: comparison and $k \to \infty$. By Step 2 and Proposition 2.10.1, applied with $A = r^k$ and $C = q^{(k+3)/2}$, we get $|\gamma| \leq q^{(k+3)/2}$ for every γ that occurs in R_n. By Step 4, every α_I occurs. Take $I = (j, \dots, j)$:

$$|\alpha_j|^k \leq q^{(k+3)/2}, \qquad \text{so} \qquad |\alpha_j| \leq q^{1/2+3/(2k)}.$$

Let $k \to \infty$ through the values allowed by (H2). This gives $|\alpha_j| \leq \sqrt{q}$. Last, $\sum_x t(x) = M_1 = -\sum_j \alpha_j$ by (H1), so $|M_1| \leq d\sqrt{q}$. □

REMARK 2.10.3 (what the sign does). The positivity of the m_γ is what is used. In general a sum of eigenvalues of size q^k could cancel down to the size of R. The sign forbids this. Purity of $\mathcal{F}$ is used only to bound $|t_n|$ by r; no bound on the size of an eigenvalue on cohomology is assumed.

REMARK 2.10.4 (Weil and Deligne). For $t(x) = \chi(f(x))$ or $t(x) = e_q(f(x))$, with the corresponding Kummer or Artin–Schreier sheaf on the open set where it is lisse, the conclusion of Theorem 2.10.2 is the Weil bound for the character sum; for a general lisse sheaf, pure of weight 0, it is Deligne's bound $|\alpha| \leq \sqrt{q}$ for the eigenvalues on H^1_c. This is the reading of "Weil and Deligne" in the note. In each case the theorem gives the bound only if (H1) and (H2) hold for the sheaf in question, and for the plain path quadric (H2) can fail for particular k in exactly these cases (Remark 2.10.7).

REMARK 2.10.5 (the condition $0 \notin U$). The condition $0 \notin U$ is not used in the deduction of Theorem 2.10.2 from (H1) and (H2). It is there for the sake of (H2): the quadric Z is singular, and its singular points lie in

the hyperplane $x_2 = 0$, so the condition keeps them out of the support of the sheaf (Remark 2.10.6). It costs nothing. Removing the point 0 from U changes $\sum_x t(x)$ by one term of size at most r. It also does not lose eigenvalues: with $U' = U \smallsetminus \{0\}$ and $j' \colon U' \to \mathbb{A}^1$, the exact sequence $0 \to j'_!(\mathcal{F}|_{U'}) \to j_!\mathcal{F} \to \mathcal{F}_0 \to 0$, in which $\mathcal{F}_0$ is a skyscraper sheaf at the origin, shows that $H^1_c(\mathbb{A}^1_{\overline{\mathbb{F}}_q}, j_!\mathcal{F})$ is a quotient of $H^1_c(\mathbb{A}^1_{\overline{\mathbb{F}}_q}, j'_!(\mathcal{F}|_{U'}))$ and that (H1) for U implies (H1) for U'.

REMARK 2.10.6 (half of (H2) is automatic). The gradient of Q at x has first coordinate x_2, so every singular point of Z has $x_2 = 0$. Since $0 \notin U$, the variety $Z \cap U^k$ is therefore smooth, of dimension $k-1$ if it is not empty, and it is affine. The sheaf $(j_!\mathcal{F})^{\boxtimes k}|_Z$ is the extension by zero of a lisse sheaf on $Z \cap U^k$. By Artin's vanishing theorem and Poincaré duality, the cohomology with compact supports of a lisse sheaf on a smooth affine variety of dimension $k-1$ vanishes in degrees below $k-1$. So the part of (H2) with $i < k-1$ always holds, and (H2) is the statement that the cohomology vanishes in the degrees $k, \dots, 2k-2$ above the middle.

The note says "Bertini gives a definite sign from middle cohomology". For the path quadric as it stands, (H2) can fail even when $\mathcal{F}$ satisfies (H1).

REMARK 2.10.7 ((H2) fails for the plain path quadric at $k = 3$). For $k = 3$ the form is $Q = x_2(x_1 + x_3)$, and $t(0) = 0$, so only the points with $x_3 = -x_1$ contribute to N:

$$N = M \cdot \sum_{x \in \mathbb{F}_q} t(x)\, t(-x).$$

Take for $\mathcal{F}$ the Artin–Schreier sheaf on $U = \mathbb{A}^1 \smallsetminus \{0\}$ with trace function $t(x) = e_q(x)$ for $x \neq 0$. It has rank one and weight 0, and it satisfies (H1) with $d = 1$ and $\alpha_1 = 1$, since $M_n = \sum_{x \neq 0} e_{q^n}(x) = -1$. Here $t_n(x)t_n(-x) = 1$ for $x \neq 0$, so

$$N_n = -(q^n - 1) = 1^n - q^n \qquad \text{for all } n \geq 1.$$

The eigenvalues 1 and q occur with opposite signs, so the cohomology of Z is not concentrated in one degree, and (H2) fails for $k = 3$.

The general statement at $k = 3$ is as follows. Since $0 \notin U$, the map $(x_1, x_2) \mapsto (x_1, x_2, -x_1)$ identifies $(U \cap -U) \times U$ with $Z \cap U^3$, and the sheaf $(j_!\mathcal{F})^{\boxtimes 3}|_Z$ with the external product of $\mathcal{F} \otimes [-1]^*\mathcal{F}$ on $U \cap -U$ and $\mathcal{F}$ on U, where $[-1]$ is the map $x \mapsto -x$. By the Künneth formula and (H1),

$$H^i_c(Z_{\overline{\mathbb{F}}_q}, (j_!\mathcal{F})^{\boxtimes 3}|_Z) = H^{i-1}_c\big((U \cap -U)_{\overline{\mathbb{F}}_q}, \mathcal{F} \otimes [-1]^*\mathcal{F}\big) \otimes H^1_c\big(\mathbb{A}^1_{\overline{\mathbb{F}}_q}, j_!\mathcal{F}\big).$$

The first factor vanishes outside the degrees 1 and 2. So (H2) fails at $k = 3$ exactly when $d \geq 1$ and $H^2_c((U \cap -U)_{\overline{\mathbb{F}}_q}, \mathcal{F} \otimes [-1]^*\mathcal{F}) \neq 0$, that is, when $\mathcal{F} \otimes [-1]^*\mathcal{F}$ has nonzero coinvariants under the geometric monodromy group. For $\mathcal{F}$ geometrically irreducible with $d \geq 1$ this is the condition that the pullback of $\mathcal{F}$ by $x \mapsto -x$ be geometrically isomorphic to the dual of $\mathcal{F}$. It

holds for the Artin–Schreier sheaf above, and for $t = \chi \circ f$ with χ a quadratic character and f an even polynomial, where $\mathcal{F} \otimes [-1]^*\mathcal{F}$ is geometrically constant.

For the same Artin–Schreier sheaf, direct computation of N over the fields $\mathbb{F}_q$ with q odd and $q \leq 27$, which include $\mathbb{F}_9$, $\mathbb{F}_{25}$ and $\mathbb{F}_{27}$, gives $N = q - 2$ for $k = 4$, $N = 3$ for $k = 5$, $N = q^2 - 2q - 4$ for $k = 6$ and $N = -2q^2 + 5q + 5$ for $k = 7$. If these formulas hold over all finite fields of odd characteristic, then (H2) also fails for $k = 4, 6, 7$, where coefficients of both signs occur, while $k = 5$ passes this test. The same formulas do not hold in characteristic 2. For $q = 2, 4, 8, 16$ the computation gives $N = -1, -13, -61, -253$ at $k = 5$, which is $3 - q^2$, and $N = 0, -12, -276, -3108$ at $k = 6$, where the formula for odd q would give $-4, 4, 44, 220$. So in characteristic 2 the test also fails at $k = 5$, on evidence of the same kind. With the coefficients $a = (1, 2, 3, 5)$ the same computation gives $N = 1$, $-2 - p$ and $3 + 3p$ for $k = 3, 4, 5$ at $p = 11, 13, 17, 19$, of the single sign $(-1)^{k-1}$ that (H2) requires. These computations are evidence and prove nothing. Theorem 2.10.2 needs (H2) only for infinitely many k, so the failures do not make it vacuous; they show that the choice of k, or of the coefficients, cannot be left unexamined.

Remark 2.10.8 (Which Bertini). The note says only "Bertini". It cannot mean Poonen's Bertini theorem over finite fields [**Poo04**]. That theorem produces smooth hypersurface sections of large degree defined over $\mathbb{F}_q$, and here the hypersurface is a quadric of fixed degree. The reading adopted here is that it is a Bertini theorem over $\overline{\mathbb{F}}_q$ with sheaf coefficients, which would give a dense open set of coefficients a for which (H2) holds for Q_a. Such a set has points over $\mathbb{F}_{q^s}$ for large s, and passing from $\mathbb{F}_q$ to $\mathbb{F}_{q^s}$ costs nothing, since it replaces α_j by α_j^s. We do not know a theorem of this kind that applies.

The geometry of Z is not the obstruction. The quadric Z is a cone, singular at the origin, and the form Q is nondegenerate only for even k, since the adjacency matrix of the path has determinant 0 for odd k and $(-1)^{k/2}$ for even k. Neither fact matters. The singular locus of Z lies in $x_2 = 0$, outside the support of the sheaf, and $Z \cap U^k$ is smooth for every k (Remark 2.10.6). What can fail is the upper half of (H2). By Poincaré duality on $Z \cap U^k$ it says that the cohomology without supports of the dual lisse sheaf vanishes in degrees at most $k - 2$, and $Z \cap U^k$ is not proper: this is a condition on the behaviour of the sheaf along the boundary. For the plain path quadric (H2) fails in cases (Remark 2.10.7). Whether it holds for infinitely many k for a given $\mathcal{F}$, or for generic coefficients, is not known to us; what generic coefficients can and cannot do is the subject of Remark 2.10.9.

Remark 2.10.9 (the forms Q_a are one torus orbit). The forms $Q_a = \sum_i a_i x_i x_{i+1}$ make up the linear system spanned by the monomials $x_i x_{i+1}$. Its members with all $a_i \neq 0$ are a single orbit of the diagonal torus: the substitution $x_i = \lambda_i y_i$ with $\lambda_1 = 1$ and $\lambda_{i+1} = (a_i \lambda_i)^{-1}$ turns $Q_a(x)$ into

$Q(y)$. So the weighted count for Q_a and t is the weighted count for the plain path form Q in which t is replaced, in the i-th variable, by $x \mapsto t(\lambda_i x)$; in terms of sheaves, $\mathcal{F}$ is replaced in the i-th factor by its pullback under $x \mapsto \lambda_i x$. Changing the coefficients therefore moves the sheaf, not the quadric, and no member of the family is more general than another as a variety.

Generic coefficients do not help every sheaf. They change nothing for a trace function that is homogeneous, $t(\lambda x) = \chi(\lambda)\, t(x)$ for a multiplicative character χ: then the count for Q_a is $\chi(\lambda_1 \cdots \lambda_k)$ times the count for Q, over $\mathbb{F}_q$ and, with $\chi(\lambda_1 \cdots \lambda_k)^n$, over $\mathbb{F}_{q^n}$, so (H2) holds for Q_a exactly when it holds for Q. Such a t is a multiple of χ, the trace function of a Kummer sheaf on $\mathbb{A}^1 \smallsetminus \{0\}$, and for nontrivial χ its sum M is zero, so Theorem 2.10.2 has no content for it. The Artin–Schreier sheaf is not of this kind: its pullbacks under distinct dilations are not geometrically isomorphic, and the counts do change with a (Remark 2.10.7).

2.10.5. What is not checked. The proof of Theorem 2.10.2 is complete as a deduction from (H1) and (H2). The following is not established.

- Hypothesis (H2) is not proved here, and no reference for it is known to us. Its lower half is automatic (Remark 2.10.6); its upper half fails for the plain path quadric at $k = 3$ in simple cases, and on numerical evidence at other small k (Remark 2.10.7). The singularity of Z and the degeneracy of the path form for odd k are not the obstruction (Remark 2.10.8). The note suggests generic coefficients, $Q_a = \sum_i a_i x_i x_{i+1}$ with $a_i \neq 0$, which amounts to moving the sheaf by dilations (Remark 2.10.9). All of the proof above works for Q_a: Proposition 2.8.2 holds for any polynomial in place of Q, and Proposition 2.9.1 only needs $ha_i \neq 0$. If (H2) fails for every Q_a, Step 4 fails and the proof gives nothing.
- The note's "vectorize to go up dimensions" suggests that the same argument, with x_i in $\mathbb{F}_q^m$, gives Deligne's theorem in higher dimension. This has not been worked out.
- That this is the argument the note intends has not been confirmed.

The first item leaves the following question.

Question 2.10.10. For which sheaves $\mathcal{F}$ satisfying (H1), and which k, does Hypothesis (H2) hold for the path quadric, or for the quadric Q_a with suitable coefficients? By Remark 2.10.7 it does not hold for all k. An answer for infinitely many k would complete the sign-only proof.

2.11. Multiplicity versus sign

The two arguments of this chapter aim at the same bound from opposite ends, the second conditionally, and the number-field argument was assembled from ideas of both. This section compares them.

2.11.1. What the sign-only proof uses. The proof of Theorem 2.10.2 uses four inputs.

(1) The identity $qN = M^k + R$. This is orthogonality of e_q and nothing else.
(2) The bound on R. This is Parseval on $\mathbb{F}_q$ and the bound $|t| \leq r$. For $t = \chi \circ f$ or $t = e_q \circ f$ the bound is $|t| \leq 1$, which is trivial.
(3) The Grothendieck–Lefschetz trace formula, to write M_n and N_n as alternating sums of n-th powers of eigenvalues.
(4) Hypotheses (H1) and (H2): the cohomology of each side lives in a single degree.

It does not use any bound on the size of an eigenvalue. Hypotheses (H1) and (H2) say in which degrees cohomology can be nonzero. They say nothing about weights, and statements of this kind are proved by geometric means: Artin vanishing, Poincaré duality, and Lefschetz or Bertini arguments. So the proof is not circular.

2.11.2. Where the sign is. The "sign" is not the sign of a number, since N is complex. It is the parity of a cohomological degree (Figure 2.22):

- every eigenvalue in M_n^k comes from degree k of the k-fold product, with sign $(-1)^k$;
- every eigenvalue in $q^n N_n$ comes from degree $k-1$, with sign $(-1)^{k-1}$.

In $R_n = q^n N_n - M_n^k$ the minus sign flips the second, so all terms carry the same sign $(-1)^{k-1}$. Without this, the identity would say nothing. A sum of eigenvalues of size q^k could cancel against another sum down to the size of R_n. The single-degree hypothesis rules this out. Then the large eigenvalues α_I must appear in R_n, and R_n is small.

Compare Section 2.6. There too everything turned on a sign: the upper bound for point counts controls the sums $\sum_i \alpha_i^n$ from one side only, and a real root of maximal modulus escapes it (Theorem 2.6.2). In the sign-only proof the eigenvalues are controlled from one side as well, but Step 4 arranges that every eigenvalue enters with a positive integer coefficient, so that Proposition 2.10.1 applies and nothing can escape.

2.11.3. The role of k. The bound on R loses a fixed factor $q^{3/2}$, and the eigenvalue α_j appears as α_j^k. Taking k-th roots turns the fixed loss into $q^{3/(2k)}$, which tends to 1 (Figure 2.23). This is the tensor power trick, the same device as Proposition 2.5.2 in the first proof, where the fixed loss was the constant A and the parameter sent to infinity was the degree n of the constant field extension. Deligne's first proof of the Weil conjectures [**Del74**] also gains by passing to large tensor powers, and also rests on a positivity that prevents cancellation. There the positivity is that of the coefficients of an L-function of $\mathcal{F} \otimes \overline{\mathcal{F}}$. Here it is the parity of a degree. This comparison is ours and is not in the note.

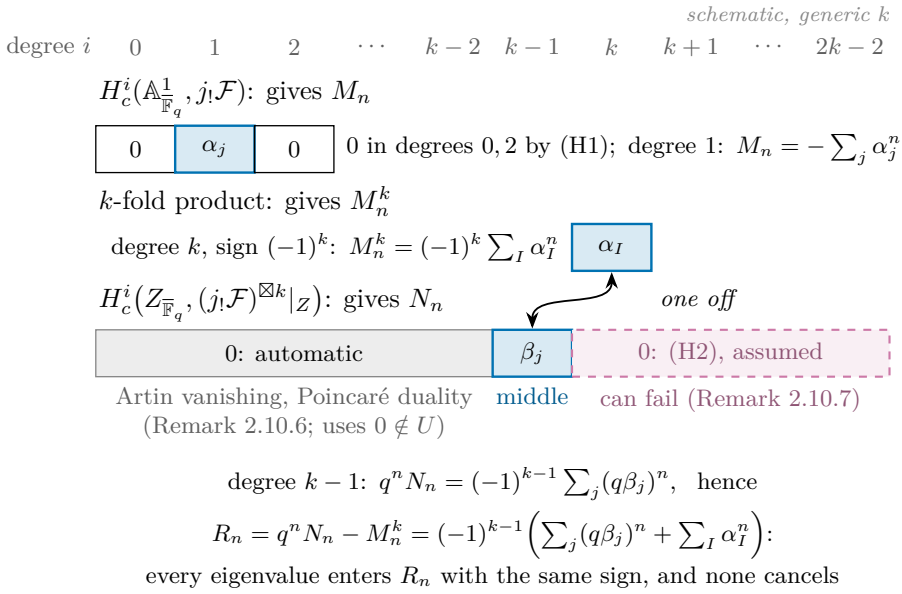


FIGURE 2.22. Where the sign is: the cohomological degrees in the proof of Theorem 2.10.2. Blue cells hold the eigenvalues of Fr that occur: the α_j on H^1_c of the curve, their products α_I in degree k of the k-fold product, and the β_j on the middle degree $k-1$ of the quadric Z, which has dimension $k-1$. The white cells of the curve, degrees 0 and 2, vanish by Hypothesis (H1). The grey cells, below the middle for Z, vanish because $0 \notin U$ (Remark 2.10.6). The dashed cells, the degrees $k, \dots, 2k-2$ above the middle, vanish only by Hypothesis (H2), which is assumed for infinitely many k, not proved, and fails for the plain path quadric at $k = 3$ in simple cases. The degrees k and $k-1$ differ by one, the note's "one off"; the minus sign in $R_n = q^n N_n - M_n^k$ then gives all eigenvalues the sign $(-1)^{k-1}$ (Steps 3 and 4 of the proof), so Proposition 2.10.1 applies. Schematic.

2.11.4. The comparison.

	Stepanov–Bombieri	Sign-only
Resource	multiplicity: one function vanishing to order m at every point	a sign: all eigenvalues enter R_n with one sign
Where $\sqrt{q}$ comes from	the balance between m and the degree it costs (Proposition 2.3.10)	the size of one Fourier sum, by Parseval
Fixed loss, and its removal	the constant $(2g+1)$, later A; removed by $n \to \infty$	the factor $q^{3/2}$; removed by $k \to \infty$
What is elementary	everything: Riemann–Roch only	the identity and the bound on R
What is not	nothing	the trace formula and (H1), (H2)
The other sign	from twists of a Galois cover (Section 2.4)	not needed

- *Where $\sqrt{q}$ comes from.* In Stepanov's method, $\sqrt{q}$ is the balance point between the multiplicity and the degree it costs. In the sign-only proof, $\sqrt{q}$ is the size of one Fourier sum, by Parseval.
- *What is elementary.* Stepanov's method needs only Riemann–Roch. The sign-only proof needs ℓ-adic cohomology for the trace formula and the vanishing hypotheses. Its Fourier analysis is elementary.
- *What carries over to $\mathbb{Q}$.* Neither proof carries over directly. There is no Frobenius on $\mathbb{Z}$ for Stepanov's m-th powers, and there is no cohomology for the sign. This is why the note looks for an automorphic source of the sign (Section 2.12), and why Chapter 1 replaces the m-th power by a chain of theta lifts (Section 1.6).

2.12. What the identity is for

This section reports what the note says about the role of the identity in the number-field argument, and what can be inferred from it. The inferences are marked. No fuller account is known to us, and item 3 of the timeline was not found in the Lean archive.

2.12.1. What "automorphy of N" means. The footnote is attached to this phrase: the definite sign of the left side of $qN = M^k + R$, "meaning holo/automorphy over a function field". Our reading is as follows. Over $\mathbb{F}_q$ the sign of qN comes from cohomology: Hypothesis (H2) puts all of N in one degree. Over $\mathbb{Q}$ there is no such cohomology. The idea was to find another source for the sign: a way to write N as an *automorphic* object, for instance a period integral of theta functions. Its sign would then come from holomorphy or automorphy, and not from the degree of a cohomology group.

Item 3 of the timeline sketches how. It says that μ is automorphic, coming from a theta function on a sixfold cover of GL_5, as a period integral of thetas. The note does not say what μ is. Item 3 also says that the quadric has a natural meaning on a CM elliptic curve such as $E = \mathbb{C}/\mathbb{Z}[\omega]$, and that one should "put Poincaré bundle on the edges to get θs". Our reading of this sentence:

- In the finite-field sum, the edge $\{i, i+1\}$ contributes the bilinear character $e_q(hx_ix_{i+1})$.
- On $E \times E$ the analogue of a bilinear pairing is the Poincaré line bundle P. Its Riemann form is a bilinear form.
- Pulling P back along $(x_i, x_{i+1})\colon E^k \to E^2$ for each edge and tensoring gives a line bundle on E^k attached to the path. Its sections are theta functions, which are automorphic.
- $\#\mu_K = 6$ for $K = \mathbb{Q}(\omega)$, which would explain "sixfold". Why the group is GL_5 is not explained.

2.12.2. What $O(2, n)$ is. The note says the decisive observation was "multiplicity 2 on $O(2, n)$" via a "natively-automorphic determinant", found by analogy with the sign of qN. The note does not define $O(2, n)$. In standard usage it is the orthogonal group of a real quadratic space of signature $(2, n)$, or a rational form of it. It is the orthogonal group whose symmetric space is Hermitian (a type IV domain), so it has holomorphic modular forms. It is the target of the classical theta lifts from SL_2, the Borcherds and Kudla–Millson lifts. Which quadratic space, which n, and which determinant are meant is not said.

Our reading: "multiplicity" is used as in Stepanov's method, as an order of vanishing or of divisibility. "Multiplicity 2" on $O(2, n)$ was the first case beyond multiplicity one. The note says that more theta lifts then gave more multiplicity.

2.12.3. Is there an N, M, k and R in the final argument? We have found no evidence that the identity $qN = M^k + R$ has a counterpart in the argument of Chapters 3–12. The note does not mention the identity after item 3. What it gives instead is the following.

- The quantity to bound is a smoothed prime sum. In the dictionary of Section 1.2, the function G_ζ of Definition 1.3.1 plays the role of a tail-smoothed sum of the numbers $q^n + 1 - \#C(\mathbb{F}_{q^n})$, with the sign reversed against the error of a point count (Proposition 2.6.10). So the point count of the curve is replaced by a weighted count of primes.
- The sign is that of the Guinand–Weil explicit formula. Item 4 of the timeline is "$W(1_S) \geq 0$ via a determinant", read there as Stepanov with multiplicity one. In the book the sign appears as the positivity of the weights in the windowed Landau criterion (Remark 4.3.7).

- Multiplicity comes from a chain of m theta lifts between definite unitary groups, with m of order $\sqrt{X}$ for a prime sum of length X. This is the analogue of $m \approx \sqrt{q}$ in Stepanov's method.

So the sign-only identity seems to have survived only as an analogy. Its definite sign became the target, a "natively automorphic determinant". One could match $k \to \infty$ with the length of the chain of lifts, since in both cases a parameter is sent to infinity to remove a fixed loss. Nothing in the note supports this match.

2.12.4. Where the quadric came from. The footnote says the quadric arose from simplifying an average over the family $\chi \mapsto \chi(\mathrm{disc})$. It names a normalized Gauss sum y of order $\#\mu_K$, which lives via θ on the $\#\mu_K$-fold cover of GL_2. Here is our reading, from what is known about these Gauss sums; the account of the literature follows the survey in [**BH13**, §1].

- Let K be a number field containing the group μ_K of roots of unity, and let $n = \#\mu_K$. For a modulus c there is an n-th order Gauss sum $g(c)$, built from the n-th power residue symbol modulo c. Let $y(c) = g(c)/|g(c)|$.
- The Fourier coefficients of the Eisenstein series on the n-fold cover of GL_2 are Dirichlet series built from Gauss sums of order n, and the theta function on that cover, introduced by Kubota, is a residue of this Eisenstein series [**BH13**, §1, pp. 3–4]. The footnote's wording suggests that the $y(c)$ are the Fourier coefficients of the theta function for every n. This is known only for $n = 3$: Patterson determined the coefficients of the cubic theta function, and at squarefree c they are the normalized cubic Gauss sums [**Pat77**]; see [**BH13**, §1, p. 4]. Kazhdan and Patterson [**KP84**] proved a periodicity property and Hecke relations for every n, and these determine all the coefficients for $n = 3$ but only some of them for $n \geq 4$ [**BH13**, §1, pp. 5–6]. For $n = 4$ the squares of the coefficients are predicted by a conjecture of Patterson. For $n = 6$, which is the case $\#\mu_K = 6$ of the note, a weaker conjecture of Chinta, Friedberg and Hoffstein [**CFH12**] relates the coefficients at squares to cubic Gauss sums, and does not determine all of them [**BH13**, §1, pp. 6–8]; see also [**Pat14**].
- y is not multiplicative. For coprime c, c' one has $y(cc') = \varepsilon(c, c')\, y(c)\, y(c')$, where $\varepsilon(c, c')$ is a power residue symbol. By the reciprocity law, ε behaves like a bilinear pairing.
- For a product of k coprime moduli this gives

$$y(c_1 \cdots c_k) \;=\; \prod_i y(c_i) \prod_{i<j} \varepsilon(c_i, c_j),$$

  with one factor for each pair $\{i, j\}$. That is the complete graph on k vertices.

- In a finite-field model a bilinear pairing becomes $e_q(x_i x_j)$. The complete graph gives $e_q(\sum_{i<j} x_i x_j)$. Keeping only the edges $\{i, i+1\}$ gives the path quadric.

The phrases "$\chi \mapsto \chi(\mathrm{disc})$" and "$a \mapsto y(\cdot)^a$" could not be reconstructed more precisely.

The footnote also says that Gabber's hypersurface in one of Katz's arguments "reminds one of the quadric". Katz [**Kat15**] uses the hypersurface $X_1^d + \sum_i X_i X_{i+1}^{d-1} = 0$ with $p \mid d$, due to Gabber, which he treats by Delsarte's method and Gauss sums. Its monomials $X_i X_{i+1}^{d-1}$ link each variable to the next, as the path does. This link is an inference.

2.13. Dictionary with Alpöge's argument

This section matches the steps of the Stepanov–Bombieri proof with the parts of the number-field argument. The argument is taken in five components, each with the place where the book develops it.

(i) A one-sided bound $G(x) \le x^{1/2+O(\varepsilon)}$ on windows for a positive combination G of smoothed Chebyshev functions (Theorem 9.4.1), fed into a windowed Landau theorem (Theorem 4.3.3).

(ii) A set of characters closed under inversion (Definition 4.4.1), so that every zero counts with the same sign (Remark 4.3.7).

(iii) Hermitian lattices of integral automorphic forms, with a slope inequality for the period maps between them in the role of "zeros at most poles" (Theorems 7.2.5 and 7.4.2; [**Bos96**]). No small sections are constructed (Remark 7.2.6).

(iv) Forced p-adic divisibility, then a chain of $m \approx \sqrt{x}$ theta lifts, with an error independent of m against a main term proportional to m (Chapter 7, Remark 9.4.3).

(v) Cyclotomic towers, which the note's outline places in the role of $\mathbb{F}_{q^n}/\mathbb{F}_q$. In the book cyclotomic fields enter in the construction of hosts (Chapter 5) and layers of a cyclotomic tower in the Tate–Shafarevich bound (Chapter 10); that they play the role of the tower of constant field extensions is not established here.

The ratings in the table are judgments: "exact" is the same statement under the usual dictionary, "loose" the same role by another mechanism, "none" no counterpart or a failing one. Each row names the statement of Chapters 3–12 against which it was rated, and the table agrees with Table 3.2 of Chapter 3. Figure 2.24 draws the table as a graph.

Step in this chapter	Component	Rating
Section w (Theorem 2.3.8); pole bound, degree	(iii) no constructed section; the period map between two given lattices, its height and the slopes (Remark 7.2.6)	loose

Step in this chapter	Component	Rating
Step 1: Riemann–Roch and pigeonhole	(iii) no Siegel lemma; arithmetic Hilbert–Samuel enters only inside the covolume bound (Theorem 8.2.3)	none
Step 4: effective divisors have nonnegative degree	(iii) slope inequality (Theorem 7.2.5), one step of the frame theorem and not used in rank one (Remark 7.4.3)	loose
Step 3: multiplicity from $w^M = \sum u_i^M v_i^q$, a ring endomorphism	(iv) forced p-adic divisibility along the chain of lifts (Proposition 9.3.1)	loose
Section 2.7: $\ell \approx \sqrt{p}$ derivative conditions	(iv) chain of $m \approx \sqrt{x}$ lifts (Remark 9.4.3)	loose
Balance $(g+1)q/M + 2gM$; Remark 2.7.6	(iv) errors x/m and m (Proposition 9.3.1)	loose
Non-vanishing (Propositions 2.3.4, 2.7.3)	not identified; no section is constructed (Remark 7.2.6)	none
Proposition 2.6.1, a finite case	(i) windowed Landau theorem (Theorem 4.3.3)	exact
Theorem 2.6.2; fails by Proposition 2.6.5	(i) "the upper bound alone gives RH"; the bound of Theorem 9.4.1 is in the other direction	none
$G(T)$, γ_K, δ_K; Theorem 2.6.9	(i) G_ζ, G_χ, $1+\gamma$, $L'/L(1,\chi)$ (Definition 4.2.1)	exact
Route 1, which also needs a base with $L = 1$	(ii) characters closed under inversion (Definition 4.4.1); Fourier analysis on a Galois group (Theorem 6.3.4)	loose
All roots of L_E enter $N(E, \mathrm{id})$ with one sign	(ii) every zero counts with one sign (Remark 4.3.7)	exact
The tower K_n (Propositions 2.2.7, 2.5.2)	(v) cyclotomic tower, role not established (Chapters 5, 10)	loose
Functional equation, Theorem 2.5.3	functional equation (Theorem 4.5.1)	exact

The sign-only proof has no row in this table. As Section 2.12 explains, the identity $qN = M^k + R$ is not known to have a counterpart in the final argument; what survives of it is the demand for a definite sign, which is component (ii).

2.13.1. Where the analogy breaks.

(B1) *Frobenius.* Step 3 of Theorem 2.3.8 uses that $z \mapsto z^M$ is additive; $\mathbb{Z}$ has no such endomorphism. Vanishing to order M at N points costs $\ell(MaP_0 + b\sigma(P_0))$ conditions, not about MN as in characteristic zero. Section 2.7 still uses $(X^p)' = 0$.

(B2) *Sign.* The bound that the number-field argument proves is, in function-field terms, a lower bound on point counts (Proposition 2.6.10), not the upper bound proved in Section 2.3. Over $\mathbb{Q}$ the corresponding statement is Remark 2.6.12.

(B3) *The base.* Route 1 needs a base with L-polynomial 1; ζ has zeros. This is consistent with the use of a Landau criterion over $\mathbb{Q}$ in place of an averaging argument.

(B4) *Finiteness.* Here there are $2g$ roots and (2.6.1) is exact; not so over $\mathbb{Q}$, where the explicit formula is an infinite sum (Remark 4.2.3).

(B5) *The centring constant.* Theorem 2.6.9 needs γ_K exactly, which Section 2.3 never produces. In the number-field argument the corresponding constants enter through a period of Chowla–Selberg type (Remark 7.5.2).

A faithful function-field model of the number-field argument would add to Theorem 2.6.9 a Stepanov-type upper bound for $G(T)$ (Question 2.6.14). Nothing in this chapter confirms or refutes Alpöge's claim.

2.14. Remarks and questions

2.14.1. The positivity that Weil uses and this proof avoids. Weil's proofs use positivity on a surface or on the Jacobian: the Castelnuovo–Severi inequality, a form of the Hodge index theorem, on $C \times C$ applied to the diagonal and the graph of Frobenius; or the positivity of the Rosati involution on $\mathrm{End}(J) \otimes \mathbb{Q}$ [**Wei48**]. Either gives both signs of $|N - q - 1| \leq 2g\sqrt{q}$ at once. Figure 2.25 contrasts the two pictures.

The Stepanov–Bombieri proof uses, on the curve itself, only Riemann–Roch and the nonnegativity of the degree of an effective divisor. This positivity is one-sided. The lower bound comes from averaging (Section 2.4), not from a second positivity. In explicit-formula terms the proof establishes an inequality for the one-parameter family of counts N_n only; it does not construct a positive form. The sharp constant $2g$ is not produced by the auxiliary section: Theorem 2.3.12 gives $2g+1$, and Proposition 2.3.10 shows that the method cannot do much better. The constant $2g$ appears only after Proposition 2.5.2.

2.14.2. Where the characteristic is used. Characteristic p enters the first proof in three places: Proposition 2.1.3(1) in (2.3.3), the only use in the upper bound; Proposition 2.1.3(3), which defines the points counted; and, in Section 2.7, $(X^{jp})' = 0$ and $x^{jp} = x^j$. Step 1 and Step 4 of Theorem 2.3.8, Proposition 2.3.4, Section 2.2, Section 2.4 apart from the definition of Frobenius places, and Propositions 2.5.2 and 2.6.1 do not use the characteristic. Stepanov treated $y^2 = f(x)$ over prime fields and W. Schmidt the general case [**Bom74**, p. 234], [**Sch73**].

2.14.3. Number fields. The following are heuristic remarks, not theorems.

(a) No endomorphism of $\mathbb{Z}$ is known that plays the role of Frobenius, and no object $\operatorname{Spec}\mathbb{Z} \times \operatorname{Spec}\mathbb{Z}$ over a base on which a graph of Frobenius could live. Item (B1) of Section 2.13 is the concrete form of this.
(b) Route 1 has no base with trivial L-function.
(c) The analytic step does transpose: a one-sided bound on a centred smoothed prime sum, with Landau's theorem, is a sound criterion (Theorem 2.6.9, Remark 2.6.12, Proposition 1.3.4). The open matter over $\mathbb{Q}$ is the bound itself.
(d) Stepanov's method has the form of the auxiliary-function arguments of Diophantine approximation (Thue, Siegel). This is a resemblance of form, not a transfer.

2.14.4. Questions. The questions that this chapter leaves open are stated where they arise. Questions 2.6.13 and 2.6.14 close the summary of Route 2 in Section 2.6, and Question 2.10.10 follows the list of Section 2.10.5.

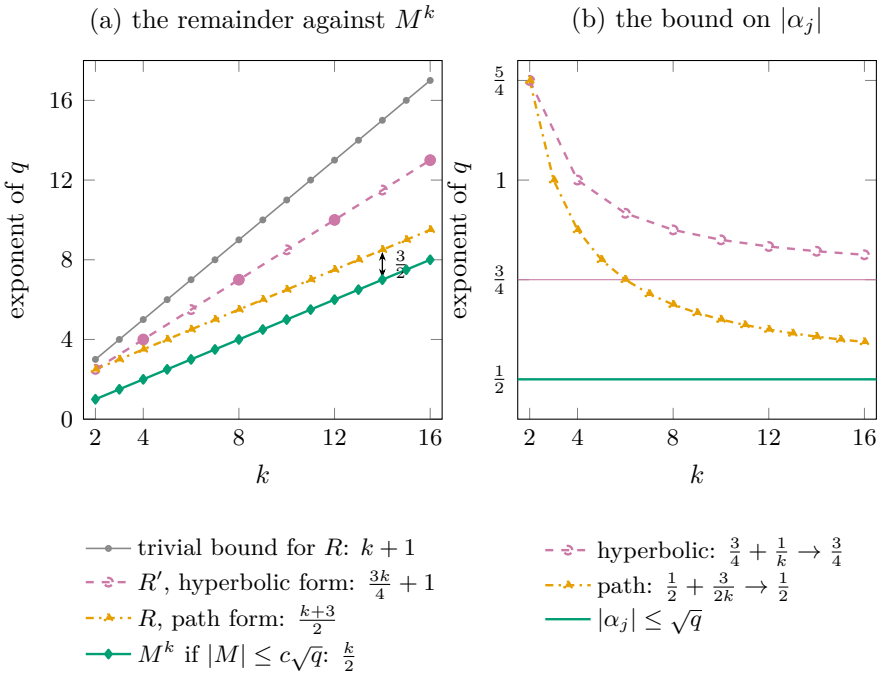


FIGURE 2.23. The role of k. Exponents of q, with the factors c^k suppressed, against the number k of variables. (a) Bounds for the remainder: the trivial bound $k+1$, the bound $(k+3)/2$ for the path form (Proposition 2.9.1), and the bound $3k/4+1$ for the hyperbolic form, k even (Remark 2.9.2); the filled circles, $4 \mid k$, mark where the latter is attained up to constants, for a quadratic character. Green: the size $k/2$ of M^k if M has the square-root size that is to be proved. The path bound exceeds it by the fixed $3/2$; the hyperbolic bound by $k/4+1$. (b) The bounds on $|\alpha_j|$ that Step 5 of the proof of Theorem 2.10.2 gives from them, as exponents of q: $1/2+3/(2k)$ for the path form and, if the proof is run with Q' under the hypothesis corresponding to (H2), $3/4+1/k$ for the hyperbolic form. The theorem uses only the k allowed by (H2), which is assumed, and lets $k \to \infty$ through them.

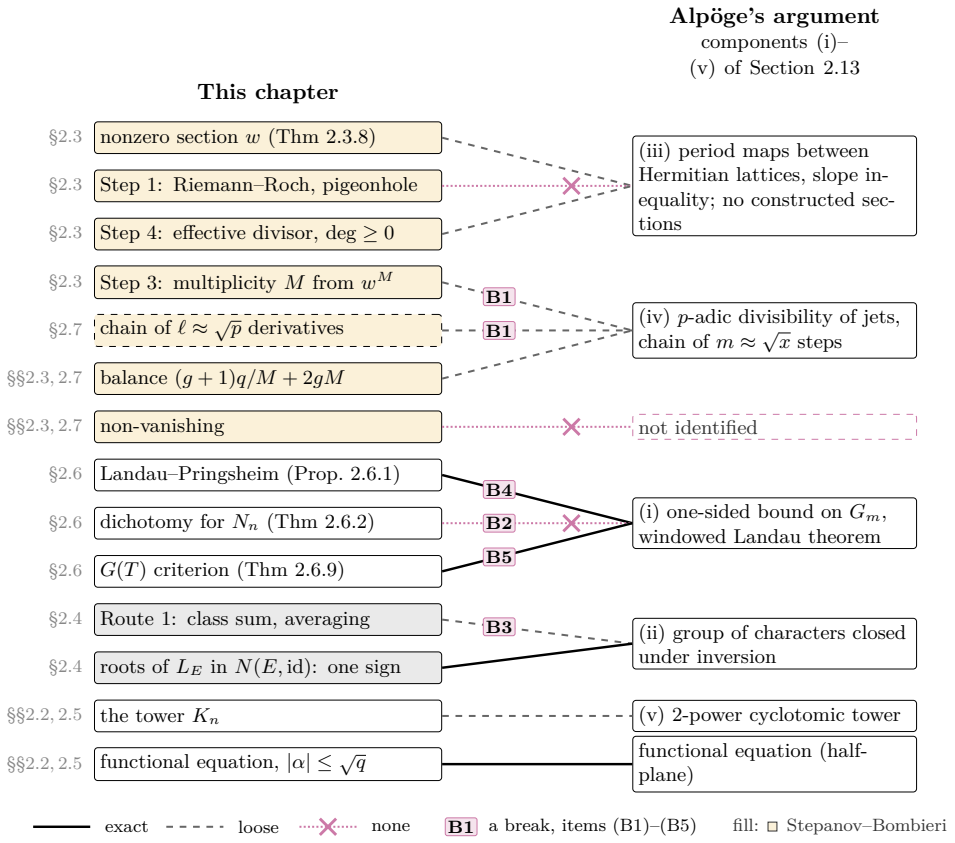


FIGURE 2.24. The dictionary of Section 2.13 as a graph. Left: steps of this chapter, with their sections in grey; orange fill marks the Stepanov–Bombieri construction (dashed border: its derivative form, Section 2.7), grey fill Route 1. Right: the components (i)–(v) of Alpöge's argument, with the functional equation and "not identified". Edges: solid black for an exact counterpart, dashed grey for a loose one, dotted purple with a cross for none; the ratings are judgments, as in the table. Purple flags: where the analogy breaks, items (B1)–(B5), placed on the edges they cut. Built from the table above; nothing here confirms or refutes Alpöge's claim.

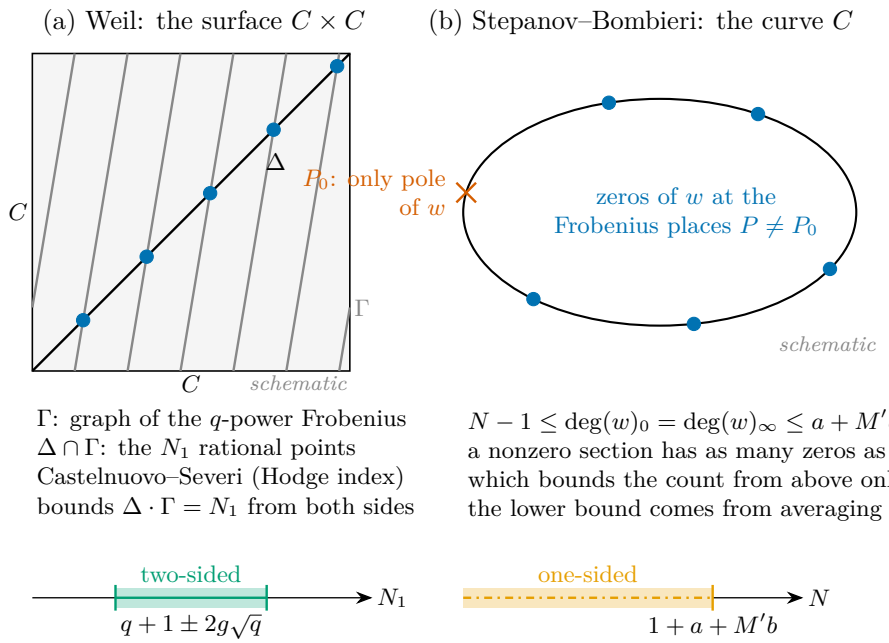


FIGURE 2.25. Where the positivity lives. (a) Weil's proof: on the surface $C \times C$, the diagonal Δ (black) and the graph Γ of the q-power Frobenius (grey) meet in the N_1 rational points (blue), and the Castelnuovo–Severi inequality, a form of the Hodge index theorem, bounds $\Delta \cdot \Gamma$ from both sides, giving the interval $|N_1 - q - 1| \leq 2g\sqrt{q}$ (green; Section 2.14). (b) This proof: on C itself, the auxiliary section w has its only pole at P_0 (vermillion cross) and vanishes at the σ-Frobenius places $P \neq P_0$ (blue), and $\deg(w)_0 = \deg(w)_\infty$ bounds their number from above only, $N(K, \sigma) \leq 1 + a + M'b$ (orange); the lower bound comes from averaging (Section 2.4). Schematic: dimensions reduced, positions arbitrary, and the number of branches of Γ is not q.

CHAPTER 3

The argument assembled

This chapter states the main theorem and gives its whole proof. Each tool that the proof uses is taken as a black box: a result proved in one of Chapters 4–12, restated here with the same hypotheses and the same conclusion, and followed by a few sentences on what it says, why the proof needs it and where it is proved. The aim is that the reader sees the flow of the argument and how its parts fit together before any of the parts is opened. Chapter 1 describes the idea informally, and Chapter 2 works out the model over a finite field. The present chapter is the precise assembly. The later chapters use from it only the statement of Theorem 3.1.1.

The order is as follows. Section 3.1 states the theorem. Section 3.2 gives the logical skeleton of the proof and its dependency diagram, and Section 3.3 recalls the terms in which the black boxes are stated. Sections 3.4–3.11 state the black boxes in the order in which the proof calls on them, and Section 3.12 is the proof. Section 3.13 compares the assembled argument with Chebyshev's and with the function-field model of Chapter 2, and Section 3.14 records the status of each input: which statements are proved in this book, which are outlined, and which are cited.

3.1. The main theorem

A *Dirichlet character* modulo $N \geq 1$ is a homomorphism $(\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$, extended to a function χ on $\mathbb{Z}$ that vanishes on the integers not prime to N. Its L-function is $L(s,\chi) = \sum_{n\geq 1} \chi(n) n^{-s}$ for $\operatorname{Re} s > 1$, continued meromorphically to $\mathbb{C}$; the continuation is entire when χ is nonprincipal, and has a simple pole at $s = 1$ and no other pole when χ is principal. A character is *primitive* if it is not induced from a character of smaller modulus. The character modulo 1 is the primitive principal character; its L-function is ζ.

THEOREM 3.1.1 (Alpöge). *Let χ be a Dirichlet character. Then $L(s,\chi) \neq 0$ for every s with $\operatorname{Re} s > 1/2$, except $s = 1$ when χ is principal, and every zero of $L(s,\chi)$ with $0 < \operatorname{Re} s < 1$ has real part $1/2$. Every zero of $\zeta(s)$ other than $-2, -4, -6, \ldots$ has real part $1/2$.*

REMARK 3.1.2 (scope of the conclusion). For the principal character the point $s = 1$ is excluded because its L-function has a pole there. The zeros of ζ at the negative even integers are its trivial zeros. For an imprimitive character, the extra Euler factors can have zeros on $\operatorname{Re} s = 0$; they do not

affect the critical-strip statement of Theorem 3.1.1. The theorem concerns ζ and Dirichlet L-functions.

REMARK 3.1.3 (the final statements of the archive). The half-plane statement of Alpöge's Lean archive [**Alp26a**] is not produced for every character directly. The final statements use Mathlib's own definitions of the Riemann zeta function and of Dirichlet L-functions, with no versions of the archive's own. The upstream output concerns primitive characters only. At the character of modulus 1 and $s = 1$ it is a statement about the finite value $(\gamma - \log 4\pi)/2$ that Mathlib assigns to ζ at its pole. Theorem 3.1.1 excludes that point. The archive was read as text and not built (Section A.9).

The proof is given in Section 3.12. Its last deduction is short: a bound on windows gives nonvanishing for primitive characters, the Euler factors extend it to imprimitive characters, and the functional equation reflects zeros across the critical line. The work of the argument is the bound on windows and the construction of the arithmetic data to which it applies.

3.2. The skeleton of the proof

The Riemann hypothesis asks for a sharp bound on the oscillation of primes around their expected density, and for a Dirichlet character the same question concerns primes in residue classes. The argument approaches these questions through an upper bound for smoothed, character-weighted Chebyshev functions. The bound is needed only on long multiplicative windows that occur arbitrarily far out, and a windowed converse to the explicit formula then excludes zeros to the right of the critical line. The difficulty lies in producing the windows and in obtaining the square-root exponent in the bound.

Fix a primitive Dirichlet character ψ. The terms used in the next four paragraphs are recalled in Section 3.3.

(I) *The analytic end.* To ψ is attached a positive combination $G_{\mathscr{L}_\psi}$ of smoothed Chebyshev functions, one for each primitive character inducing a power of ψ. If, for some $\sigma_0 \in (0,1)$, the inequality $G_{\mathscr{L}_\psi}(x) \le x^{\sigma_0}$ holds on windows $[Y, Y^{3/2}]$ with Y arbitrarily large, then $L(s,\psi) \neq 0$ for $\operatorname{Re} s > \sigma_0$. This is the windowed Landau criterion (Section 3.10). Positivity of the weights makes a one-sided upper bound sufficient, and the bound is not required at every large x.

(II) *The inequality.* For every small $\varepsilon > 0$ one needs $G_{\mathscr{L}_\psi}(x) \le x^{1/2+9\varepsilon}$ on such windows. This is the conclusion of the Chebyshev inequality on windows (Section 3.9). Its hypothesis is the existence of a *family* of automorphic data at ε, indexed by the points x of a set Win with windows, that is *tied* at every large $x \in \mathrm{Win}$ and satisfies an explicit-formula condition for $\mathscr{L}_\psi$.

(III) *Existence.* Tied families exist: this is the top existence statement, which is proved from the existence of packages and of admissible-family data (Section 3.8). It needs two inputs. The first is a supply of host fields for ψ on windows, meeting the explicit-formula condition and carrying a small Sha budget (Section 3.6). The second is the Sha bound on windows, which provides that budget (Section 3.5).

(IV) *The inputs behind the inputs.* The Sha bound rests on Galois representations with a truncated local–global compatibility (Section 3.4). The admissible-family data rest on the same Galois representations and on a covolume bound for lattices of integral automorphic forms, which is where arithmetic geometry enters (Section 3.7).

Given (I)–(III), a zero ρ of $L(s,\psi)$ with $\operatorname{Re}\rho > 1/2$ is excluded by choosing ε so small that $1/2 + 9\varepsilon < \operatorname{Re}\rho$. The passage from primitive characters to all characters, and from the half-plane to the critical line, is classical (Section 3.11).

The dependencies are displayed in Figure 3.1. The proof in Section 3.12 reads the figure from the bottom right to the top: Galois representations, the Sha bound, the supply of hosts, the existence of tied families, zero-freeness for primitive characters, and the passage to all characters.

Three of the boxes in Figure 3.1 carry almost all of the mathematics that is specific to this argument. The first is Step 2, the Chebyshev inequality, in which a chain of m theta lifts is compared on its archimedean and finite sides and m is taken near $\sqrt{x}$. The other two are the inputs at the bottom right: the Sha bound on windows and the Galois representations behind it. The assembly itself, as Remark 3.12.2 records, is a chain of implications.

3.3. The vocabulary of the black boxes

The black boxes are stated in terms that are defined in Chapters 4–12. This section recalls those terms, in the order in which they are needed. The definitions cited are the reference; where a definition is long, only the part that the statements of this chapter use is recalled, and this is said. Nothing in this section is new.

Smoothed Chebyshev functions. Let Λ be the von Mangoldt function and γ Euler's constant. For $x > 0$ the smoothed Chebyshev function of ζ and its normalised form are

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

For a nonprincipal Dirichlet character χ and $y > 0$ the complex twisted function is

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

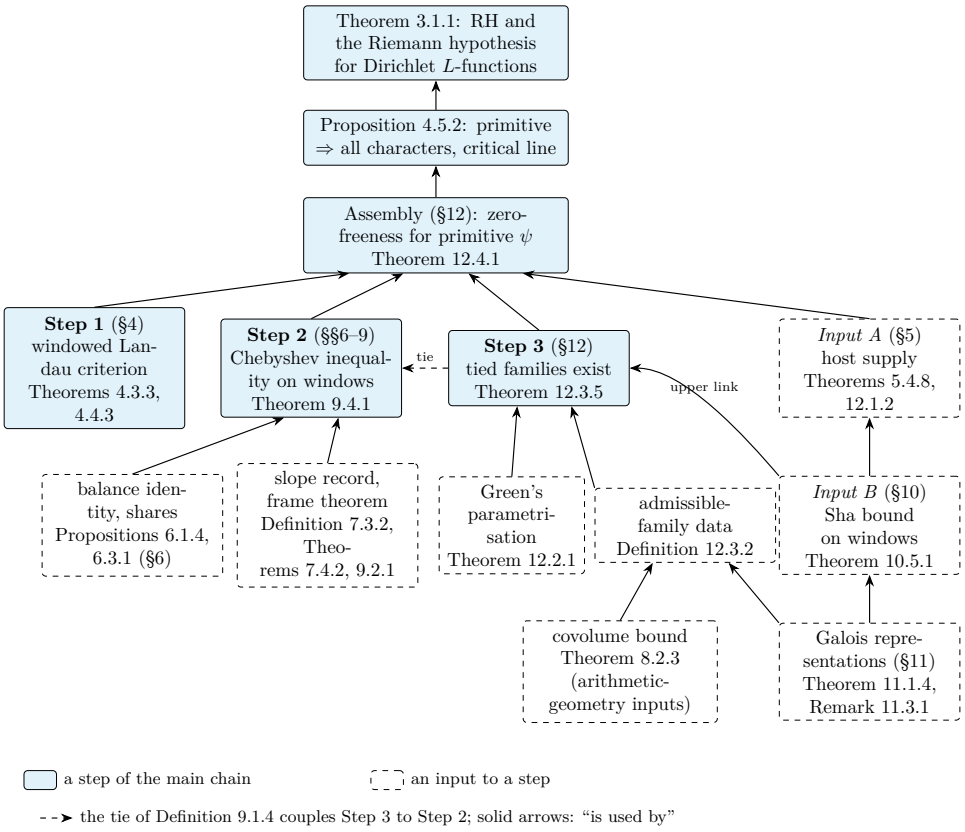


FIGURE 3.1. Dependencies in the proof of Theorem 3.1.1. Solid arrows point from inputs to results that use them. The dashed arrow marks the tie between a family and its slope record. A number after the sign § is the number of the chapter that develops the step.

and for the character modulo 1 one puts $g_1 = G_\zeta$ (Definition 4.2.1 and Proposition 4.2.2). The terms added to the prime sums remove the contribution of $s = 1$. A *character family* F is a finite list of pairs (χ_i, a_i), where each χ_i is a Dirichlet character that is either the character modulo 1 or nontrivial and each $a_i > 0$ is a real weight; its weighted Chebyshev function is $G_F(y) = \sum_i a_i \operatorname{Re} g_{\chi_i}(y)$ (Definition 4.3.1).

Levels. A *level* $\mathscr{L}$ consists of a finite set $\mathrm{Kept}(\mathscr{L})$ of primitive Dirichlet characters that contains the character modulo 1 and is closed under inversion, an optional exceptional character $\psi_{\mathrm{exc}} \in \mathrm{Kept}(\mathscr{L})$, and a natural number $o(\mathscr{L})$. The weight of $\psi \in \mathrm{Kept}(\mathscr{L})$ is $\mathrm{wt}(\psi) = 1$ if $\psi = \psi_{\mathrm{exc}}$ and $\mathrm{wt}(\psi) = (3o(\mathscr{L}) + 1)^{-1}$ otherwise, and the weighted Chebyshev function of

the level is

$$G_{\mathscr{L}}(y) = G_\zeta(y) + \sum_{\psi \in \mathrm{Kept}(\mathscr{L}),\ \psi \neq 1} \mathrm{wt}(\psi)\, \mathrm{Re}\, g_\psi(y).$$

For a primitive character ψ the *cyclic level* $\mathscr{L}_\psi$ has as kept set the primitive characters inducing the powers of ψ, no exceptional character, and $o(\mathscr{L}_\psi)$ equal to the order of ψ (Definition 4.4.1). In particular $\psi \in \mathrm{Kept}(\mathscr{L}_\psi)$, and all weights of a level are positive.

Windows. A set $S \subset \mathbb{R}$ *has windows of exponent* $\lambda > 1$ if for every y there is $Y \geq y$ with $[Y, Y^\lambda] \subset S$. A level $\mathscr{L}$ *satisfies the window criterion at* λ if for every $\sigma_0 \in (0,1)$ the following holds: if $\{x : G_{\mathscr{L}}(x) \leq x^{\sigma_0}\}$ has windows of exponent λ, then $L(s,\psi) \neq 0$ for every $\psi \in \mathrm{Kept}(\mathscr{L})$ and every s with $\mathrm{Re}\, s > \sigma_0$, where $s \neq 1$ when ψ is the character modulo 1 (Definition 4.4.2).

Scales and window pairs. Throughout, $\varepsilon \in (0, 1/52)$. A *scale* is a triple (x, ε, J) with $x > 1$ and $J \geq 1$; one puts $b = \exp(x^\varepsilon)$ and $a = xb$. A *system of scales* of exponent λ is a J together with a set $\mathrm{Win} \subset \mathbb{R}$ that has windows of exponent λ. An integer ν is *adapted* to (ε, x) if $x^{2+8\varepsilon} \leq 2^\nu \leq 2x^{2+8\varepsilon}$. A pair of primes (q, d) is a *window pair* for (ε, Y) if $q \equiv 3 \pmod 8$ and $Y^{\varepsilon/5}/2 < \log q \leq Y^{\varepsilon/5}$; $d \equiv 7 \pmod{16}$ and $Y^{\varepsilon/5} \log^2 Y/2 < \log d \leq Y^{\varepsilon/5} \log^2 Y$; $q \mid d+1$; and every odd prime $p \leq Y^{\varepsilon/5}$ divides $d+1$ (Definition 5.3.1). The upper bound $\varepsilon < 1/52$ is fixed once and for all; only small values of ε are used in the end.

Hosts. The geometry of the argument lives over an abelian CM field E attached to ψ. A *frame* ϕ fixes the shape of its Galois group: an integer $\nu \geq 0$, a finite abelian group A with a surjection onto $(\mathbb{Z}/2^{\nu+2}\mathbb{Z})^\times/\{\pm 1\}$, and a cyclic subgroup $A_0 \subset A$, the core, of index o; the Galois model is $\mathfrak{G} = \{\pm 1\}^2 \times A$ (Definition 5.1.1).

A *host* $\mathcal{H}$ for ϕ consists of auxiliary primes $q \equiv 3 \pmod 8$ and $d \equiv 7 \pmod{16}$ with $(-d/q) = 1$, an integer Q_S, a number field E that is Galois over $\mathbb{Q}$ with group $\mathfrak{G}$, contains $K = \mathbb{Q}(\sqrt{-d})$ and embeds in $\mathbb{Q}(\zeta_{Q_S qd})$, and an Artin map $(\mathbb{Z}/Q_S qd\mathbb{Z})^\times \to \mathfrak{G}$, subject to the conditions of Definition 5.1.2. A character θ of $\mathfrak{G}$ then determines a primitive Dirichlet character χ_θ. A *ψ-field* for (ψ, ν, q, d) is the subfield of a cyclotomic field cut out by the conditions $a \equiv \pm 1 \pmod{2^{\nu+2}}$, $\psi(a) = 1$, $(a/q) = 1$ and $(a/d) = 1$ (Definition 5.1.3); the host fields used for ψ are of this kind, with ν adapted to the scale, so their degree is about $x^{2+8\varepsilon}$.

A host is used together with a finite set $\mathcal{T} \subset A$ of *top places*. The host *satisfies the explicit-formula condition* $\mathrm{EF}(\mathscr{L}, \mathcal{T}, x, \varepsilon, C)$ if its index is $o(\mathscr{L})$, the characters of $\mathrm{Kept}(\mathscr{L})$ are among the characters χ_θ in the way prescribed by Definition 5.3.3, the set $\mathcal{T}$ has prescribed shares in the cosets

of A_0 up to $x^{-1+\varepsilon}$, and every *bounded* character θ of $\mathfrak{G}$ satisfies

$$|g_{\chi_\theta}(x)| \le x^{1/2+\varepsilon} \qquad \text{and} \qquad \Big|\frac{L'(1,\chi_\theta)}{L(1,\chi_\theta)}\Big| \le C\log x.$$

The bounded characters are characters of $\mathfrak{G}$ trivial on the core; they include every such character that is not kept (Definition 5.2.1). The pointwise bound is required at the single scale x, not for all large arguments.

Host requirements. Let $\lambda > 1$. A *host requirement* $\mathscr{R}$ is a property of a scale (x, ε, J), a frame, a host $\mathcal{H}$ for that frame, a finite set $\mathcal{T} \subset A$ and a positive integer t. It is *supplied on windows of exponent* λ if for every $\varepsilon \in (0, 1/52)$ there are constants and a system of scales of exponent λ such that at every large scale of the system there are ranks $\mathbf{k} = (k, m, t)$, a host $\mathcal{H}$ for $\mathbf{k}$ with top places $\mathcal{T}_{\mathcal{H}}$ that is *good* at the scale, and $\mathscr{R}$ holds for them (Definition 12.1.1). The goodness of a host at a scale is a list of size, splitting and character conditions; it is part of the schematic content of that definition and is not recalled here. For a primitive character ψ, the *requirement of* ψ is the host requirement $\mathscr{R}_\psi$: the host satisfies $\mathrm{EF}(\mathscr{L}_\psi, \mathcal{T}, x, \varepsilon, 4)$, and its field E has Sha budget $(t, b^{1/2})$ in the sense recalled in Section 3.5.

Packages, families and ties. A *ranks datum* for a frame is a triple $\mathbf{k} = (k, m, t)$ of integers subject to the inequalities of Definition 7.1.1; m is the number of theta lifts. Over a host field there is a tower of hermitian spaces: definite spaces $X_0, \dots, X_m$, whose unitary groups carry algebraic automorphic forms, and spaces $W_1, \dots, W_m$ of signature $(p_j, q_j) = (k + j - 1, m - j)$, whose unitary groups carry Shimura varieties (Definition 7.1.2). A *package* $\mathcal{P}$ is the automorphic input at one scale: a host, a ranks datum, the tower, forms on the unitary groups of the tower with their integral structures and theta lifts, and the hermitian lattices and period maps through which the lifts are measured (Definition 7.2.3). A *family at* ε, written $\mathfrak{F}$, consists of a threshold, an unbounded set $\mathrm{Win} \subset \mathbb{R}$, and for each x a finite set of members, each a package with a size datum at scale x, subject to uniformity conditions; it *has windows of exponent* λ if Win does (Definition 8.1.1).

The inequalities that a package satisfies are collected in a *slope record*: a list of real numbers subject to eleven inequalities (Definitions 7.3.1 and 7.3.2). A family is *tied at* x, with constants $e_D, e_L, c_W, c_S, e_K, C_L$, if there is a slope record at x whose numbers are the slopes, heights and local sizes of members of the family at x, one member for each shift, and which satisfies the hypotheses listed in Definition 9.1.4. The family satisfies the *uniform explicit-formula condition* for a level $\mathscr{L}$ if, with one constant C, the host of every member at every large x satisfies $\mathrm{EF}(\mathscr{L}, \mathcal{T}, x, \varepsilon, C)$ (Definition 9.1.3).

Classes of families and the top existence statement. Let $\lambda > 1$ and let $\mathscr{R}$ be a host requirement. The following terms are those of Definition 12.3.2.

(1) A family *satisfies the Sha hypothesis* if there is x_0 such that for every $x \geq x_0$ the host field of every member at x has Sha budget $(t, b^{1/2})$, with t the rank parameter and $b = \exp(x^{\varepsilon})$.
(2) The *class of $\mathscr{R}$ at λ*, written $\mathcal{C}_{\mathscr{R}}(\lambda)$, consists of the families with windows of exponent λ all of whose members beyond the threshold meet $\mathscr{R}$.
(3) A class $\mathcal{C}$ of families *supplies tied families* if there are real numbers e_D, e_L and $\varepsilon_1 > 0$ such that for every $\varepsilon \in (0, \varepsilon_1)$ some family at ε in $\mathcal{C}$ is tied at every large $x \in \mathrm{Win}$, with these e_D, e_L and some constants.
(4) The *top existence statement* is: for every $\lambda > 1$ and every host requirement $\mathscr{R}$, if $\mathscr{R}$ is supplied on windows of exponent λ and every family of $\mathcal{C}_{\mathscr{R}}(\lambda)$ satisfies the Sha hypothesis, then $\mathcal{C}_{\mathscr{R}}(\lambda)$ supplies tied families.
(5) *Admissible-family data* for a class $\mathcal{A}$ of families consist of realisation data, constants and a list of propositions about the families of $\mathcal{A}$, in four groups: automorphic forms, theta lifts and mass formulas; Galois representations and local–global compatibility; Arakelov geometry and covolume; counting and design.

The definitions of packages, size data, families, the covolume condition, host supply and admissible-family data are *schematic* in the sense of Section 4.1: they list the parts of the data and the conditions that the argument uses, and omit the remaining compatibility conditions. The consequences for the statements of this chapter are collected in Remark 3.14.2.

3.4. Black box: Galois representations

The first box supplies Galois representations attached to automorphic representations over CM fields. Two notions are needed to state it; both are defined in Chapter 11, and both are weaker than their counterparts in the literature.

A *weak local Langlands map* for a number field E and a finite place w is a map from the irreducible admissible representations of the groups $\mathrm{GL}_n(E_w)$, $n \geq 1$, to Weil–Deligne representations of the Weil group of E_w that satisfies a list of laws: it preserves dimension, is local class field theory on characters, is compatible with the Zelevinsky classification, and has further properties on a class of *good* representations (Definition 11.1.1). It is not required to be bijective, and no condition on L- or ϵ-factors is imposed. *Truncated local–global compatibility* is the statement that for every CM field E, every prime l, every isomorphism $\iota\colon \overline{\mathbb{Q}}_l \cong \mathbb{C}$, every $a \geq 1$ and every regular algebraic polarized cuspidal automorphic representation π of $\mathrm{GL}_a(\mathbb{A}_E)$ there is a *compatible system* for π: a continuous semisimple representation $r\colon G_E \to \mathrm{GL}_a(\overline{\mathbb{Q}}_l)$ with prescribed Hodge–Tate weights that matches π_w, through every weak local Langlands map, at the finite places w in the cases

listed in Definition 11.1.3. Nothing is asserted about the monodromy operator except where it vanishes, and nothing about purity.

THEOREM 3.4.1 (Alpöge; restated from Theorem 11.1.4). *(a) For every number field E and every finite place w of E, a weak local Langlands map for (E, w) exists.*
(b) Truncated local–global compatibility (Definition 11.1.3) holds.

What it says. Regular algebraic polarized cuspidal automorphic representations over CM fields have Galois representations with the local properties listed in Definition 11.1.3, and local Langlands maps exist in the weak sense above.
Why it is needed. It is used twice. The Sha bound of Section 3.5 bounds the eigenspaces of the wrong parity through automorphic points of ordinary deformation rings; these points are the Galois representations attached to regular algebraic polarized cuspidal representations of GL_3 and GL_4 over imaginary quadratic fields, with the local properties that part (b) supplies (Remark 11.3.1). The admissible-family data of Section 3.8 use the weak local Langlands maps of part (a), and their congruence link is listed with the propositions on Galois representations and local–global compatibility.
Where it is proved. Chapter 11, where the construction of part (b) is outlined at the level of its named steps and compared with the published constructions; the proof of part (a) is not outlined there. The statement is not taken as a hypothesis from the literature (Remark 11.1.5).

3.5. Black box: the Sha bound on windows

Let E be a number field, l a prime and $e \in \mathbb{Z}$. Let S_l be the set of places of E above l and G_{E,S_l} the Galois group of the maximal extension of E unramified outside S_l and the infinite places. The *Tate–Shafarevich group* of the twist e is

$$Ш_l(E, e) = Ш^1(G_{E,S_l}, \mathbb{Q}_l/\mathbb{Z}_l(-e)),$$

the group of classes in $H^1(G_{E,S_l}, \mathbb{Q}_l/\mathbb{Z}_l(-e))$ whose restrictions to the decomposition groups at S_l vanish, and $\mathrm{len}_l(e) \in \mathbb{Z}_{\geq 0} \cup \{\infty\}$ is its length as a $\mathbb{Z}_l$-module (Definition 10.1.1).

Let $t \geq 1$ be an integer and $B \geq 0$ real. The field E has *Sha budget* (t, B) if $\mathrm{len}_p(e) < \infty$ for every prime p and every $0 < e < t$, and

$$\sum_{l \in S} \log l \sum_{1 \leq e < t} \mathrm{len}_l(e) \leq B$$

for every finite set S of primes. Let $C_t > 0$. A primitive Dirichlet character ψ has *Sha bounds on windows with constant C_t* if the following holds for every $\varepsilon \in (0, 1/52)$ and every function ν on $[1, \infty)$ such that $\nu(x)$ is adapted to (ε, x) for every $x \geq 1$. There are $J \geq 1$ and X_0 such that for every $X \geq X_0$ and every window pair (q, d) for $(\varepsilon, X^{1/J})$ there is a real Y with

$X^{1/J} \le Y$ and $Y^2 \le X$ such that, for every $x \in [Y, Y^2]$, every ψ-field for $(\psi, \nu(x), q, d)$ has Sha budget $(\lfloor x^{C_t} \rfloor, \exp(x^{\varepsilon})^{1/2})$ (Definition 10.1.2).

THEOREM 3.5.1 (Alpöge; restated from Theorem 10.5.1). *Every primitive Dirichlet character has Sha bounds on windows with constant* 401 *(Definition 10.1.2).*

What it says. On some window $[Y, Y^2]$ inside every range $[X^{1/J}, X]$, the CM fields of degree about $x^{2+8\varepsilon}$ attached to ψ satisfy $\sum_l \log l \sum_{1 \le e < x^{401}} \mathrm{len}_l(e) \le \exp(x^{\varepsilon}/2)$. The window may depend on X and on the pair (q, d), but the exponent J may not. By Poitou–Tate duality this is a bound, uniform in l, on the subgroups Ш^2 of the groups $H^2(G_{E,S_l}, \mathbb{Z}_l(e+1))$ (Remark 10.1.3).
Why it is needed. The congruence defects at the twisted places of the families of Section 3.8 are bounded by lengths of Selmer groups, and these by the Sha budget of the host field. The budget enters the argument as part of the requirement $\mathscr{R}_\psi$ on hosts, so the box is the hypothesis of the supply of hosts (Section 3.6); through the Sha hypothesis on families it is then spent in the divisibility bound of the upper link (Remark 12.3.3). The box is also the reason for the windows. For a bounded number of exceptional Tate twists the only bound available is doubly exponential, and a pigeonhole argument gives the bound only on windows of scales (Remark 10.5.2). The size of the bound is generous: heuristics predict a total polynomial in x, far below $\exp(x^{\varepsilon}/2)$ (Remark 10.1.4). The difficulty is uniformity in the pair (q, d) and in l, above all for the eigenspaces of the wrong parity, where no L-value controls the group.
Where it is proved. Chapter 10. The proof splits the budget into a 2-part, a minus part and a plus part, and sorts the plus eigenspaces into six cases. Its automorphy input is Theorem 3.4.1, part (b). The remaining inputs are classical Galois cohomology and the theorems stated in that chapter.

3.6. Black box: the supply of hosts

The host fields are produced in two stages, and the second stage is the box that the proof of Theorem 3.1.1 calls. The first stage, in Chapter 5, shows that for every primitive ψ, every large frame exponent and every suitable pair of auxiliary primes (q, d) there is a host whose characters trivial on the core are the twists $\xi \chi_q^i \chi_d^j$ of the powers ξ of ψ by the quadratic characters of q and d (Theorem 5.2.4). The bounded characters are twists with $(i, j) \ne (0, 0)$ (Proposition 5.2.3), and the explicit-formula condition asks for a bound on them at the scale x. This is arranged by the choice of (q, d). A window pair (q, d) for (ε, Y) with $\gcd(M, qd) = 1$ is *good* for (M, ε, Y, X) if the twists of the characters modulo M satisfy $|g_\chi(y)| \le y^{1/2+\varepsilon/2}$ and $|L'(1, \chi)/L(1, \chi)| \le 4 \log y$ for every $y \in [Y, X]$, and the twists of the real primitive characters of conductor dividing $8M$ have zero-free boxes of prescribed heights (Definition 5.4.2).

THEOREM 3.6.1 (Alpöge; restated from Theorem 5.4.8). *Let $M > 0$ be an integer, $0 < \varepsilon < 1/52$ and $\Delta \geq 1$. There is Y_1 such that for every $Y \geq Y_1$ some window pair for (ε, Y) is good for $(M, \varepsilon, Y, Y^{\Delta})$.*

The proof of Theorem 3.6.1, outlined in Chapter 5, is an averaging over the auxiliary primes: since $\log q \approx Y^{\varepsilon/5}$, a test point $y \in [Y, Y^{\Delta}]$ is polylogarithmic in the conductors of the twists, and at such a scale a twisted prime sum is small for most q and d by a mean-square estimate. No hypothesis of Riemann type enters (Remark 5.5.1).

THEOREM 3.6.2 (Alpöge; restated from Theorem 12.1.2). *Let ψ be a primitive Dirichlet character. If ψ has Sha bounds on windows with constant 401 (Definition 10.1.2), then $\mathscr{R}_{\psi}$ is supplied on windows of exponent 3/2.*

What it says. Given the Sha bound for ψ, there is, for every $\varepsilon \in (0, 1/52)$, a set of scales with windows $[W, W^{3/2}]$ at every large scale of which there is a good host for ψ that satisfies the explicit-formula condition $\mathrm{EF}(\mathscr{L}_{\psi}, \mathcal{T}, x, \varepsilon, 4)$ and whose field has Sha budget $(t, b^{1/2})$.
Why it is needed. It is one of the two hypotheses of the zero-freeness theorem for primitive characters (Theorem 3.12.1): the families to which the Chebyshev inequality is applied are built on these hosts, and the explicit-formula condition and the Sha budget of the hosts become the uniform explicit-formula condition and the Sha hypothesis of the families.
Where it is proved. Chapter 12, by an outline that combines Theorem 5.2.4, Theorem 3.6.1 and Proposition 5.4.9 of Chapter 5 with the Sha bound. One pair (q, d) serves both the goodness and the Sha window. The Sha bound gives windows $[Y, Y^2]$, and the supply of hosts is stated with windows $[W, W^{3/2}]$. The step that turns the exponent 2 into 3/2 is not reconstructed in this book. It is Question 12.1.3, which follows the proof of Theorem 12.1.2 in Chapter 12.

3.7. Black box: the covolume bound

For a member of a family and a lift index $1 \leq j \leq m$, let M_j be the lattice of integral forms on the indefinite unitary group $U(W_j)$, of signature $(p_j, q_j) = (k + j - 1, m - j)$, formed at a *point* of the member, that is, at a choice of auxiliary integral data. The *point conditions* are compatibility conditions on such a choice. A point carries weights $w(j, \rho)$, a period Ω and a positive integer deg, its degree. It satisfies the *covolume inequality* with constant C at x if for every lift index j

$$\frac{\log \mathrm{covol}(M_j)}{\deg} \leq \sum_{\rho \in \Phi} (p_j - q_j)\, w(j, \rho) \log \Omega + C\, B_0 \log\log x,$$

where $\log \mathrm{covol}(M_j)$ is the normalised logarithmic covolume of M_j and B_0 is a size quantity of the member. The *covolume condition* is the statement that this inequality, for all large x and all large C, is a *closed* property on a fixed class $\mathcal{C}_{\mathrm{cov}}$ of families: for ε small it holds for every family of the class

whose constants exceed thresholds chosen before the family (Definitions 8.1.2 and 8.2.1). The *member hypothesis* of that definition bounds the size of the coefficient field of a package and fixes its local types at the two norm primes; the period Ω is the CM period of the slope record, with

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

where v is the half root discriminant of the host field (Remark 7.5.2).

THEOREM 3.7.1 (Alpöge; covolume bound; restated from Theorem 8.2.3). *There is $\varepsilon_1 > 0$ such that for every $\varepsilon \in (0, \varepsilon_1]$ there is a threshold κ with the following property. Let $\mathfrak{F}$ be a family at ε in $\mathcal{C}_{\mathrm{cov}}$ and c a real number such that $\mathfrak{F}$ is large for (κ, c). Then there is $C_0 \geq 0$ such that for all large x, every member $(\mathcal{P}_{x,i}, Z_{x,i})$ satisfying the member hypothesis, every point of $\mathcal{P}_{x,i}$ satisfying the point conditions and every $C \geq C_0$, the point satisfies the covolume inequality with constant C at x (Definition 8.2.1(3)).*

What it says. The covolume condition (Definition 8.2.1) holds: the property of a family stated in the last sentence of the theorem is closed on $\mathcal{C}_{\mathrm{cov}}$. For the lattice M_j of integral cusp forms of the weights $w(j, \rho)$ on the Shimura variety of $U(W_j)$, at the level prescribed by the package, the normalised logarithmic covolume is at most $\sum_{\rho \in \Phi} (p_j - q_j)\, w(j, \rho) \log \Omega$, up to an error $O(B_0 \log \log x)$, for all large x and uniformly over the members of a family whose field of coefficients is not too large.
Why it is needed. It is the one result through which arithmetic geometry enters the argument (Remark 8.3.1). Summed over j, its main terms give the term $-m a_\tau \log \Omega$ of the divisor bound of the upper link (Definition 7.5.1), which is the archimedean side of the comparison. In the assembly it is an input to the admissible-family data: the propositions of their Arakelov group are obtained from the covolume condition (Theorem 3.8.2, part (c)).
Where it is proved. Chapter 8. The proof is not reproduced in this book; the chapter lists its inputs, which are cited theorems of arithmetic geometry (among them toroidal compactifications, resolution of singularities, the arithmetic Hilbert–Samuel theorem, a bigness theorem, vanishing theorems and Grothendieck–Messing theory), and records the form in which each is used (Remark 8.3.4).

3.8. Black box: packages, Green's parametrisation, tied families

Three results produce the families. The first is a theorem of the literature on the characters of $\mathrm{GL}_n(\mathbb{F}_p)$.

THEOREM 3.8.1 (Green [**Gre55**]; restated from Theorem 12.2.1). *Let p be a prime and $n \geq 1$. There is a map $P \mapsto \chi_P$ from the monic irreducible polynomials $P \in \mathbb{F}_p[X]$ of degree n with $P(0) \neq 0$ to the cuspidal irreducible characters of $\mathrm{GL}_n(\mathbb{F}_p)$ such that*

(a) $\chi_P(1) = \prod_{i=1}^{n-1}(p^i - 1)$ *for every such* P*;*
(b) $P \mapsto \chi_P$ *is injective;*
(c) *every cuspidal irreducible character of* $\mathrm{GL}_n(\mathbb{F}_p)$ *is* χ_P *for some such* P*.*

Green's theorem is not an analytic input, and the window exponent 3/2 does not come from it. It enters only through local types at the norm primes ϖ_0, ϖ_1 of a host: it is used to show that the rank defect of each member of a family is small, and to label the local types in the ties (Remark 12.2.3).

THEOREM 3.8.2 (Alpöge; restated from Theorem 12.3.4). *Let* $\varepsilon \in (0, 1/52)$.

(a) *For every scale* (x, ε, J) *with* x *sufficiently large there are packages (Definition 7.2.3) whose automorphic forms have the weights, level and local types prescribed at that scale, on the tower of unitary groups of Definition 7.1.2 over a CM field of degree about* $x^{2+8\varepsilon}$*.*
(b) *The members of the families built from these packages have points, and for each lift index* j *the lattice* M_j *of the covolume condition (Definition 8.2.1) is the lattice of integral cusp forms on* $U(W_j)$*.*
(c) *There is a class of families of these packages that carries admissible-family data (Definition 12.3.2).*

What it says. At every large scale there are automorphic forms on the tower of unitary groups over a CM field of degree about $x^{2+8\varepsilon}$ with the prescribed weights, level and local types. Their integral cusp forms give the lattices M_j of the covolume bound, and families of these packages satisfy the propositions that make up admissible-family data.

THEOREM 3.8.3 (Alpöge; restated from Theorem 12.3.5). *Suppose that some class* $\mathcal{A}$ *of families carries admissible-family data (Definition 12.3.2). Let* $\lambda > 1$*, and let* $\mathscr{R}$ *be a host requirement (Definition 12.1.1) that is supplied on windows of exponent* λ *and such that every family of* $\mathcal{C}_{\mathscr{R}}(\lambda)$ *satisfies the Sha hypothesis. Then* $\mathcal{C}_{\mathscr{R}}(\lambda)$ *supplies tied families.*

What it says. Suppose some class of families carries admissible-family data. Let $\mathscr{R}$ be a host requirement that is supplied on windows of exponent λ, and suppose the host fields of the families built on these hosts satisfy the Sha bound. Then for every small ε there is a family of packages, built on the supplied hosts, that is tied at every large scale of its windows. For the families produced, $e_D = 2$ and $e_L = 0$.
Why they are needed. Theorem 3.8.2, part (c), provides the first hypothesis of Theorem 3.8.3, the class $\mathcal{A}$. The rest of Theorem 3.8.3, read for every $\lambda > 1$ and every host requirement $\mathscr{R}$, is the top existence statement of Section 3.3, and this statement is a hypothesis of the zero-freeness theorem for primitive characters (Theorem 3.12.1). The top existence statement does not mention the class $\mathcal{A}$, so one class carrying admissible-family data suffices (Remark 12.3.6). Supplying tied families cannot hold vacuously: a closed

property chooses its threshold before the family, the family produced is large for that threshold, and an empty class supplies no tied family.
Where they are proved. Chapter 12. For Theorem 3.8.2 the chapter indicates the standard method for parts (a) and (b) and lists the inputs of part (c); among them are the covolume bound (Theorem 3.7.1) and the weak local Langlands maps of Theorem 3.4.1. Theorem 3.8.3 is proved in outline. The links of a family (Definition 12.3.1) are closed properties, so one threshold serves for all of them. Green's parametrisation enters in the closed form of the irreducibility link. The design statement of the data, fed by the supplied hosts, gives a family that is large for the threshold. The global link shows that this family is tied.

In terms of the objects of Chapters 7–9, the data encode nine steps:

(1) build the packages and their lattices;
(2) form the period map;
(3) apply the slope inequality (Theorem 7.4.2);
(4) bound the archimedean side;
(5) bound the covolume (Theorem 3.7.1);
(6) bound the finite side;
(7) balance the two sides (Theorem 3.9.3);
(8) bound the defects at the twisted places by Sha lengths and the Sha hypothesis;
(9) supply near-orthogonality, good local types, masses and class numbers.

3.9. Black box: the Chebyshev inequality on windows

This box contains the main inequality of the argument. Its hypotheses are stated in the terms of Section 3.3: a level, a family at ε, the uniform explicit-formula condition and the ties.

THEOREM 3.9.1 (Alpöge; Chebyshev inequality on windows; restated from Theorem 9.4.1). *Let $\mathscr{L}$ be a level, e_D and e_L real numbers and $\varepsilon > 0$. Let $\mathfrak{F}$ be a family at ε that satisfies the uniform explicit-formula condition for $\mathscr{L}$ (Definition 9.1.3). Suppose that there are real numbers c_W, c_S, e_K, C_L and x_0 such that $\mathfrak{F}$ is tied (Definition 9.1.4) with constants $e_D, e_L, c_W, c_S, e_K, C_L$ at every $x \in \mathrm{Win}$ with $x \geq x_0$. Then there is x_1 such that*

$$G_{\mathscr{L}}(x) \leq x^{1/2+(e_D+e_L+7)\varepsilon}$$

for every $x \in \mathrm{Win}$ with $x \geq x_1$.

What it says. If a family is tied at every large point of its window set, with exponents e_D and e_L that do not depend on the point, then the weighted Chebyshev function of the level is at most $x^{1/2+O(\varepsilon)}$ on that set. The bound is one-sided and holds only for $x \in \mathrm{Win}$.
Why it is needed. It is the only source of a square-root bound in the argument. Applied to $\mathscr{L}_\psi$ and to the tied families of Theorem 3.8.3, for which

$e_D + e_L + 7 = 9$, it gives $G_{\mathscr{L}_\psi}(x) \le x^{1/2+9\varepsilon}$ on windows of exponent $3/2$, which is the hypothesis of the windowed Landau criterion. The theorem would say nothing for a family whose window set had no large elements, or if e_D and e_L could grow as $\varepsilon \to 0$. Neither happens: the window set of a family is unbounded, and in the top existence statement e_D and e_L are fixed before ε is sent to 0 (Remark 9.4.2).

Where it is proved. Chapter 9, in outline, from the four results restated below. The proof first forgets the geometry: at each $x \in \mathrm{Win}$ the tie provides a slope record, the arithmetic data of a host and auxiliary real numbers, and after that only inequalities between real numbers are used.

Inside the box. The inequality is assembled from an identity and three bounds. They are restated here because the route through them is the core of the argument; the proof of Theorem 3.1.1 in Section 3.12 calls only Theorem 3.9.1. Their statements use the following terms. Let H be a finite abelian group, $V_4 = (\mathbb{Z}/2\mathbb{Z})^2$ and $\mathfrak{G} = V_4 \times H$, with the index-two subgroup Φ of the elements whose first coordinate is 0, the *CM type*. *Arithmetic data* D for $\mathfrak{G}$ consist of natural numbers q_1 and d and, for each prime p, an inertia subgroup $I_p \subset \mathfrak{G}$ and an element $\mathrm{Frob}_p \in \mathfrak{G}$; for a host they are read off from the Artin map, with $H = A$ and $\Phi = \mathrm{Gal}(E/K)$ (Definitions 6.1.1 and 6.1.2). For a function w on $\mathfrak{G}$ that is positive on Φ, a point $\tau \in H$ and $a > 0$, Definition 6.1.3 attaches to D a *prime sum*

$$\mathrm{PS}_{\tau,w}(a) = \sum_{p \le N \text{ inert}} 2\log p \sum_{\substack{1 \le i \le N \\ i\,\mathrm{Frob}_p \notin \Phi}} \Big(\frac{a}{p^i} - m_\tau(i\,\mathrm{Frob}_p)\Big)_+ + (\text{a term for the prime } d),$$

a *main term*

$$\mathrm{MT}_w(a) = a\Big(\log a - 1 - \gamma + \ell(\chi_K) - \frac{1}{f}\sum_{\rho \in \Phi} \log w(\rho)\Big),$$

and a *defect* $\mathrm{DF}_{\tau,w}(a)$, which is a combination, with coefficients $\overline{\theta(g)}\, m_\tau(g)/f$, of the values $G_\theta^{\mathrm{ar}}(a/m_\tau(g))$ over all characters θ of $\mathfrak{G}$ and all $g \notin \Phi$. Here m_τ is the partner weight, $f = |\Phi|$, and G_θ^{ar} is the arithmetic Chebyshev function attached to θ by D and a function ℓ on the characters of $\mathfrak{G}$. It is built from the *averaged values* $\theta_D(n)$ of θ at the integers $n \ge 1$ (Definition 6.1.2). We write θ_0 for the trivial character of $\mathfrak{G}$. The function $G_{\theta_0}^{\mathrm{ar}}$ is G_ζ. For a host one has $\theta_D(n) = \chi_\theta(n)$, and with $\ell(\theta) = L'(1,\chi_\theta)/L(1,\chi_\theta)$ the function G_θ^{ar} is g_{χ_θ} for $\theta \ne \theta_0$.

Proposition 3.9.2 (balance identity; restated from Proposition 6.1.4). *Let D be arithmetic data for $\mathfrak{G} = V_4 \times H$ that are split outside d on a finite set S of primes, let ℓ be a function from the characters of $\mathfrak{G}$ to $\mathbb{C}$, let $a > 0$ and $\tau \in H$, let $w\colon \mathfrak{G} \to \mathbb{R}$ be positive on Φ, and let N be an integer with $a/w(\rho) \le N$ for every $\rho \in \Phi$. Then*

$$\mathrm{PS}_{\tau,w}(a) - \mathrm{MT}_w(a) = \mathrm{DF}_{\tau,w}(a).$$

The prime sum has the shape of Legendre's formula for $\log\lfloor a\rfloor!$, restricted to the primes inert in K and smoothed (Remark 6.1.5). The identity is finite Fourier analysis on $\mathfrak{G}$ and is proved completely in Chapter 6. Through it, an inequality PS $\le$ MT + (small) is an upper bound for a combination of the functions G_θ^{ar}. For a slope record with shifts $u_0 \in H$, the *gap* at u_0 is the sum over the top places τ and the lift indices $j < m$ of the differences PS $-$ Re MT at $a = a_\tau$ and $w = w(u_0, j)$ (Definition 6.2.1 and Proposition 9.1.2), and the *error bracket* Err is the explicit quantity of Definition 6.2.3.

THEOREM 3.9.3 (Alpöge; gap bound; restated from Theorem 9.2.1). *Let H be a finite abelian group of order n and $\mathfrak{G} = V_4 \times H$. Let a slope record with shifts $U_0 = H$, positive integers m and f and size constants be given (Definitions 7.3.1 and 7.3.2); write x, ε, a, b and k for the quantities of the record, and c_{sp}, e_{div} and e_{loss} for its size constants. Let c_W and e_K be real numbers. Let D be arithmetic data for $\mathfrak{G}$, ℓ a function from the characters of $\mathfrak{G}$ to $\mathbb{C}$, $N \ge 1$ an integer, $\mathcal{T}, \mathcal{T}_2 \subset H$ finite sets, $a\colon H \to \mathbb{R}$ and $w\colon H \times \{0, \dots, m-1\} \times \mathfrak{G} \to \mathbb{R}$. Suppose that the record satisfies the size data (Definition 7.3.2), that the identification clause, the window conditions with c_W and e_K and the bounds on B_0 and N_c (Definition 9.1.1) hold, and that*

$$\log x \le x^{\varepsilon} \qquad \text{and} \qquad 200(1 + c_W)^2(1 + c_{\mathrm{sp}})(1 + e_K) \le x^{\varepsilon}.$$

Then for every $u_0 \in H$

$$\mathrm{gap}(u_0) \le x^{(e_{\mathrm{div}} + e_{\mathrm{loss}} + 3)\varepsilon}\,\mathrm{Err},$$

with Err *the error bracket of Definition 6.2.3 for $n = |H|$, $n_T = |\mathcal{T}|$, $n_{T_2} = |\mathcal{T}_2|$, the given b, k and m, and $a_E = a$.*

The slope record is spent in the gap bound. The record contains two bounds for one number, the frame slope s of the period lattice data. A lower bound comes from the divisor bound and the quotient slopes; it carries the finite-side gain, which bounds the prime sum, and the archimedean sum. An upper bound comes from the slope inequality for the period maps into the test lattice. Subtracting them removes the slope of the test lattice and the heights, and leaves an inequality that one may read as "finite-side gain at most archimedean main term" (Remark 9.2.2). The rest of the proof is linear arithmetic under the window conditions. The slope inequality has the form of the conclusion of the following theorem, in which μ, $\mu_{\max}$, $\mu_{\min}$ are the slope and the extreme slopes of a hermitian lattice, ht is the height of a map (Definition 7.2.1), and the frame operator and the Hilbert–Schmidt defect are those of Definition 7.4.1.

THEOREM 3.9.4 (Alpöge; frame theorem; restated from Theorem 7.4.2). *Let L be a totally complex number field and $\sigma\colon L \hookrightarrow \mathbb{C}$ an embedding. Let W be an L-vector space with a basis indexed by a finite set J, with a positive definite hermitian Gram matrix at each infinite place, and let $R \subset W$ be a finitely generated $\mathcal{O}_L$-submodule spanning W. Let I be a nonempty finite*

set, and for $i \in I$ let V_i be an L-vector space with a basis indexed by a finite set ι_i and positive definite hermitian Gram matrices, $N_i \subset V_i$ a finitely generated $\mathcal{O}_L$-submodule spanning V_i, $B_i \colon V_i \to W$ an L-linear map, and h_i a real number; assume $\sum_i |\iota_i| > 0$. Let J_1 be a finite set, and let Y' and Y_i ($i \in I$) be invertible complex matrices indexed by J and by ι_i. For each i let $\widetilde{B}_i$ be a complex matrix with rows indexed by $J \sqcup J_1$ and columns by ι_i, whose J-rows form the matrix $Y'^{-1}[B_i]_\sigma Y_i$, where $[B_i]_\sigma$ is the matrix of B_i in the given bases, embedded by σ. Let $c_i > 0$ for $i \in I$, let Q be the frame operator of $(c_i, \widetilde{B}_i)$ (Definition 7.4.1), and let $s, u, \Delta, \beta_0, \beta_1, \beta_2$ be real numbers. Suppose:

- *(a) $B_i(N_i) \subset R$ for all i, and $\operatorname{ht}(B_i) \le h_i$ whenever $B_i \neq 0$;*
- *(b) $s \sum_i \operatorname{rk} N_i \le \sum_i \operatorname{rk} N_i \, (\mu(N_i) - h_i)$; $u \ge 0$ and $\mu_{\max}(N_i) - h_i \le s + u$ for all i; $\Delta \ge 0$ and $s - \mu_{\min}(R) \le \Delta$;*
- *(c) $0 \le \beta_0 < 1$ and $(1 - \beta_0)|J \sqcup J_1| \le |J|$;*
- *(d) the Hilbert–Schmidt defect of $(c_i, \widetilde{B}_i)$ is at most $\beta_2 \sum_i \operatorname{rk} N_i$, with $0 \le \beta_2 \le 1/4$; and $\operatorname{tr}(Q^2) \le (1 + \beta_1)(\operatorname{tr} Q)^2/|J \sqcup J_1|$, with $\beta_1 \ge 0$.*

Then

$$s - \log|I| - \frac{1}{1-\beta_0}\Big(4u \log\Big(1 + \frac{\Delta}{u}\Big) + (\beta_1 + 4\sqrt{\beta_2})\Delta\Big) \le \mu(R),$$

where the term $4u \log(1 + \Delta/u)$ is read as 0 when $u = 0$.

The frame theorem compares a rank-weighted mean of the slopes of many source lattices with the slope of one target lattice. It bounds an average by the slope of the target, and not by its maximal slope. So it is not Bost's slope inequality applied to each period map. Bost's inequality is one step of its proof (Remark 7.4.3). No small sections are constructed. The period map between two given lattices is shown to be integral and of bounded height, and the slope inequality is applied to it (Remark 7.2.6). Chapter 7 states the theorem and gives an account of the method.

The next two results turn the gap bound into a bound for $G_{\mathscr{L}}$. For arithmetic data D, a function ℓ, a subgroup $A_0 \subset H$, a set Ξ of characters of $\mathfrak{G}$, the kept set, and a real number $\varkappa$, the *extracted function* G_A is the combination of the functions $\operatorname{Re} G_\theta^{\mathrm{ar}}$, $\theta \in \Xi$, with weights 1 and $\varkappa$, and the *rest* Σ_A collects the characters trivial on $0 \times A_0$ that are not in Ξ (Definition 6.2.2).

THEOREM 3.9.5 (Alpöge; core Chebyshev bound; restated from Theorem 6.3.4). *Let $C \ge 0$, $\varepsilon > 0$, $\varkappa$ with $|\varkappa| \le 1$, and C_E, C_o, M be real numbers. There is x_1 such that the following holds for every $x \ge x_1$. Let H be a finite abelian group, $\mathfrak{G} = V_4 \times H$, m a natural number, D arithmetic data for $\mathfrak{G}$ that are split outside d on a finite set of primes, ℓ a function from the characters of $\mathfrak{G}$ to $\mathbb{C}$, $N \ge 1$ an integer, $A_0 \subset H$ a subgroup, $\mathcal{T}, \mathcal{T}_2 \subset H$ finite sets, $a \colon H \to \mathbb{R}$, $w \colon H \times \{0, \dots, m-1\} \times \mathfrak{G} \to \mathbb{R}$, $b, b', \eta, X^\flat, \vartheta, k, a_E$ real numbers and Ξ a set of characters of $\mathfrak{G}$. Let G_A and Σ_A be the extracted function and the rest (Definition 6.2.2) for D, ℓ, A_0, Ξ and $\varkappa$. Suppose that*

(i) $b > 0$, $a_E \geq 0$, $k > 0$, $m \leq x^{1/2+\varepsilon/2}$ *and* $xm \leq k$;
(ii) *the size conditions and the thin-scale bound of Definition 6.2.1 hold, and* $\mathcal{T}_2 \subset A_0$;
(iii) *the cost conditions hold with* $f = |\Phi|$, *and the rest-size conditions hold with* $n = |H|$, $n_T = |\mathcal{T}|$ *and* $n_{T_2} = |\mathcal{T}_2|$ *(Definition 6.2.3);*
(iv) *the character conditions of Definition 6.2.2 hold with the given* $\varkappa, M, C_o, C_E, \varepsilon$ *and* x;
(v) $\mathrm{gap}(u_0) \leq x^{C\varepsilon}\,\mathrm{Err}$ *for every* $u_0 \in A_0$, *with* Err *the error bracket of Definition 6.2.3 for these* $n, n_T, n_{T_2}, m, b, k, a_E$.

Then $G_A(x) \leq x^{1/2+(C+3)\varepsilon} + \Sigma_A(x)$.

THEOREM 3.9.6 (Alpöge; restated from Theorem 6.3.3). *Let* $\mathscr{L}$ *be a level,* ϕ *a frame, and take a host for* ϕ *(Definition 5.1.2) with its arithmetic data* D *for* $\mathfrak{G} = V_4 \times A$. *Let* $\mathcal{T} \subset A$ *be finite and* $x > 0$, ε *and* C *real, and suppose that the host satisfies the explicit-formula condition* $\mathrm{EF}(\mathscr{L}, \mathcal{T}, x, \varepsilon, C)$ *(Definition 5.3.3). Let* ℓ *be a function from the characters of* $\mathfrak{G}$ *to* $\mathbb{C}$ *such that* $\ell(\theta) = L'(1,\chi)/L(1,\chi)$ *whenever* $\theta \neq \theta_0$ *and* χ *is a primitive Dirichlet character with* $\chi(n) = \theta_D(n)$ *for all* $n \geq 1$. *Let* Ξ *be the set of characters* θ *of* $\mathfrak{G}$ *for which some primitive Dirichlet character* χ *with* $\chi(n) = \theta_D(n)$ *for all* $n \geq 1$ *lies in* $\mathrm{Kept}(\mathscr{L})$, *let* A_0 *be the core of* ϕ, *and let* G_A *and* Σ_A *be the extracted function and the rest (Definition 6.2.2) for* D, ℓ, A_0, Ξ *and* $\varkappa = (3o(\mathscr{L}) + 1)^{-1}$. *Then:*

(a) $G_A(y) = G_{\mathscr{L}}(y)$ *for every* $y > 0$;
(b) $\Sigma_A(x) \leq 4o(\mathscr{L})\, x^{1/2+\varepsilon}$.

In the core bound the gap is written through the balance identity, the arguments a_τ/w are moved to x, and averages over the shifts $u_0 \in A_0$ and the top places $\tau \in \mathcal{T}$ project onto the characters trivial on the core. The kept characters come out with the positive weights 1 and $(3o+1)^{-1}$. This is why an upper bound for the gap gives an upper bound for G_A. Theorem 3.9.6 then identifies G_A with $G_{\mathscr{L}}$ for a host that satisfies the explicit-formula condition, and bounds the rest by the pointwise bounds on the bounded characters. The three results combine as follows. Since $e_{\mathrm{div}} \leq e_D$ and $e_{\mathrm{loss}} \leq e_L$, Theorem 3.9.3 bounds the gap by $x^{(e_D+e_L+3)\varepsilon}\mathrm{Err}$. Theorem 3.9.5 with $C = e_D+e_L+3$ gives $G_A(x) \leq x^{1/2+(e_D+e_L+6)\varepsilon} + \Sigma_A(x)$. Theorem 3.9.6 then gives

$$G_{\mathscr{L}}(x) \leq x^{1/2+(e_D+e_L+6)\varepsilon} + 4o\, x^{1/2+\varepsilon} \leq x^{1/2+(e_D+e_L+7)\varepsilon}$$

for large $x \in \mathrm{Win}$. The last step absorbs the term $4o\, x^{1/2+\varepsilon}$, and for this it uses $e_D + e_L \geq 0$, which holds because $e_D, e_L \geq 0$ by the window conditions.

REMARK 3.9.7 (the square-root gain). The factorial comparison alone yields Chebyshev bounds of the right order, without a square-root error. Here the gain comes from the dependence of the slope-record errors on the number of theta lifts. After normalisation, one part contributes x/m and

another contributes m (Proposition 9.3.1). Their balance is the source of the exponent $1/2$, as explained in Remark 9.4.3.

The bookkeeping behind Remark 3.9.7 is elementary. In it $m \approx \sqrt{x}$ is chosen.

PROPOSITION 3.9.8 (the bookkeeping behind $m \approx \sqrt{x}$; restated from Proposition 9.3.1). *Let $x \geq 1$, $\varepsilon > 0$, $b > 0$, $k > 0$ and $a_E \geq 0$ be real numbers and n, n_T, n_{T_2}, m positive integers. Suppose that $a_E \leq 2xb$, $n_T \geq x^{-\varepsilon/8} n/2$, $x^{1/2-\varepsilon} \leq m \leq x^{1/2+\varepsilon/2}$, $k \leq n^2 x^{1/2-\varepsilon}$, $xm \leq k$ and $n_{T_2} \leq n$. Then the error bracket* Err *of Definition 6.2.3 satisfies*

$$\frac{\mathrm{Err}}{n_T m b} \leq \frac{2x}{m} + 2 + 4m\, x^{9\varepsilon/8} + 2x^{1/2+9\varepsilon/8} \leq 10\, x^{1/2+13\varepsilon/8}.$$

The constraints on m and k in Proposition 3.9.8 are hypotheses; for the families of Theorem 3.8.3 they are part of the ties.

3.10. Black box: the windowed Landau criterion

The analytic end of the argument is a converse to the explicit formula. Landau's theorem on Laplace–Stieltjes integrals with a monotone integrator (Theorem 4.3.2) gives such a converse when the bound $G_F(y) \leq Cy^{\sigma_0}$ is assumed for all large y. The version used here assumes it only on windows.

THEOREM 3.10.1 (Alpöge; restated from Theorem 4.3.3). *Let $F = ((\chi_i, a_i))_i$ be a character family, and let $0 < \sigma_0 < 1$, $\delta > 0$ and $C \in \mathbb{R}$. Suppose that for every Y_0 there is $Y \geq Y_0$ with $G_F(y) \leq Cy^{\sigma_0}$ for all $y \in [Y, Y^{1+\delta}]$. Then $L(s, \chi_i) \neq 0$ for every i and every s with $\operatorname{Re} s > \sigma_0$, where $s \neq 1$ when χ_i is the character modulo 1.*

THEOREM 3.10.2 (Alpöge; restated from Theorem 4.4.3). *Let $\mathscr{L}$ be a level and $\lambda > 1$. Then $\mathscr{L}$ satisfies the window criterion at λ.*

What they say. An upper bound x^{σ_0} for a positive combination of smoothed Chebyshev functions, on windows $[Y, Y^{1+\delta}]$ with Y arbitrarily large, excludes zeros with real part greater than σ_0 for every character in the combination. Theorem 3.10.2 is the case of the combination $G_{\mathscr{L}}$ of a level, with $\delta = \lambda - 1$ and $C = 1$.

Why they are needed. They convert the Chebyshev inequality into zero-freeness. Two features are used. Only an upper bound is assumed. This suffices because, with positive weights, the residues of the relevant Laplace transform at real zeros in $(0, 1)$ all have one sign, so a hypothetical real zero cannot be cancelled by another (Remark 4.3.7). The bound is also assumed only on windows. The classical form would not be enough, because the Chebyshev inequality is proved only on the windows that the Sha bound provides (Remark 4.3.8). Any exponent $\lambda > 1$ is accepted, and the argument uses $\lambda = 3/2$.

Where they are proved. Chapter 4. Theorem 3.10.1 is proved in outline, by comparing two estimates for the Taylor coefficients of a Laplace transform at a point far to the right; its steps are elementary complex analysis. Theorem 3.10.2 is deduced from it there by a complete proof. Whether the windowed criterion is new as stated is left open (Question 4.3.9).

3.11. Black box: from a half-plane to the critical line

For a primitive Dirichlet character χ modulo N, let $\mathfrak{a} \in \{0,1\}$ be given by $\chi(-1) = (-1)^{\mathfrak{a}}$, and put $\xi(s,\chi) = (N/\pi)^{(s+\mathfrak{a})/2}\Gamma((s+\mathfrak{a})/2)L(s,\chi)$. We write (H) for the following hypothesis: $L(s,\chi') \neq 0$ for every primitive Dirichlet character χ' and every s with $\operatorname{Re} s > 1/2$, where $s \neq 1$ when χ' is the character modulo 1.

THEOREM 3.11.1 (Riemann, Hecke [**Dav00, Rie59**]; restated from Theorem 4.5.1). *Let χ be a primitive Dirichlet character modulo N. Then there is $\epsilon(\chi) \in \mathbb{C}$ with $|\epsilon(\chi)| = 1$ and $\xi(s,\chi) = \epsilon(\chi)\xi(1-s,\overline{\chi})$ for all $s \neq 0, 1$.*

PROPOSITION 3.11.2 (restated from Proposition 4.5.2). *Let χ be a Dirichlet character modulo N, induced by the primitive character χ^* modulo N^*.*

- *(a) For $s \neq 1$, and for every $s \in \mathbb{C}$ if χ is nonprincipal, $L(s,\chi) = L(s,\chi^*)\prod_{p|N}(1-\chi^*(p)p^{-s})$. In particular, for such s with $\operatorname{Re} s \neq 0$, $L(s,\chi) = 0$ if and only if $L(s,\chi^*) = 0$.*
- *(b) The L-function of the character modulo 1 is ζ on $s \neq 1$.*
- *(c) Under (H), $L(s,\chi) \neq 0$ for every s with $\operatorname{Re} s > 1/2$, where $s \neq 1$ when χ is principal.*
- *(d) Under (H), every zero of $L(s,\chi)$ with $0 < \operatorname{Re} s < 1$ has $\operatorname{Re} s = 1/2$, and every zero $s \neq 1$ of ζ that is not in $\{-2,-4,-6,\dots\}$ has $\operatorname{Re} s = 1/2$.*

What they say. Zero-freeness in $\operatorname{Re} s > 1/2$ for primitive characters implies the same for all characters, and the functional equation then places the zeros of the critical strip on the line $\operatorname{Re} s = 1/2$.
Why they are needed. The argument produces (H). The statement of Theorem 3.1.1 concerns all Dirichlet characters and the critical line.
Where they are proved. Chapter 4. Theorem 3.11.1 is cited, and Proposition 3.11.2 is proved completely. The step to the Riemann hypothesis needs neither all Dirichlet characters nor the property $|\epsilon(\chi)| = 1$ of the root numbers. The statement about ζ uses only (H) for the character modulo 1 and Riemann's functional equation, and where the functional equation of a primitive character is used the root number multiplies zero (Remark 4.5.3).

3.12. Proof of the main theorem

All the boxes are now stated. They meet in one statement, which turns the two existence results into zero-freeness for primitive characters.

THEOREM 3.12.1 (Alpöge; restated from Theorem 12.4.1). *Let $\lambda > 1$. Assume that the top existence statement (Definition 12.3.2) holds and that $\mathscr{R}_\psi$ is supplied on windows of exponent λ for every primitive Dirichlet character ψ. Then $L(s,\psi) \neq 0$ for every primitive ψ and every s with $\operatorname{Re} s > 1/2$, except $s = 1$ when ψ is trivial.*

Theorem 3.12.1 is proved completely in Chapter 12, from the Chebyshev inequality on windows and the windowed Landau criterion. The argument is short, and it is the place where the boxes of Sections 3.9 and 3.10 are joined to those of Sections 3.6 and 3.8, so we go through it here.

PROOF OF THEOREM 3.12.1 FROM THE BLACK BOXES. Fix a primitive ψ, and let $\mathcal{C} = \mathcal{C}_{\mathscr{R}_\psi}(\lambda)$ be the class of the requirement of ψ at λ.

We check that the class meets the hypotheses of the top existence statement. Every member of a family in $\mathcal{C}$ beyond the threshold meets $\mathscr{R}_\psi$. The Sha budget $(t, b^{1/2})$ is part of $\mathscr{R}_\psi$, so every family in $\mathcal{C}$ satisfies the Sha hypothesis. The explicit-formula condition $\mathrm{EF}(\mathscr{L}_\psi, \mathcal{T}, x, \varepsilon, 4)$ is also part of $\mathscr{R}_\psi$, so every family in $\mathcal{C}$ satisfies the uniform explicit-formula condition for $\mathscr{L}_\psi$, with $C = 4$. By assumption $\mathscr{R}_\psi$ is supplied on windows of exponent λ.

The top existence statement therefore applies to λ and $\mathscr{R}_\psi$, and $\mathcal{C}$ supplies tied families: there are e_D, e_L and $\varepsilon_1 > 0$ such that for every $\varepsilon \in (0, \varepsilon_1)$ some family $\mathfrak{F}$ at ε in $\mathcal{C}$ is tied at every large $x \in \mathrm{Win}$, with these e_D, e_L and some constants. The numbers e_D and e_L do not depend on ε.

Now let s satisfy $\operatorname{Re} s > 1/2$, with $s \neq 1$ when ψ is trivial, put $\mu = \min(\operatorname{Re} s - 1/2, 1/2) > 0$, and choose $\varepsilon \in (0, \varepsilon_1)$ with $(e_D + e_L + 7)\varepsilon \leq \mu/2$. By the previous paragraph there is a family $\mathfrak{F}$ at this ε in $\mathcal{C}$ and there are real numbers c_W, c_S, e_K, C_L and x_0 such that $\mathfrak{F}$ is tied with constants $e_D, e_L, c_W, c_S, e_K, C_L$ at every $x \in \mathrm{Win}$ with $x \geq x_0$. The family $\mathfrak{F}$ satisfies the uniform explicit-formula condition for $\mathscr{L}_\psi$, since it lies in $\mathcal{C}$. By Theorem 3.9.1, applied to $\mathscr{L}_\psi$ and $\mathfrak{F}$, there is x_1 such that

$$G_{\mathscr{L}_\psi}(x) \leq x^{1/2 + (e_D + e_L + 7)\varepsilon} \leq x^{\sigma_0}, \qquad \sigma_0 = \tfrac{1}{2} + \tfrac{\mu}{2},$$

for every $x \in \mathrm{Win}$ with $x \geq x_1$.

It remains to apply the Landau criterion. Since $\mathfrak{F}$ lies in $\mathcal{C}$, the set Win has windows of exponent λ. Hence the set $\{x : G_{\mathscr{L}_\psi}(x) \leq x^{\sigma_0}\}$, which contains every large element of Win, has windows of exponent λ, and $\sigma_0 \in (0,1)$. By Theorem 3.10.2, $\mathscr{L}_\psi$ satisfies the window criterion at λ, so $L(s', \chi) \neq 0$ for every $\chi \in \mathrm{Kept}(\mathscr{L}_\psi)$ and every s' with $\operatorname{Re} s' > \sigma_0$, where $s' \neq 1$ when χ is the trivial character. The character ψ is kept by $\mathscr{L}_\psi$, and $\operatorname{Re} s > \sigma_0$ because $\mu \leq \operatorname{Re} s - 1/2$. Hence $L(s, \psi) \neq 0$. □

The construction depends on the order of the quantifiers in this proof. The exponents e_D, e_L and the window exponent λ are fixed first. The zero s is given next, and it determines ε. The family, its window set and all remaining

constants depend on ε. The bound $x^{1/2+(e_D+e_L+7)\varepsilon}$ is used only on the window set of that one family.

Proof of Theorem 3.1.1. We chain the black boxes.

(1) By Theorem 11.1.4, weak local Langlands maps exist and the truncated local–global compatibility of Definition 11.1.3 holds for them. This is Theorem 3.4.1.
(2) By Theorem 10.5.1, whose automorphy input is step (1) (Remark 11.3.1), every primitive Dirichlet character ψ has Sha bounds on windows with constant 401. This is Theorem 3.5.1.
(3) By Theorem 12.1.2, applied to each primitive ψ with step (2) as its hypothesis, $\mathscr{R}_\psi$ is supplied on windows of exponent $3/2$ for every primitive ψ. This is Theorem 3.6.2; the good pairs of Theorem 3.6.1 are used in its proof.
(4) By Theorem 12.3.4(c), whose proof uses the covolume bound (Theorem 8.2.3) and the maps of step (1), some class of families carries admissible-family data. By Theorem 12.3.5, applied to that class, the top existence statement holds; Green's parametrization (Theorem 12.2.1) enters its proof. These are Theorems 3.7.1, 3.8.1, 3.8.2 and 3.8.3.
(5) By Theorem 12.4.1 with $\lambda = 3/2$, whose two hypotheses are steps (4) and (3), $L(s,\psi) \neq 0$ for every primitive ψ and every s with $\operatorname{Re} s > 1/2$, except $s = 1$ when ψ is trivial. This is Theorem 3.12.1, whose proof above uses Theorems 3.9.1 and 3.10.2. The conclusion is the hypothesis (H) of Section 3.11.
(6) By Proposition 4.5.2, the nonvanishing extends to all Dirichlet characters, and the functional equation (Theorem 4.5.1) puts the zeros in $0 < \operatorname{Re} s < 1$ on the line $\operatorname{Re} s = 1/2$, for ζ and for every $L(s,\chi)$. In detail, let χ be a Dirichlet character. Part (c) of Proposition 3.11.2 gives $L(s,\chi) \neq 0$ for $\operatorname{Re} s > 1/2$, with $s = 1$ excluded when χ is principal. Part (d) gives that every zero of $L(s,\chi)$ with $0 < \operatorname{Re} s < 1$ has real part $1/2$, and that every zero of ζ outside $\{-2,-4,-6,\dots\}$ has real part $1/2$. The point $s = 1$ is a pole of ζ and not a zero.

This proves Theorem 3.1.1 relative to the black boxes. Section 3.14 records their status. □

Remark 3.12.2 (where the content is). The deduction of Theorem 3.1.1 from its ingredients carries no substantial argument of its own: the proof above is a chain of implications. The substantive inputs are the windowed Landau criterion (Theorem 4.3.3), the frame theorem (Theorem 7.4.2), the covolume bound (Theorem 8.2.3), the Sha bound on windows (Theorem 10.5.1, with Theorem 11.1.4), the supply of hosts (Theorems 5.4.8 and 12.1.2), and the existence of packages and tied families (Theorems 12.3.4 and 12.3.5). The analytic core that turns these into

$G_{\mathscr{L}}(x) \le x^{1/2+O(\varepsilon)}$ on windows is the gap bound, Theorem 9.2.1, with the core bound, Theorem 6.3.4. Once the gap bound is in place, the Chebyshev inequality on windows (Theorem 9.4.1) is linear arithmetic.

Remark 3.12.3 (the constants of the chain). Five numbers pass from box to box, and each has one origin.

(1) The bound $\varepsilon < 1/52$ is fixed once, in the definition of a scale. Only small values of ε are used in the end.
(2) The constant 401 of the Sha bound makes the budget cover $\lfloor x^{401} \rfloor$ Tate twists; the number of twists needed is the rank parameter t, which satisfies $t \le x^{400}$ (Remark 10.5.2).
(3) The window exponent is 2 in the Sha bound, and 3/2 in the supply of hosts (step (3)) and in the application of the Landau criterion (step (5)). How the 2 becomes 3/2 is Question 12.1.3. The exponent 3/2 does not come from Green's theorem, which enters at step (4), between the two. It is the window exponent of the supply of hosts, and Green's theorem is not an analytic input (Remark 12.2.3). The Landau criterion accepts any exponent greater than 1.
(4) The constant 4 in $\mathrm{EF}(\mathscr{L}_\psi, \mathcal{T}, x, \varepsilon, 4)$ is the constant of the pointwise clause of a good pair (Definition 5.4.2).
(5) The exponent $1/2 + (e_D + e_L + 7)\varepsilon$ has $e_D = 2$ and $e_L = 0$ for the families of Theorem 3.8.3, so it is $1/2 + 9\varepsilon$. Of the 7, the gap bound contributes 3, the core bound 3, and the absorption of the rest term $4o\,x^{1/2+\varepsilon}$ one more.

3.13. Chebyshev's comparison and the function-field model

Chebyshev compared the archimedean size of a factorial, given by Stirling's formula, with its prime factorisation, given by Legendre's formula. The comparison bounds a prime sum because the two calculations concern the same integer. Here the objects are periods of successive theta lifts of algebraic automorphic forms. Their archimedean sizes contain a Stirling term and a CM period; their finite sizes record divisibility at primes inert in an imaginary quadratic field. A map between lattices of forms and its slope estimates take the place of the factorial identity. Table 3.1 records the parallel and the point at which the present argument gains its sharper error. Each row of its right column lies inside one of the black boxes. The first five rows lie in the Chebyshev inequality on windows, through Proposition 3.9.2 and Theorems 3.9.3 and 3.9.4. The row on multiplicity lies in Proposition 3.9.8, and the last two rows lie in Theorems 3.9.1 and 3.10.1.

Chapter 2 proves the Riemann hypothesis for a function field over a finite field by the method of Stepanov and Bombieri, and Table 3.2 places each part of that proof beside the black box that plays its role. The table is a guide to roles. It asserts no implication between its two columns.

	Chebyshev 1852	The present argument
Object	the integer $n!$ (or $\binom{2n}{n}$)	periods of m successive theta lifts of algebraic forms of weight $a \approx xb$ on definite unitary groups (Definitions 7.1.4 and 7.2.2)
Archimedean size	Stirling: $\log n! = n\log n - n + O(\log n)$	Stirling's term $m(a\log a - a)$ minus $a\log\Omega$, with $\log\Omega = \gamma + \log(\pi/v) - L'(1,\chi_{-d})/L(1,\chi_{-d})$ (clause (ii) of Definition 7.3.2, Remark 7.5.2)
p-adic size	Legendre: $v_p(n!) = \sum_i \lfloor n/p^i \rfloor$	Legendre-type gains $2\log p \sum_i (a/p^i - m_\tau)_+$ at primes p inert in $K = \mathbb{Q}(\sqrt{-d})$, over i with $i\,\mathrm{Frob}_p \notin \Phi$ (Definition 6.1.3, Remark 6.1.5)
How the sides are linked	unique factorisation: an identity	inequalities between hermitian lattices: an archimedean upper link and a slope clause of the shape of the frame theorem; $\mu(R^\vee)$ cancels (Theorem 7.4.2, Remark 9.2.2)
Characters	none	characters of $\mathrm{Gal}(E/\mathbb{Q})$; the balance identity and the shares give positive weights 1, $(3o+1)^{-1}$ (Propositions 6.1.4 and 6.3.1)
Multiplicity	one factorial	a chain of $m \approx \sqrt{x}$ lifts; errors $\asymp x/m$ and m balance at $\sqrt{x}$ (Proposition 9.3.1)
Output	$c_1 x \le \mathrm{Ch}(x) \le c_2 x$, two-sided, $c_1 < 1 < c_2$	one-sided $G_{\mathscr{L}}(x) \le x^{1/2+O(\varepsilon)}$, only on windows of scales (Theorem 9.4.1)
To zeros	not reached: the constants are not 1	the windowed Landau criterion (Theorem 4.3.3)

TABLE 3.1. Chebyshev's factorial comparison and the corresponding parts of the present argument. The slope comparison and the inert-prime gain are developed in Chapters 7–8.

Three points of the comparison need care. Section 2.13 of Chapter 2 lists further places where the analogy breaks.

The first point is the direction of the inequality. The Chebyshev inequality on windows is an upper bound for $G_{\mathscr{L}}$. In the model, an upper bound for the centred sum $G(T)$ corresponds to a *lower* bound for the point counts (Proposition 2.6.10), so the signs are opposite to those of the Stepanov–Bombieri upper bound for point counts. The inequality that the assembled argument proves is therefore the counterpart of the hypothesis of Theorem 2.6.9. It does not correspond to the inequality that the Stepanov–Bombieri method supplies. In Chapter 2 the lower bound is obtained from the upper bound by Galois averaging (Theorem 2.4.6). In the assembled argument the upper bound for $G_{\mathscr{L}}$ is proved directly, and the Landau criterion does the rest.

The second point is the auxiliary function. The Stepanov–Bombieri proof constructs a section with prescribed vanishing and counts its zeros against its poles. The assembled argument constructs no space of small sections for its slope inequality. The inequality is applied to the period map between two given lattices (Remark 7.2.6).

The third point is the sign-only argument. The second argument of Chapter 2, through the identity $qN = M^k + R$, is conditional on a hypothesis that is not proved there (Theorem 2.10.2, hypothesis (H2)). It gets its bound

In Chapter 2	In the assembled argument	Black box
The centred smoothed sum $G(T)$ of the point counts (Definition 2.6.7)	the weighted Chebyshev function $G_{\mathscr{L}}$	Section 3.3
A bound on a set containing arbitrarily long intervals forces $\lvert\alpha_i\rvert = \sqrt{q}$ (Proposition 2.6.1, Theorem 2.6.9)	a bound on windows excludes zeros with $\operatorname{Re} s > \sigma_0$	Theorem 3.10.1
The Stepanov–Bombieri inequality, from an auxiliary section vanishing at the Frobenius places (Theorem 2.3.8)	the gap bound, from a slope inequality for period maps; no auxiliary section is constructed	Theorems 3.9.3, 3.9.4
The error $(g+1)q/M + 2gM$, balanced at $M = \sqrt{q}$ (Proposition 2.3.9, Theorem 2.3.12)	the errors x/m and m, balanced at $m \approx \sqrt{x}$	Proposition 3.9.8
Averaging over a Galois group to pass between upper and lower bounds (Theorem 2.4.6)	Fourier analysis on $\mathrm{Gal}(E/\mathbb{Q})$, with positive weights for the kept characters	Proposition 3.9.2, Theorem 3.9.5
The functional equation places the reciprocal roots on $\lvert\alpha\rvert = \sqrt{q}$ (Theorem 2.5.3)	the functional equation places the zeros on $\operatorname{Re} s = 1/2$	Theorem 3.11.1, Proposition 3.11.2
The M-th power map (Frobenius), the source of multiplicity	the chain of m theta lifts; $\mathbb{Z}$ has no such endomorphism (item (B1) of Section 2.13)	Chapter 7, Proposition 3.9.8
No counterpart	hosts on windows, the Sha bound, Galois representations, the covolume bound, packages and tied families	Sections 3.4–3.8

TABLE 3.2. The function-field model of Chapter 2 beside the black boxes of this chapter.

from a sign that prevents cancellation. No identity of this form occurs among the black boxes of this chapter. In its place the assembly has the positivity of the weights of a level. Because the weights are positive, the contributions of real zeros cannot cancel, and a one-sided bound suffices (Remark 4.3.7).

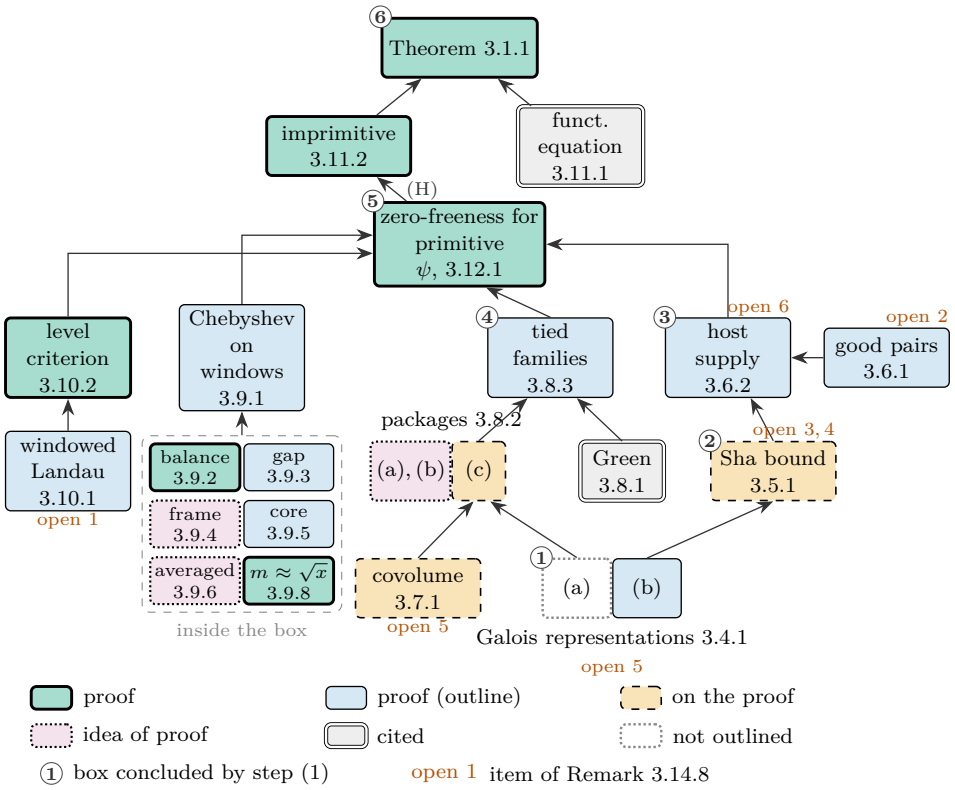


FIGURE 3.2. The black boxes of this chapter on the dependency structure of the proof of Theorem 3.1.1, filled by the kind of proof that Table 3.3 records (Remark 3.14.1). An arrow points from a box to a box whose proof uses it, following steps (1)–(6) of that proof (compare Figure 3.1); the six results restated inside the Chebyshev box are grouped without their internal arrows. A circled number k marks the box that step (k) of the proof of Theorem 3.1.1 concludes, and a tag "open k" is item (k) of Remark 3.14.8. "Proof" means complete relative to the results quoted; nothing here establishes the boxes of the other kinds. Schematic.

3.14. The status of the inputs

The proof of Theorem 3.1.1 in Section 3.12 is complete relative to its black boxes. This section records what this book provides for each box, and the caveats attached to them in the chapters where they are treated.

REMARK 3.14.1 (attributions and kinds of proof). The conventions are those of Section 4.1. A statement marked (Alpöge) is a statement of Alpöge's argument. Alpöge claims it, and it is stated to be formalized in his Lean archive [**Alp26a**]; the archive was not built or checked for this book, and

nothing here establishes the statement (Remark 1.7.1). This book states it and explains it. A statement carrying another name is cited from the literature, and an unattributed statement is proved in the text. Each proof block is of one of four kinds.

(1) A *proof* is complete.
(2) A *proof (outline)* gives Alpöge's route with its steps named.
(3) A block *on the proof* gives the inputs of the proof and the shape of its route, without an argument.
(4) An *idea of proof* is an explanatory account of a standard method. It is not Alpöge's proof, and nothing later depends on it.

Not every box is closed by a complete proof in the later chapters. Figure 3.2 places every box on the dependency structure of the proof of Theorem 3.1.1. Only the boxes marked "proof" in Table 3.3 are proved completely in this book. For the others the book gives what the table says, and for a proof it can only refer to the archive, which it has not checked, or to the cited source.

Remark 3.14.2 (schematic statements). The definitions of packages, size data, the upper and lower links, families, the covolume condition, host supply, the links and admissible-family data (Definitions 7.2.3, 7.2.4, 7.5.1, 7.6.1, 8.1.1, 8.2.1, 12.1.1, 12.3.1 and 12.3.2) are schematic: each lists the parts of the data that the argument uses and the conditions on them that the text uses, and omits the remaining compatibility conditions. So not every black box can be checked against its definitions line by line. Theorems 3.6.2, 3.7.1, 3.8.2, 3.8.3, 3.9.1 and 3.12.1 are statements about data given by schematic definitions, and are schematic to the same extent; where the later chapters give an informal statement in standard terms, it is repeated here under "what it says". The other black boxes of this chapter are stated in full, apart from two clauses of Definition 11.1.1 that are summarised there.

Remark 3.14.3 (no hypothesis is stated). Theorem 3.1.1 is stated without hypothesis. In particular, none of the following appears as an assumption of a black box: a bound of Riemann type for the auxiliary characters, local–global compatibility for the Galois representations, or the local Langlands correspondence. This is a statement about the form of the argument. The proofs of the boxes not marked "proof" in Table 3.3 are not checked in this book, so nothing here establishes that the boxes hold (Remark 1.7.1).

(1) The analytic input on the twists is the explicit-formula condition at the single scale x, which is polylogarithmic in the conductor. It follows from the averaging over auxiliary primes, and neither it nor the zero-free boxes of the real twists has the strength of the Riemann hypothesis near $\operatorname{Re} s = 1/2$ (Remark 5.5.1). The classical inputs to the averaging are listed in Remark 5.5.2; none assumes a hypothesis of Riemann type.

Black box	Source	Chapter	What the book gives
Theorem 3.4.1	Alpöge	11	proof (outline) of (b); (a) not outlined
Theorem 3.5.1	Alpöge	10	on the proof
Theorem 3.6.1	Alpöge	5	proof (outline)
Theorem 3.6.2	Alpöge	12	proof (outline)
Theorem 3.7.1	Alpöge	8	on the proof; not reproduced
Theorem 3.8.1	Green	12	cited
Theorem 3.8.2	Alpöge	12	idea of proof, on the proof
Theorem 3.8.3	Alpöge	12	proof (outline)
Theorem 3.9.1	Alpöge	9	proof (outline)
Proposition 3.9.2	—	6	proof
Theorem 3.9.3	Alpöge	9	proof (outline)
Theorem 3.9.4	Alpöge	7	idea of proof
Theorem 3.9.5	Alpöge	6	proof (outline)
Theorem 3.9.6	Alpöge	6	idea of proof
Proposition 3.9.8	—	9	proof
Theorem 3.10.1	Alpöge	4	proof (outline)
Theorem 3.10.2	Alpöge	4	proof
Theorem 3.11.1	Riemann, Hecke	4	cited
Proposition 3.11.2	—	4	proof
Theorem 3.12.1	Alpöge	12	proof

TABLE 3.3. The black boxes of this chapter, the chapter in which each is treated, and the kind of proof given there (Remark 3.14.1). The proofs marked "proof" are complete relative to the results they quote.

(2) The Galois-theoretic inputs are the two parts of Theorem 3.4.1, which is a statement of the argument and not an assumption (Remark 11.1.5). Truncated compatibility asks for less than the published theorems, and the local Langlands correspondence is replaced by the existence of weak local Langlands maps (Remark 11.1.2).
(3) Every result of Chapter 10 is stated without hypothesis; its only input beyond classical Galois cohomology and the theorems stated there is Theorem 3.4.1, part (b) (Remark 11.3.1).

REMARK 3.14.4 (the scan of the archive for assumptions). The 4081 naive hits of a scan of the archive for assumptions are not assumptions. The scans find no axiom declaration, no proof placeholder and no unsafe

or kernel-bypassing construct. The hits are macros, imports and metaprogramming commands. The scans also count 37 declarations with conclusion *False*, 124 uses of kernel-checked decision and 11 declarations of a trivially true type, of which only a sample was read. The scans are static: the archive was not built (Section A.9).

Remark 3.14.5 (imported results and the forms in which they are used). The black boxes cite the literature at the following places.

(1) The functional equation (Theorem 3.11.1) and Green's parametrisation (Theorem 3.8.1) are used as stated. Landau's theorem (Theorem 4.3.2) is the model for the windowed criterion and is not itself used.
(2) The covolume bound rests on cited theorems of arithmetic geometry. Each is used in a specific formulation, not necessarily in the generality of the theorem it is named after, ranging from near-general to special or weakened (Remark 8.3.4). Each formulation is implied by the cited theorem, with two qualifications recorded there: for a nef line bundle the asymptotic Hilbert–Samuel formula is due to Moriwaki, and for the bigness theorem the implication needs an axiomatic intersection pairing to be the arithmetic intersection pairing; this is not checked, and the implication is not claimed.
(3) Bost's slope inequality (Theorem 7.2.5) is one step of the proof of the frame theorem; the slope clause of a slope record is not Bost's inequality (Remark 7.4.3).
(4) Poitou–Tate duality, the finiteness of local Galois cohomology and local Tate duality are used in the Sha bound. The existence of packages and of admissible-family data draws on the Siegel–Weil formula, Howe duality, results of Mínguez on theta integrals, the Casselman–Wallach theory and Minkowski's bound on finite subgroups of $\mathrm{GL}_n(\mathbb{Z})$ (Definition 12.3.2 and the discussion of Theorem 12.3.4).
(5) Some named results are not used. Clozel's purity lemma, the main theorem of complex multiplication and the theorem of Deligne and Ribet do not enter the Sha bound (Remark 10.5.3), and no height of a CM point and no Colmez-type sum of logarithmic derivatives enters the main term (Remark 7.5.3).

Remark 3.14.6 (transcribed theorems in the archive). The statements of the archive that transcribe published theorems are neither all genuine nor all hollow. Of 30 sampled, 27 are genuine statements about Mathlib objects. Of the 195 theorems that produce them, 182 cite theorems proved inside the archive.

Remark 3.14.7 (imports in the archive). Long import chains do not show that the named theorems of the archive are logically used. All 175,187

modules are import-reachable from the main file, so import reachability says nothing about term dependency, which is unknown.

Remark 3.14.8 (points left open in the later chapters). The chapters that treat the black boxes leave the following points open, and the reader of the assembly should know of them.

(1) Whether the windowed Landau criterion is new as stated is a question (Question 4.3.9).
(2) In the supply of good pairs, the constant 4 of the pointwise clause must be uniform in the exponent Δ. The route described in Chapter 5 gives a constant depending on Δ, and the sharper bound on $L'(1,\chi)/L(1,\chi)$ that uniformity requires is not reconstructed there.
(3) The constant 401 leaves a margin of one in the exponent over the x^{400} twists that are needed; no step that uses the margin is identified (Remark 10.5.2).
(4) The bounds for the eigenspaces of the wrong parity inside the Sha bound are of the type of Vandiver's and Greenberg's conjectures, and Chapter 10 records that no published result of this strength is known (Remark 10.4.10). A vanishing theorem of Kato type used there (Theorem 10.4.6) is stated without an outline of proof, and whether the twist it needs lies in the range that Kato treats is not settled there.
(5) The proof of the covolume bound is not reproduced, the proof of part (a) of Theorem 3.4.1 is not outlined, and the compact-group statement used in the construction of the Galois representations (Proposition 11.2.3) is stated without an outline of proof.
(6) The step from the window exponent 2 of the Sha bound to the exponent 3/2 of the supply of hosts is Question 12.1.3; see also Remark 12.1.4.

CHAPTER 4

Smoothed Chebyshev functions and a windowed Landau criterion

This chapter supplies the analytic end of the argument. Its main result is the windowed Landau criterion, Theorem 4.3.3: a one-sided bound $G_F(y) \le Cy^{\sigma_0}$ for a positive combination G_F of smoothed twisted Chebyshev functions, assumed only on windows $[Y, Y^{1+\delta}]$ with Y arbitrarily large, forces the L-functions in the combination to have no zero in $\operatorname{Re} s > \sigma_0$. Theorem 4.4.3 specialises the criterion to the combinations $G_{\mathscr{L}}$ attached to levels, which are the functions that the later chapters bound, and Proposition 4.5.2 passes from primitive characters and a half-plane to all Dirichlet characters and the critical line. The proof of Theorem 3.1.1 in Chapter 3 (Section 3.12) uses both as black boxes. Theorem 4.4.3 enters through Theorem 12.4.1 of Chapter 12, the step that turns growth bounds on windows into zero-freeness. Proposition 4.5.2, with the functional equation (Theorem 4.5.1), is the last step. Section 4.1 fixes the conventions for statements and proofs used in Chapter 3 and in Chapters 4–12.

The argument proves zero-freeness through one-sided growth. Its analytic end is a converse to the explicit formula: if a suitable smoothed prime sum is bounded above by x^{σ_0}, then the corresponding L-functions have no zero in $\operatorname{Re} s > \sigma_0$. Two features of the converse are used throughout. It needs only an *upper* bound, provided the prime sum is a combination of twisted sums with positive weights. It needs the bound only on *windows* $[Y, Y^{\lambda}]$ for arbitrarily large Y, not for all large x.

4.1. Conventions for Chapters 3–12

The conventions of this section hold in Chapter 3 and in Chapters 4–12. A statement marked (Alpöge) is a statement of Alpöge's argument. Alpöge claims it, and it is stated to be formalized in his Lean archive [**Alp26a**]; the archive was not built or checked for this book, and nothing here establishes the statement (Remark 1.7.1). We state it and explain it. A statement carrying another name is cited from the literature, and an unattributed statement is proved in the text.

Each proof block carries one of four labels.

(i) *Proof*: a complete proof.
(ii) *Proof (outline)*: Alpöge's route, with its steps named; steps that are standard, or that are cited theorems, are said to be.

(iii) *On the proof*: the inputs of the proof and the shape of its route, without an argument.
(iv) *Idea of proof*: our explanatory account of a standard method for statements of this kind. It is not Alpöge's proof, and nothing later depends on it.

Some definitions in Chapters 7, 8 and 12 are marked *schematic*. A schematic definition lists the parts of the data that the argument uses, and the conditions on them that the text uses; the remaining compatibility conditions among the data are omitted. A result about data given by a schematic definition is stated schematically, and an informal statement in standard terms accompanies it.

4.2. Smoothed Chebyshev functions

DEFINITION 4.2.1 (smoothed Chebyshev functions). Let Λ be the von Mangoldt function, γ Euler's constant and $\mathrm{Ch}(t) = \sum_{n \le t} \Lambda(n)$ Chebyshev's function. Let $x, y > 0$ and let χ be a Dirichlet character modulo N.

(1) The *smoothed Chebyshev function* of ζ is

$$S_\zeta(x) = \sum_{1 \le n \le x} \Lambda(n)\Big(\frac{x}{n} - 1\Big),$$

and its *normalised form* is $G_\zeta(x) = S_\zeta(x) - x \log x + (1+\gamma)x$.

(2) The *real twisted function* of χ is

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

(3) The *complex form* for ζ is

$$G(y) = \sum_{n \le y} \Lambda(n) \max\Big(\frac{y}{n} - 1, 0\Big) - y(\log y - 1 - \gamma).$$

(4) The *complex twisted function* of χ is $g_\chi(y) = G(y)$ if χ is principal (we then write $\chi = 1$), and, for $\chi \ne 1$,

$$g_\chi(y) = \sum_{1 \le n \le y} \Lambda(n)\chi(n) \max\Big(\frac{y}{n} - 1, 0\Big) + y\,\frac{L'(1,\chi)}{L(1,\chi)}.$$

The letter ψ is kept for characters throughout, which is why Chebyshev's function is written Ch. For a nonprincipal χ the value $L(1,\chi)$ is nonzero, so the term $yL'(1,\chi)/L(1,\chi)$ is defined. The extra terms in G_ζ, G_χ and g_χ remove the contribution of $s = 1$, as Remark 4.2.3 explains.

PROPOSITION 4.2.2. *Let $y > 0$. Then:*

(a) $G(y) = G_\zeta(y)$;
(b) if $\chi \ne 1$ is a Dirichlet character, then $\operatorname{Re} g_\chi(y) = G_\chi(y)$;
(c) if $y \ge 1$, then $S_\zeta(y) = y \int_1^y \mathrm{Ch}(t) t^{-2}\,dt$.

PROOF. *Proof of (a).* For $1 \le n \le y$ we have $y/n - 1 \ge 0$, so $\max(y/n - 1, 0) = y/n - 1$. This gives $G(y) = S_\zeta(y) - y\log y + (1+\gamma)y = G_\zeta(y)$.
Proof of (b). Since $\Lambda(n)$ and $y/n - 1$ are real, the real part of each summand of $g_\chi(y)$ is $\Lambda(n)\operatorname{Re}\chi(n)(y/n-1)$, and $\operatorname{Re}(yL'(1,\chi)/L(1,\chi)) = y\operatorname{Re}(L'(1,\chi)/L(1,\chi))$. Hence $\operatorname{Re} g_\chi(y) = G_\chi(y)$.
Proof of (c). The function $\mathrm{Ch}(t) = \sum_{n\le t}\Lambda(n)$ is a finite sum on $[1,y]$, so

$$y\int_1^y \mathrm{Ch}(t)t^{-2}\,dt = y\sum_{n\le y}\Lambda(n)\int_n^y t^{-2}\,dt = \sum_{n\le y}\Lambda(n)\Big(\frac{y}{n}-1\Big) = S_\zeta(y). \quad \square$$

REMARK 4.2.3 (what the functions measure). Perron's formula with the kernel $x^s/(s(s-1))$ gives the explicit formula for $\mathrm{Ch}(x)$ [**Dav00, MV07**] in integrated form:

$$S_\zeta(x) = \frac{1}{2\pi i}\int_{(2)}\Big(-\frac{\zeta'}{\zeta}(s)\Big)\frac{x^s}{s(s-1)}\,ds \qquad (x>1).$$

The double pole at $s=1$ contributes exactly $x\log x - (1+\gamma)x$, the pole at $s=0$ contributes $\log 2\pi$, and the trivial zeros contribute $O(1/x)$. Hence

$$G_\zeta(x) = \log 2\pi - \sum_\rho \frac{x^\rho}{\rho(\rho-1)} + O(1/x),$$

the sum running over the nontrivial zeros of ζ and converging absolutely. For $\chi \ne 1$ the term $xL'(1,\chi)/L(1,\chi)$ removes the simple pole of the integrand at $s=1$, and $g_\chi(x) = -\sum_\rho x^\rho/(\rho(\rho-1)) + O(\log x)$, the sum now over the nontrivial zeros of $L(s,\chi)$. Consequently the Riemann hypothesis for $L(s,\chi)$ gives $g_\chi(x) = O(x^{1/2})$, and the Riemann hypothesis gives $G_\zeta(x) = O(x^{1/2})$. The rest of this chapter is about the converse.

Figure 4.1 shows $G_\zeta(x)/\sqrt{x}$ for $x \le 10^7$ against the size that the Riemann hypothesis predicts.

4.3. The windowed Landau criterion

The criterion is stated for finite positive combinations of the functions g_χ.

DEFINITION 4.3.1 (character families). A *character family* F is a finite list of pairs (χ_i, a_i), where each χ_i is a Dirichlet character modulo some N_i that is either the character modulo 1 or nontrivial, and each $a_i > 0$ is a real *weight*. Its *weighted Chebyshev function* is $G_F(y) = \sum_i a_i \operatorname{Re} g_{\chi_i}(y)$.

The classical model for a one-sided converse is Landau's theorem on Laplace–Stieltjes integrals with a monotone integrator.

THEOREM 4.3.2 (Landau [**Lan05**]). *Let α be a nondecreasing function on $[0,\infty)$ and suppose that $F(s) = \int_0^\infty e^{-su}\,d\alpha(u)$ has a finite abscissa of convergence σ_c. Then F has no analytic continuation to any neighbourhood of $s = \sigma_c$.*

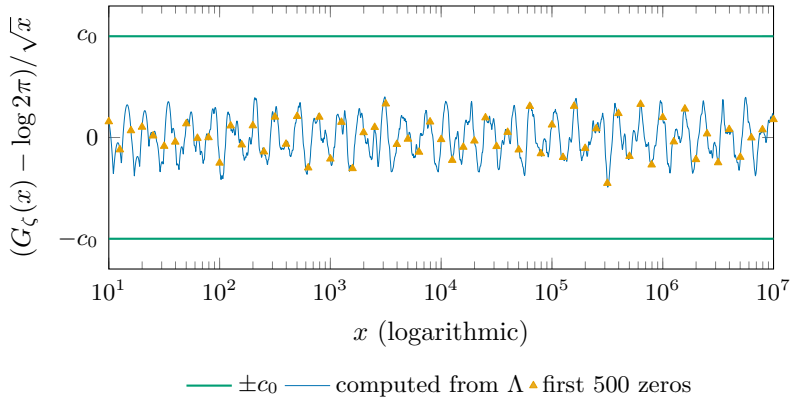


FIGURE 4.1. The function $G_\zeta(x) = \sum_{n \le x} \Lambda(n)(x/n - 1) - x \log x + (1 + \gamma)x$, centred at its constant $\log 2\pi$ and divided by $\sqrt{x}$, for $10 \le x \le 10^7$. The line is computed from Λ, the triangles from the explicit formula with the first 500 zeros of ζ. The band $\pm c_0$, with $c_0 = \sum_\rho 1/|\rho(\rho - 1)|$, is the envelope that the explicit formula gives when every zero lies on the critical line. Numerics at finite x do not bear on the asymptotic statement.

See also [**Wid41**] and, for Dirichlet series, [**MV07**]. Applied to $\int_1^\infty (Cy^{\sigma_0} - G_F(y))y^{-s-1}\,dy$, it gives the statement below with the bound assumed for *all* large y. The version used in the argument assumes it only on windows.

THEOREM 4.3.3 (Alpöge). *Let $F = ((\chi_i, a_i))_i$ be a character family, and let $0 < \sigma_0 < 1$, $\delta > 0$ and $C \in \mathbb{R}$. Suppose that for every Y_0 there is $Y \ge Y_0$ with $G_F(y) \le Cy^{\sigma_0}$ for all $y \in [Y, Y^{1+\delta}]$. Then $L(s, \chi_i) \ne 0$ for every i and every s with $\operatorname{Re} s > \sigma_0$, where $s \ne 1$ when χ_i is the character modulo 1.*

Step (2) of the proof below uses how much of the mass of a Gamma kernel lies in one window. The next two statements give the correct rate and show that it does not improve as the expansion point moves to the right.

For $k \ge 0$ and $\lambda > 0$ let $\mu_{k,\lambda}$ be the probability measure $\frac{\lambda^{k+1}}{k!} u^k e^{-\lambda u}\,du$ on $(0, \infty)$. It is the law of Z/λ, where Z has the Gamma distribution with density $z^k e^{-z}/k!$; its mean is $(k + 1)/\lambda$.

PROPOSITION 4.3.4 (Gamma tails). *Let $k \ge 0$, $\lambda > 0$ and $0 < t < 1$, and put $h(t) = t - \log(1 + t)$. Then $h(t) > 0$ and*

$$\mu_{k,\lambda}\Big(\Big\{u : \Big|\frac{\lambda u}{k+1} - 1\Big| \ge t\Big\}\Big) \le 2e^{-(k+1)h(t)}.$$

PROOF. The function h vanishes at 0 and has derivative $t/(1 + t) > 0$, so $h(t) > 0$. The variable $Z = \lambda u$ has moment generating function $\int_0^\infty e^{\theta z} z^k e^{-z}\,dz/k! = (1 - \theta)^{-(k+1)}$ for $\theta < 1$. Put $n = k + 1$. For $0 < \theta < 1$, Markov's inequality applied to $e^{\theta Z}$ gives $\mu_{k,\lambda}(Z \ge n(1 + t)) \le (1 -$

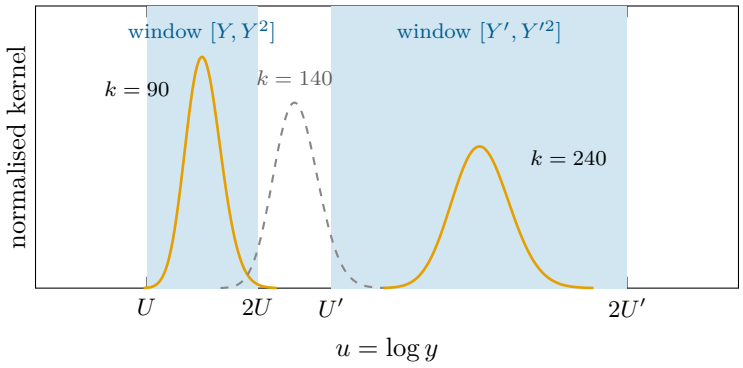


FIGURE 4.2. Step (2) of the proof of Theorem 4.3.3, with illustrative values $\sigma = 40$ and $\delta = 1$. The curves are the kernels $u^k e^{-\sigma u}$ of the Taylor coefficients $c_k(\sigma)$, scaled to total mass 1; the kernel for k peaks at $u = k/\sigma$. In the variable $u = \log y$ a window $[Y, Y^{1+\delta}]$ is the interval $[U, (1+\delta)U]$, $U = \log Y$; here $U = 1.5$ and $U' = 4$. The kernel for $k = 90$ has a proportion $2.7 \cdot 10^{-3}$ of its mass outside $[U, 2U]$, and the kernel for $k = 240$ a proportion $1.7 \cdot 10^{-6}$ outside $[U', 2U']$, so each of these coefficients sees essentially one window. The kernel for $k = 140$ has about 91% of its mass outside both windows; the argument needs no bound on such coefficients, because the lower bound of step (4) holds for every k. The proportions are computed for these illustrative values only.

$\theta)^{-n} e^{-\theta n(1+t)}$; with $\theta = t/(1+t)$ the right side is $e^{-nh(t)}$. Applied to $e^{-\theta Z}$ with $\theta = t/(1-t)$, it gives $\mu_{k,\lambda}(Z \le n(1-t)) \le (1+\theta)^{-n} e^{\theta n(1-t)} = e^{-n(-t-\log(1-t))}$. The power series $-t - \log(1-t) = \sum_{j\ge 2} t^j/j$ and $t - \log(1+t) = \sum_{j\ge 2} (-1)^j t^j/j$ show $-t - \log(1-t) \ge h(t)$. Adding the two tails gives the claim. □

The rate $e^{-(k+1)h(t)}$ does not depend on λ, and nothing better is true: the next proposition shows that the mass outside a window is not $O(\lambda^{-k})$, which is what an earlier version of step (2) asserted.

PROPOSITION 4.3.5. *For every $\lambda > 0$, $U > 0$ and $k \ge 1$,*

$$\mu_{k,\lambda}((0,\infty) \setminus [U, 2U]) \ge 8^{-(k+1)}.$$

In particular, for fixed $\lambda > 8$ there is no constant C' such that, for arbitrarily large U, some $k \ge U$ has $\mu_{k,\lambda}((0,\infty) \setminus [U,2U]) \le C'\lambda^{-k}$.

PROOF. Put $a = \lambda U$; in the variable $z = \lambda u$ the claim is that $Z \notin [a, 2a]$ with probability at least $8^{-(k+1)}$. If $a \ge k/2$, then

$$\mu(Z < a) \ge \frac{1}{k!} \int_0^{k/2} z^k e^{-z}\, dz \ge \frac{e^{-k/2}(k/2)^{k+1}}{(k+1)!} \ge e^{-k/2} 4^{-(k+1)} \ge 8^{-(k+1)},$$

using $(k+1)! \le (k+1)^{k+1} \le (2k)^{k+1}$ and $e^{-1/2} \ge 1/2$. If $a < k/2$, then $2a < k$ and

$$\mu(Z > 2a) \ge \frac{1}{k!}\int_k^{k+1} z^k e^{-z}\,dz \ge \frac{k^k}{k!}e^{-(k+1)} \ge e^{-(k+1)} \ge 8^{-(k+1)}.$$

For the second sentence, $8^{-(k+1)} \le C'\lambda^{-k}$ forces $C' \ge (\lambda/8)^k/8$, which is unbounded as $k \ge U \to \infty$ when $\lambda > 8$. □

The proof compares two estimates for the Taylor coefficients of a Laplace transform at a point far to the right: one coming from the bound on windows, the other from a hypothetical zero. Its steps are elementary complex analysis.

PROOF (OUTLINE). Put $g(u) = G_F(e^u)$ and, for $\operatorname{Re} s > 2$,

$$\widehat{g}_F(s) = \int_0^\infty g(u)e^{-su}\,du = \int_1^\infty G_F(y)y^{-s-1}\,dy.$$

(1) *Continuation and poles.* For a single nontrivial χ, a direct computation gives $\int_1^\infty g_\chi(y)y^{-s-1}\,dy = -\frac{L'}{L}(s,\chi)\frac{1}{s(s-1)} + \frac{L'}{L}(1,\chi)\frac{1}{s-1}$, in which the pole at $s=1$ cancels; for $\chi = 1$ the correction $y(\log y - 1 - \gamma)$ in G cancels the double pole of $-\frac{\zeta'}{\zeta}(s)/(s(s-1))$. Since $\operatorname{Re} g_\chi = (g_\chi + g_{\overline{\chi}})/2$ and no $L(s,\chi_i)$ vanishes on $\operatorname{Re} s \ge 1$, the transform $\widehat{g}_F$ continues meromorphically to $\operatorname{Re} s > 0$, with simple poles exactly at the zeros of the $L(s,\chi_i)$ and $L(s,\overline{\chi}_i)$. At a zero ρ the residue is $-c_\rho/(\rho(\rho-1))$, where $c_\rho > 0$ is the sum of the weights times the multiplicities. For real $\rho \in (0,1)$ this residue is positive.

(2) *The Gamma kernel.* Up to the sign $(-1)^k$, the k-th Taylor coefficient of $\widehat{g}_F$ at a point s_0 with $\operatorname{Re} s_0 = \sigma > 2$ is $c_k(s_0) = \frac{1}{k!}\int_0^\infty g(u)u^k e^{-s_0 u}\,du$. The kernel $u^k e^{-\sigma u}$ peaks at $u = k/\sigma$. Fix the window $W = [U, (1+\delta)U]$ with $U \ge 1$; in the variable $y = e^u$ it is the window $[Y, Y^{1+\delta}]$ with $Y = e^U$ (Figure 4.2). Put $\lambda = \sigma - 2$, assume $\lambda \ge 4/\delta$, and choose k with $k+1 = \lceil \lambda(1+\delta/2)U \rceil$, so that $k+1 \ge \lambda U$. With $t = \delta/(4+3\delta)$, every $u \notin W$ has $|\lambda u/(k+1) - 1| \ge t$, so by Proposition 4.3.4

$$\mu_{k,\lambda}\big((0,\infty) \setminus W\big) \le 2e^{-(k+1)h(t)}.$$

The rate depends on δ but not on σ; by Proposition 4.3.5 no bound of the form $C'\sigma^{-k}$ holds.

(3) *The window bound.* We may take $C \ge 0$. By Proposition 4.2.2(c) and $\operatorname{Ch}(t) \le 2t$, there is A with $|g(u)| \le Ae^{2u}$ for $u \ge 0$. On W we have $g \le R(u) := Ce^{\sigma_0 u}$, hence $|g| + g \le 2R$ there, and off W we have $|g| + g \le 2Ae^{2u}$. Since g is real,

$$|c_k(\sigma + it_0)| + \operatorname{Re} c_k(\sigma) \le \frac{1}{k!}\int_0^\infty (|g|+g)u^k e^{-\sigma u}\,du \le \frac{2C}{(\sigma-\sigma_0)^{k+1}} + \frac{4Ae^{-(k+1)h(t)}}{(\sigma-2)^{k+1}}.$$

If σ is so large in terms of δ that $e^{-h(t)}(\sigma-\sigma_0) < \sigma - 2$, the right side is at most $C''(\sigma-\sigma_0)^{-(k+1)}$ with $C'' = 2C + 4A$.

(4) *The contradiction.* A zero with $\operatorname{Re}\rho > \sigma_0$ gives, through the pole of $\widehat{g}_F$ nearest to σ (real, or an isolated nearest complex pole after a choice of t_0), a lower bound $c\,r^{-(k+1)}$ on the Taylor coefficients for all large k, with $c > 0$ and r the distance from the expansion point to the pole used. That pole has real part $> \sigma_0$, so $r < \sigma - \sigma_0$, and the lower bound contradicts step (3) once k is large, that is, once U is large ($k + 1 \geq (\sigma - 2)U$). One window per coefficient suffices, because the lower bound holds for all large k and not only along a subsequence. □

REMARK 4.3.6 (a correction). An earlier version of step (2) asserted that all but a proportion C'/σ^k of the kernel's mass lies in the window. That is false (Proposition 4.3.5). Hilaf Hasson found the error, with an AI-assisted reading of the earlier version, and gave the counterexample of Proposition 4.3.5. Steps (2) and (3) above use the correct rate of Proposition 4.3.4 instead. The rate $e^{-(k+1)h(t)}$ suffices because $k + 1$ is of size σU, so it beats the trivial growth e^{2u} of g once σ is large in terms of δ. Steps (1) and (4) are unchanged, and the outline remains an outline: step (4), including the choice of t_0 when the nearest pole is not real, is not written out here.

REMARK 4.3.7 (why the weights must be positive). Theorem 4.3.3 assumes only an upper bound for G_F. This suffices because, with positive weights, the residues of $\widehat{g}_F$ at real zeros in $(0, 1)$ all have one sign, so a hypothetical real zero cannot be cancelled by another. The argument must therefore produce an upper bound for a *positive* combination of G_ζ and the functions $\operatorname{Re} g_\psi$. The weights that occur are 1 and $(3o + 1)^{-1}$ (Definition 4.4.1); they arise from the shares of Proposition 6.3.1.

Figure 4.3 pictures the mechanism.

REMARK 4.3.8 (windows are needed). The classical, non-windowed form of Theorem 4.3.3, which follows from Theorem 4.3.2 (see the paragraph after it), does not suffice for the argument. The growth bounds that the argument proves hold only on windows: the Sha bound of Chapter 10 controls the congruence defects only for x in a set with windows $[Y, Y^2]$ (Theorem 10.4.8), and the Chebyshev inequality of Chapter 9 is proved only there. So the windowed form is the one used. The closest result in the literature known to us, [**BR26**, Thm. 3.1], assumes a pointwise bound on $|F(s)|$ near $\sigma = 1$, not a one-sided bound on windows.

QUESTION 4.3.9. Is the windowed criterion of Theorem 4.3.3 new as stated?

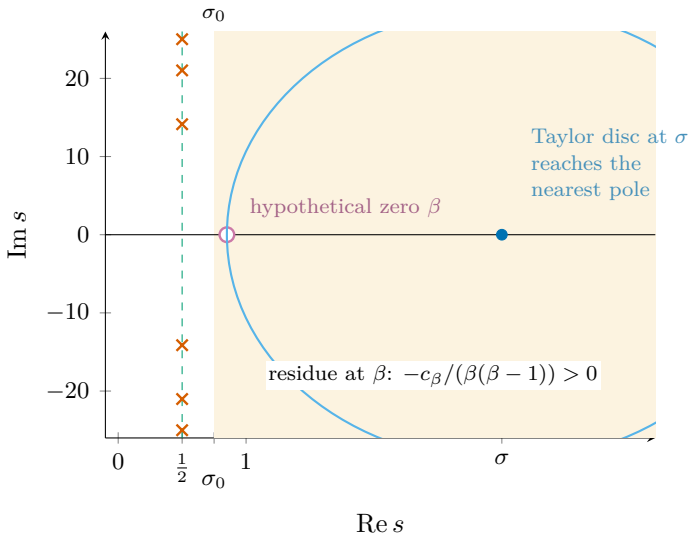


FIGURE 4.3. The mechanism of Theorem 4.3.3. The Laplace transform $\widehat{g}_F$ of the proof has poles at the zeros of the $L(s, \chi_i)$; the crosses are the first zeros of ζ. A hypothetical real zero β with $\frac{1}{2} < \sigma_0 < \beta$ (hollow point) would be a pole with residue $-c_\beta/(\beta(\beta-1)) > 0$, where $c_\beta > 0$ is the sum of the weights times the multiplicities of β (step (1) of the proof), and positive weights give all such residues one sign. The Taylor expansion at a point σ far to the right then has every coefficient bounded below, which contradicts the bound on windows. The disc is drawn as an ellipse because the two axes have different scales.

4.4. Levels

The positive combinations that the argument controls are attached to finite sets of primitive characters closed under inversion.

DEFINITION 4.4.1 (levels). (1) A *level* $\mathscr{L}$ consists of a finite set $\mathrm{Kept}(\mathscr{L})$ of primitive Dirichlet characters that contains the character modulo 1 and is closed under inversion, an optional *exceptional character* $\psi_{\mathrm{exc}} \in \mathrm{Kept}(\mathscr{L})$, and a natural number $o(\mathscr{L})$. The elements of $\mathrm{Kept}(\mathscr{L})$ are its *kept characters*.

(2) The *weight* of $\psi \in \mathrm{Kept}(\mathscr{L})$ is $\mathrm{wt}(\psi) = 1$ if $\psi = \psi_{\mathrm{exc}}$ and $\mathrm{wt}(\psi) = (3o(\mathscr{L}) + 1)^{-1}$ otherwise.

(3) The *weighted Chebyshev function* of $\mathscr{L}$ is

$$G_{\mathscr{L}}(y) = G_\zeta(y) + \sum_{\psi \in \mathrm{Kept}(\mathscr{L}),\ \psi \neq 1} \mathrm{wt}(\psi)\, \mathrm{Re}\, g_\psi(y).$$

(4) For a primitive character ψ, the *cyclic level* $\mathscr{L}_\psi$ has as kept set the primitive characters inducing the powers of ψ, no exceptional character, and $o(\mathscr{L}_\psi)$ equal to the order of ψ.

All weights of a level are positive, and the trivial character enters with weight 1 through G_ζ. Thus $G_{\mathscr{L}_\psi} = G_\zeta + (3o+1)^{-1}\sum_\chi \operatorname{Re} g_\chi$, the sum over the primitive characters $\chi \neq 1$ that induce powers of ψ, where o is the order of ψ.

DEFINITION 4.4.2 (windows). Let $S \subset \mathbb{R}$, $\lambda > 1$, and let $\mathscr{L}$ be a level.

(1) S *has windows of exponent* λ if for every y there is $Y \geq y$ with $[Y, Y^\lambda] \subset S$.
(2) $\mathscr{L}$ *satisfies the window criterion at* λ if for every $\sigma_0 \in (0,1)$ the following holds: if $\{x : G_{\mathscr{L}}(x) \leq x^{\sigma_0}\}$ has windows of exponent λ, then $L(s,\psi) \neq 0$ for every $\psi \in \operatorname{Kept}(\mathscr{L})$ and every s with $\operatorname{Re} s > \sigma_0$, where $s \neq 1$ when ψ is the character modulo 1.

THEOREM 4.4.3 (Alpöge). *Let $\mathscr{L}$ be a level and $\lambda > 1$. Then $\mathscr{L}$ satisfies the window criterion at λ.*

PROOF. Let F be the character family whose pairs are $(1,1)$, for the character modulo 1, and $(\psi, \operatorname{wt}(\psi))$ for $\psi \in \operatorname{Kept}(\mathscr{L})$, $\psi \neq 1$. A primitive character is either the character modulo 1 or nontrivial, and every weight is positive, so F is a character family. Since $g_1 = G = G_\zeta$ on $y > 0$ (Proposition 4.2.2(a)), $G_F = G_{\mathscr{L}}$. Let $\sigma_0 \in (0,1)$ and suppose that $\{x : G_{\mathscr{L}}(x) \leq x^{\sigma_0}\}$ has windows of exponent λ. Then for every Y_0 there is $Y \geq Y_0$ with $G_F(y) \leq y^{\sigma_0}$ on $[Y, Y^\lambda]$. This is the hypothesis of Theorem 4.3.3 with $\delta = \lambda - 1$ and $C = 1$, whose conclusion is the required nonvanishing. □

The windows that occur are pictured in Figure 4.4: the Sha bound gives exponent 2, the supply of hosts is stated with 3/2 (Question 12.1.3), and Theorem 4.4.3 accepts any exponent $\lambda > 1$.

4.5. From a half-plane to the critical line

The passage from zero-freeness in $\operatorname{Re} s > 1/2$ to the critical line is classical. For a primitive Dirichlet character χ modulo N, let $\mathfrak{a} \in \{0,1\}$ be given by $\chi(-1) = (-1)^{\mathfrak{a}}$, and put $\xi(s,\chi) = (N/\pi)^{(s+\mathfrak{a})/2}\Gamma((s+\mathfrak{a})/2)L(s,\chi)$. We write (H) for the following hypothesis: $L(s,\chi') \neq 0$ for every primitive Dirichlet character χ' and every s with $\operatorname{Re} s > 1/2$, where $s \neq 1$ when χ' is the character modulo 1.

THEOREM 4.5.1 (Riemann, Hecke [**Dav00, Rie59**]). *Let χ be a primitive Dirichlet character modulo N. Then there is $\epsilon(\chi) \in \mathbb{C}$ with $|\epsilon(\chi)| = 1$ and $\xi(s,\chi) = \epsilon(\chi)\xi(1-s,\overline{\chi})$ for all $s \neq 0, 1$.*

For $N = 1$ this is Riemann's functional equation.

PROPOSITION 4.5.2. *Let χ be a Dirichlet character modulo N, induced by the primitive character χ^* modulo N^*.*

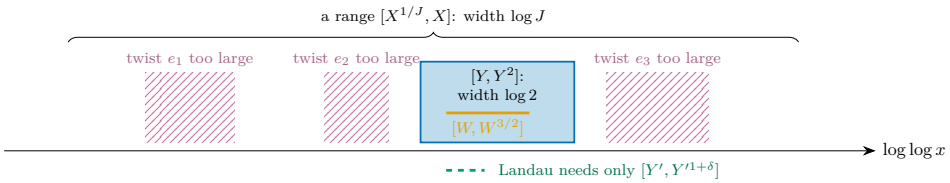


FIGURE 4.4. Windows on a $\log\log x$ axis, on which $[Y, Y^\lambda]$ has fixed width $\log\lambda$. The Sha bound gives windows $[Y, Y^2]$: in every range $[X^{1/J}, X]$ a pigeonhole argument over the at most R exceptional Tate twists finds a window on which all of them are small (Theorem 10.4.8, Remark 10.5.2). The supply of hosts is stated with windows $[W, W^{3/2}]$; a window $[Y, Y^2]$ contains $[Y, Y^{3/2}]$, and the step between the two statements is not reconstructed (Question 12.1.3). Theorem 4.3.3 needs only windows $[Y, Y^{1+\delta}]$ for arbitrarily large Y. Here $R = 3$; positions and widths are illustrative.

(a) *For $s \neq 1$, and for every $s \in \mathbb{C}$ if χ is nonprincipal, $L(s,\chi) = L(s,\chi^*)\prod_{p\mid N}(1-\chi^*(p)p^{-s})$. In particular, for such s with $\operatorname{Re} s \neq 0$, $L(s,\chi) = 0$ if and only if $L(s,\chi^*) = 0$.*

(b) *The L-function of the character modulo 1 is ζ on $s \neq 1$.*

(c) *Under (H), $L(s,\chi) \neq 0$ for every s with $\operatorname{Re} s > 1/2$, where $s \neq 1$ when χ is principal.*

(d) *Under (H), every zero of $L(s,\chi)$ with $0 < \operatorname{Re} s < 1$ has $\operatorname{Re} s = 1/2$, and every zero $s \neq 1$ of ζ that is not in $\{-2,-4,-6,\dots\}$ has $\operatorname{Re} s = 1/2$.*

PROOF OF (A). For $\operatorname{Re} s > 1$ the Euler products give $L(s,\chi) = \prod_p(1-\chi(p)p^{-s})^{-1}$ with $\chi(p) = \chi^*(p)$ for $p \nmid N$ and $\chi(p) = 0$ for $p \mid N$, which is the identity. Both sides are meromorphic on $\mathbb{C}$ and holomorphic off $s = 1$, so the identity theorem extends it to $s \neq 1$. If χ is nonprincipal, then χ^* is nontrivial, $L(s,\chi)$ and $L(s,\chi^*)$ are entire, and the identity theorem on $\mathbb{C}$ gives the identity for every s. A factor $1-\chi^*(p)p^{-s}$ vanishes only if $|p^{-s}| = 1$, that is, $\operatorname{Re} s = 0$. □

PROOF OF (B). The character modulo 1 takes the value 1 at every $n \geq 1$, so the two Dirichlet series coincide on $\operatorname{Re} s > 1$. Conclude by the identity theorem. □

PROOF OF (C). Let $\operatorname{Re} s > 1/2$, with $s \neq 1$ if χ is principal. If χ is nonprincipal, then χ^* is nontrivial, so $L(s,\chi^*) \neq 0$ by (H), also at $s = 1$. If χ is principal, then χ^* is the character modulo 1 and $s \neq 1$, so again $L(s,\chi^*) \neq 0$. In both cases part (a) applies at s, and since $\operatorname{Re} s \neq 0$ it gives $L(s,\chi) \neq 0$. □

PROOF OF (D). First let χ be primitive, $0 < \operatorname{Re} s < 1$ and $L(s,\chi) = 0$. The factor $(N/\pi)^{(s+\mathfrak{a})/2}\Gamma((s+\mathfrak{a})/2)$ is finite and nonzero, because $\operatorname{Re}((s+\mathfrak{a})/2) > 0$; the same holds at $1-s$. Apply Theorem 4.5.1 to $\overline{\chi}$ at the point $1-s$: $\xi(1-s,\overline{\chi}) = \epsilon(\overline{\chi})\xi(s,\chi) = 0$, so $L(1-s,\overline{\chi}) = 0$, and $\overline{\chi}$ is primitive. If $\operatorname{Re} s > 1/2$ this contradicts (H) for χ; if $\operatorname{Re} s < 1/2$ it contradicts (H) for $\overline{\chi}$ at $1-s$ (note $s \neq 1$ and $1-s \neq 1$). So $\operatorname{Re} s = 1/2$. For an arbitrary χ, a zero of $L(s,\chi)$ in the strip is a zero of $L(s,\chi^*)$ by part (a), so it lies on the line.

Now let $\zeta(s) = 0$ with $s \neq 1$ and $s \notin \{-2,-4,\dots\}$. By part (b) and (H) for the character modulo 1, $\operatorname{Re} s \leq 1/2$. If $0 < \operatorname{Re} s < 1$, then $\operatorname{Re} s = 1/2$ by the first paragraph. If $\operatorname{Re} s \leq 0$, then $\operatorname{Re}(1-s) \geq 1$, so $\zeta(1-s) \neq 0$ by (H). For $s \neq 0$ the functional equation $\pi^{-s/2}\Gamma(s/2)\zeta(s) = \pi^{-(1-s)/2}\Gamma((1-s)/2)\zeta(1-s)$ then has a nonzero finite right-hand side, so $\zeta(s) = 0$ forces a pole of $\Gamma(s/2)$, that is, $s \in \{-2,-4,\dots\}$, which was excluded. The case $s = 0$ does not occur, as $\zeta(0) = -1/2$. □

REMARK 4.5.3 (the step to the Riemann hypothesis). The step from the half-plane to the Riemann hypothesis for ζ does not use all Dirichlet characters, their functional equations or the property $|\epsilon(\chi)| = 1$ of the root numbers. The statement about ζ in Proposition 4.5.2(d) uses only the hypothesis for the character modulo 1 and Riemann's functional equation. Where the functional equation of a primitive character is used, in the first paragraph of the proof of part (d), it is applied to $\overline{\chi}$ at a zero of $\xi(s,\chi)$, so the root number multiplies zero and none of its properties is used.

CHAPTER 5

Host fields and averaging over auxiliary primes

This chapter constructs the host fields attached to a primitive character ψ and bounds the characters of their Galois groups other than the kept ones. It has two outputs. Theorem 5.2.4 gives, for every primitive ψ and suitable auxiliary primes q, d, a host with twist form for ψ. Theorem 5.4.8 gives, at every large scale, a window pair (q, d) for which every twist obeys a pointwise bound at the test scale and the real twists have a zero-free box. Through Proposition 5.4.9 this is the pointwise clause of the explicit-formula condition (Definition 5.3.3), the hypothesis through which this chapter reaches Chapter 6. These three results are combined in the supply of hosts, Theorem 12.1.2 of Chapter 12. In Chapter 3, the supply of hosts is a step of the proof of Theorem 3.1.1 (Section 3.12), and Theorem 5.4.8 is one of the inputs that Remark 3.12.2 lists as substantive. No hypothesis of Riemann type is used (Remark 5.5.1).

The geometry of Chapters 7–9 lives over an abelian CM field E attached to the primitive character ψ whose L-function is to be controlled. The Galois group of E has many characters besides the powers of ψ, and the balance identity of Chapter 6 brings all of them in. Every character of $\mathrm{Gal}(E/\mathbb{Q})$ that is trivial on a certain subgroup is a twist $\xi\chi_q^i\chi_d^j$ of a power ξ of ψ by the quadratic characters of two auxiliary primes q and d. The twists with $(i, j) \neq (0, 0)$ need to be bounded only at a test scale x that is polylogarithmic in their conductor, and at such a scale a twisted prime sum is small for most choices of q and d, by a mean-square estimate over the auxiliary prime.

Two letters are fixed for this chapter. The 2-adic exponent of a frame is ν; the bare letter e is kept for Tate twists in Chapter 10. The exponent of the interval $[Y, Y^{\Delta}]$ on which the auxiliary primes are tested is Δ.

5.1. Frames and hosts

Definition 5.1.1 (frames). For an integer $\nu \geq 0$ let H_ν be the quotient of $(\mathbb{Z}/2^{\nu+2}\mathbb{Z})^\times$ by $\{\pm 1\}$.

(1) A *frame* ϕ consists of an integer $\nu \geq 0$, a finite abelian group A, written additively, a surjection $h_0 \colon A \to H_\nu$, and an element $\alpha_0 \in A$ with $\langle \alpha_0 \rangle \cap \ker h_0 = 0$.

(2) Its *core* is $A_0 = \langle \alpha_0 \rangle$, with order $n_0 = |A_0|$, and its *index* is $o = [A : A_0]$.

(3) Its *Galois model* is $\mathfrak{G} = \{\pm 1\}^2 \times A$, with the elements $c = ((-1,-1),0)$ and $\iota = ((-1,1),0)$.

The group $\mathfrak{G}$ will be the Galois group of the host field, with c complex conjugation. The first sign records the action on $\sqrt{-d}$ and the second the action on a second quadratic subfield; the group A carries the 2-power cyclotomic part through h_0 and the part cut out by ψ.

Definition 5.1.2 (hosts). Let ϕ be a frame.

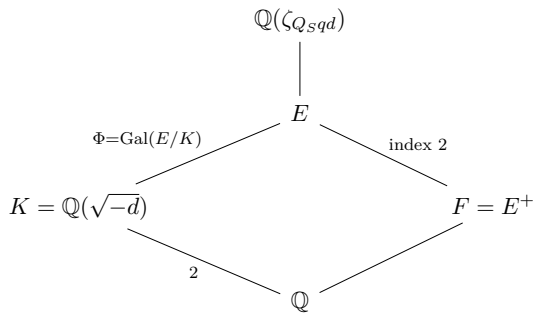
(1) A *host* $\mathcal{H}$ for ϕ consists of:
   (i) primes $q \equiv 3 \pmod 8$ and $d \equiv 7 \pmod{16}$ with $(-d/q) = 1$, the *auxiliary primes*;
   (ii) an integer $Q_S > 0$ with $2^{\nu+2} \mid Q_S$ and $\gcd(Q_S, qd) = 1$, every odd prime divisor of which splits in $K = \mathbb{Q}(\sqrt{-d})$;
   (iii) a number field E, Galois over $\mathbb{Q}$, with an isomorphism $\mathrm{Gal}(E/\mathbb{Q}) \cong \mathfrak{G}$, containing $\sqrt{-d}$, on which $\mathfrak{G}$ acts through the first sign of $\{\pm 1\}^2$, and with an embedding $E \hookrightarrow \mathbb{Q}(\zeta_{Q_S qd})$;
   (iv) an Artin map $(\mathbb{Z}/Q_S qd\mathbb{Z})^\times \to \mathfrak{G}$, compatible with the cyclotomic action, sending -1 to c, compatible with h_0 on the 2-power part, with inertia at d of order at most 2 and inertia at q of exponent 2.
(2) The *modulus* of $\mathcal{H}$ is $N_{\mathcal{H}} = Q_S qd$.
(3) For a character θ of $\mathfrak{G}$, χ_θ is the primitive Dirichlet character that agrees with θ composed with the Artin map on the integers prime to $N_{\mathcal{H}}$.
(4) ω_q is the character of $\mathfrak{G}$ given by the second sign.
(5) The *half root discriminant* of $\mathcal{H}$ is $v = \frac{1}{2}|\operatorname{disc} E|^{1/[E:\mathbb{Q}]}$.

From these conditions the Artin map is surjective, E is a CM field and c is complex conjugation. The subgroup $\Phi = \mathrm{Gal}(E/K)$ has index 2; it is the CM type induced from K (Definition 6.1.1). Figure 5.1 summarises the fields and the group.

The host field is cut out of a cyclotomic field by explicit congruence conditions.

Definition 5.1.3 (ψ-fields). Let ψ be a primitive character of conductor N_ψ, let $\nu \geq 0$, and let q, d be primes. A number field E is a *ψ-field* for (ψ, ν, q, d) if there are an integer N with $2^{\nu+2}, N_\psi, q, d \mid N$ and an embedding $E \hookrightarrow \mathbb{Q}(\zeta_N)$ whose image is the fixed field of the group U of classes $a \in (\mathbb{Z}/N\mathbb{Z})^\times$ with $a \equiv \pm 1 \pmod{2^{\nu+2}}$, $\psi(a) = 1$, $(a/q) = 1$ and $(a/d) = 1$.

Proposition 5.1.4. *Let ψ be a primitive character, $\nu \geq 0$ an integer and q, d primes. Let E be a ψ-field for (ψ, ν, q, d), with integer N and embedding $E \hookrightarrow \mathbb{Q}(\zeta_N)$. Then the characters of* $\mathrm{Gal}(E/\mathbb{Q})$*, viewed through the embedding as Dirichlet characters modulo N, are exactly the products $\chi_{(2)}\, \xi\, \chi_q^i \chi_d^j$ with $\chi_{(2)}$ an even character modulo $2^{\nu+2}$, $\xi \in \langle \psi \rangle$, $\chi_q = (\cdot/q)$, $\chi_d = (\cdot/d)$ and $i, j \in \{0, 1\}$.*



$\mathrm{Gal}(E/\mathbb{Q}) \cong \mathfrak{G} = \{\pm 1\}^2 \times A$
$c = ((-1,-1),0)$: complex conjugation
$\iota = ((-1,1),0)$: acts on $\sqrt{-d}$ only
$A_0 = \langle \alpha_0 \rangle \subset A$, $o = [A : A_0]$
$h_0 \colon A \twoheadrightarrow H_\nu$ (2-power part)
first sign: action on $\sqrt{-d}$; second sign: ω_q

FIGURE 5.1. The host field of Definitions 5.1.1 and 5.1.2. The field E is abelian over $\mathbb{Q}$, contains $\sqrt{-d}$ and embeds in $\mathbb{Q}(\zeta_{Q_S q d})$, where $q \equiv 3 \pmod 8$ and $d \equiv 7 \pmod{16}$ are the auxiliary primes. The group $\Phi = \mathrm{Gal}(E/K)$ has index 2 and is the CM type induced from K; E is a CM field, and c is complex conjugation.

PROOF. Let $U \subset (\mathbb{Z}/N\mathbb{Z})^\times$ be the group of classes a with $a \equiv \pm 1 \pmod{2^{\nu+2}}$, $\psi(a) = 1$, $(a/q) = 1$ and $(a/d) = 1$, so that E is the fixed field of U (Definition 5.1.3). By the Galois correspondence for $\mathbb{Q}(\zeta_N)$, $\mathrm{Gal}(E/\mathbb{Q}) \cong (\mathbb{Z}/N\mathbb{Z})^\times/U$, so its characters are the Dirichlet characters modulo N trivial on U [**Was97**]. Let X be the group generated by the even characters modulo $2^{\nu+2}$, ψ, χ_q and χ_d. A class a lies in the kernel of every even character modulo $2^{\nu+2}$ if and only if $a \equiv \pm 1 \pmod{2^{\nu+2}}$, because these characters are the characters of $(\mathbb{Z}/2^{\nu+2}\mathbb{Z})^\times/\{\pm 1\}$. Hence $U = \bigcap_{\chi \in X} \ker \chi$. For a finite abelian group G and a subgroup X of its dual, the characters trivial on $\bigcap_{\chi \in X} \ker \chi$ form X, by duality of finite abelian groups. Every element of X has the stated form, because X is abelian, generated by these characters, and $\chi_q^2 = \chi_d^2 = 1$. □

5.2. Kept and bounded characters

The balance identity of Chapter 6 produces a combination of the functions g_{χ_θ} over characters θ of $\mathfrak{G}$ trivial on A_0. Those that belong to the level are kept; the others must be bounded.

DEFINITION 5.2.1 (kept and bounded characters). Let $\mathcal{H}$ be a host for a frame ϕ, and let S be a finite set of primitive characters.

(1) X_+ is the set of characters θ of $\mathfrak{G}$ trivial on A_0 with $\theta(\iota) = 1$, and X_- the set of those with $\theta(\iota) = -1$.
(2) A character $\theta \in X_+ \cup X_-$ is *kept* if $\chi_\theta \in S$.
(3) A character θ is *bounded* if $\theta \in X_+$ and θ is not kept, or $\theta \in X_-$.
(4) $\mathcal{H}$ has *kept characters in* X_+ if every element of S is χ_θ for some $\theta \in X_+$.

(5) Let ψ_{exc} be an optional element of S. The host *keeps the exceptional character* ψ_{exc} if $\psi_{\mathrm{exc}} = \chi_{\omega_q}$ when ψ_{exc} is given, and ω_q is not kept when it is not given.

For a level $\mathscr{L}$ we take $S = \mathrm{Kept}(\mathscr{L})$ and ψ_{exc} the exceptional character of $\mathscr{L}$ (Definition 4.4.1).

DEFINITION 5.2.2 (twist form). Let ψ be a primitive character and $\mathcal{H}$ a host for a frame ϕ, with auxiliary primes q, d, and put $\chi_q = (\cdot/q)$ and $\chi_d = (\cdot/d)$. The host has *twist form for* ψ if every character θ of $\mathfrak{G}$ trivial on A_0 admits $\xi \in \langle\psi\rangle$ and $i, j \in \{0, 1\}$ with $\theta(\iota) = (-1)^j$ and $\theta(a) = \xi(a)\chi_q(a)^i\chi_d(a)^j$ for every integer a prime to $N_{\mathcal{H}}$, θ being composed with the Artin map.

PROPOSITION 5.2.3. *Let ψ be a primitive character and $\mathcal{H}$ a host with twist form for ψ, with auxiliary primes q, d, and put $\chi_q = (\cdot/q)$ and $\chi_d = (\cdot/d)$. Let θ be bounded for $S = \mathrm{Kept}(\mathscr{L}_\psi)$. Let $\xi \in \langle\psi\rangle$ and $i, j \in \{0, 1\}$ satisfy $\theta(\iota) = (-1)^j$ and $\theta(a) = \xi(a)\chi_q(a)^i\chi_d(a)^j$ for every integer a prime to $N_{\mathcal{H}}$. Then $(i, j) \neq (0, 0)$.*

PROOF. Suppose $(i, j) = (0, 0)$. Then $\theta(\iota) = 1$, so $\theta \notin X_-$, and therefore θ is in X_+ and not kept. But θ agrees with ξ on the integers prime to $N_{\mathcal{H}}$, so χ_θ is the primitive character inducing $\xi \in \langle\psi\rangle$, which lies in $\mathrm{Kept}(\mathscr{L}_\psi)$. Thus θ is kept, a contradiction. □

Table 5.1 lists the four kinds of characters.

Hosts with all these properties exist for every primitive ψ.

THEOREM 5.2.4 (Alpöge). *Let ψ be a primitive character of conductor N_ψ. There are an integer $c_\psi \geq 1$ and ν_0 with the following property. Let $\nu \geq \nu_0$, let $z \geq 4N_\psi$, and let q, d be primes with $q \equiv 3 \pmod 8$, $q \neq 3$, $d \equiv 7 \pmod{16}$, $(-d/q) = 1$ and $\gcd(N_\psi, qd) = 1$, such that every odd prime $p \leq z$ splits in $\mathbb{Q}(\sqrt{-d})$. Then there are a frame ϕ with exponent ν and $|A| = c_\psi 2^\nu$, and a host $\mathcal{H}$ for ϕ with auxiliary primes q, d, such that, for $S = \mathrm{Kept}(\mathscr{L}_\psi)$: the index of ϕ is $o(\mathscr{L}_\psi)$, the host has kept characters in X_+, it keeps the exceptional character of $\mathscr{L}_\psi$, and it has twist form for ψ.*

IDEA OF PROOF. We describe the standard route. One takes for E a ψ-field (Definition 5.1.3) inside $\mathbb{Q}(\zeta_N)$ for a suitable N divisible by $2^{\nu+2}$, N_ψ, q and d. By Proposition 5.1.4 the character group of $\mathrm{Gal}(E/\mathbb{Q})$ is generated by the even characters modulo $2^{\nu+2}$, ψ, χ_q and χ_d, and the frame is read off from this description: the two signs come from χ_q and χ_d (the first sign is the action on $\sqrt{-d}$), and A carries the 2-power and ψ-parts, so that the characters trivial on A_0 are exactly the twists of Definition 5.2.2. The splitting in K of the odd primes $p \leq z$ is a hypothesis of the theorem. The conditions on Q_S are congruence conditions, handled by the Chinese remainder theorem and quadratic reciprocity. All of this is standard class field theory over $\mathbb{Q}$. In the application the splitting hypothesis is arranged

(i,j)	character θ	$\theta(\iota)$	status for Kept($\mathscr{L}_\psi$)	controlled by
$(0,0)$	ξ	$+1$ (X_+)	kept (proof of Prop. 5.2.3): weights 1 (trivial) and $(3o+1)^{-1}$	— (they are the target)
$(1,0)$	$\xi\chi_q$	$+1$ (X_+)	bounded when $q \nmid N_\psi$ (then not kept)	good pair, clause (1); clause (2) if ξ is real
$(0,1)$	$\xi\chi_d$	-1 (X_-)	bounded, as $\theta(\iota) = -1$	same
$(1,1)$	$\xi\chi_q\chi_d$	-1 (X_-)	bounded, as $\theta(\iota) = -1$	same

TABLE 5.1. Characters θ of $\mathfrak{G}$ trivial on A_0, for a host with twist form for ψ. Each is $\xi\chi_q^i\chi_d^j$ with $\xi \in \langle\psi\rangle$ and $\theta(\iota) = (-1)^j$ (Definition 5.2.2). By Proposition 5.2.3 a bounded character has $(i,j) \neq (0,0)$. Conversely, $(0,1)$ and $(1,1)$ are bounded because $\theta(\iota) = -1$, and $(1,0)$ is bounded when $q \nmid N_\psi$, since it is then not kept (Definition 5.2.1). The bounded characters are controlled by the pointwise bound at the test scale, clause (1) of a good pair (Definition 5.4.2); for real twists, clause (2) adds a zero-free box.

by the choice of a window pair (Definition 5.3.1 and the paragraph after it). □

5.3. Scales and the explicit-formula condition

The bounded characters are needed only at one scale x, for top places with prescribed shares.

DEFINITION 5.3.1 (scales and window pairs). Let $\varepsilon \in (0, 1/52)$.

(1) A *scale* is a triple (x, ε, J) with $x > 1$ and $J \geq 1$. Put $b = \exp(x^\varepsilon)$ and $a = xb$.
(2) A *system of scales* of exponent λ is a J together with a set $\mathrm{Win} \subset \mathbb{R}$ that has windows of exponent λ (Definition 4.4.2).
(3) An integer ν is *adapted* to (ε, x) if $x^{2+8\varepsilon} \leq 2^\nu \leq 2x^{2+8\varepsilon}$.
(4) A pair of primes (q,d) is a *window pair* for (ε, Y) if $q \equiv 3 \pmod 8$ and $Y^{\varepsilon/5}/2 < \log q \leq Y^{\varepsilon/5}$; $d \equiv 7 \pmod{16}$ and $Y^{\varepsilon/5}\log^2 Y/2 < \log d \leq Y^{\varepsilon/5}\log^2 Y$; $q \mid d+1$; and every odd prime $p \leq Y^{\varepsilon/5}$ divides $d+1$.

The upper bound $\varepsilon < 1/52$ is fixed once and for all and recurs in the statements below. Only small values of ε are used in the end: the proof of Theorem 12.4.1 lets $\varepsilon \to 0$.

For a window pair, $d \equiv -1 \pmod p$ for every odd prime $p \le Y^{\varepsilon/5}$ and for $p = q$, so all these primes split in $\mathbb{Q}(\sqrt{-d})$, and $(-d/q) = 1$. Thus a window pair meets the conditions of Theorem 5.2.4 with $z = Y^{\varepsilon/5}$ once Y is large. Since $\log q \approx Y^{\varepsilon/5}$, a test point $y \in [Y, Y^{\Delta}]$ is polylogarithmic in q, and hence in the conductors of the twists (Figure 5.2).

A host also carries two further primes, used in the congruence modules of Chapters 7–9. They are fixed here because their sizes depend on the scale.

Definition 5.3.2 (norm primes). Let $\mathcal{H}$ be a host with field E and auxiliary primes q, d, let $K = \mathbb{Q}(\sqrt{-d}) \subset E$, let $k \ge 1$ be an integer and let $b > 1$ be real.

(1) A pair of *norm primes* of $\mathcal{H}$ at (k, b) is a pair of distinct primes ϖ_0, ϖ_1 such that, for $i \in \{0, 1\}$, $(kb)^8 \le \varpi_i \le 2(kb)^8$ and $\varpi_i = \pi_i \overline{\pi_i}$ for some $\pi_i \in \mathcal{O}_K$ with $\pi_i \equiv 1 \pmod{N_{\mathcal{H}}\mathcal{O}_E}$.
(2) For such a pair, $\lg = \log(\varpi_0 \varpi_1)$.

In applications k is the first entry of a ranks datum (Definition 7.1.1) and $b = \exp(x^{\varepsilon})$ is the parameter of the scale. From $\pi_i \equiv 1$ modulo $N_{\mathcal{H}}$ one gets $\varpi_i \equiv 1 \pmod{N_{\mathcal{H}}}$, so ϖ_i splits completely in $\mathbb{Q}(\zeta_{N_{\mathcal{H}}})$ and hence in E. The norm primes are level primes at which the local type of the automorphic forms is prescribed; they enter the later chapters only through lg, in the upper link (Definition 7.5.1) and in the ties (Definition 9.1.4). They are distinct from the auxiliary primes q, d and from the core generator α_0 of the frame.

Definition 5.3.3 (shares and the explicit-formula condition). Let $\mathscr{L}$ be a level, $\mathcal{H}$ a host for a frame ϕ with core A_0 and index o, $\mathcal{T} \subset A$ a finite set, and $x, \varepsilon, C, \delta$ real.

(1) The *share* of $\alpha \in A$ is $\lambda_\phi(\alpha) = 4/(3o+1)$ if $\alpha \in A_0$ and $3/(3o+1)$ otherwise.
(2) $\mathcal{T}$ *has shares up to* δ if for every $\alpha \in A$,

$$\big|\#\{\tau \in \mathcal{T} : \tau - \alpha \in A_0\} - \lambda_\phi(\alpha)|\mathcal{T}|\big| \le \delta|\mathcal{T}|.$$

(3) The host *satisfies the explicit-formula condition* $\mathrm{EF}(\mathscr{L}, \mathcal{T}, x, \varepsilon, C)$ if $o = o(\mathscr{L})$; the host has kept characters in X_+ and keeps the exceptional character of $\mathscr{L}$, for $S = \mathrm{Kept}(\mathscr{L})$; $\mathcal{T}$ has shares up to $x^{-1+\varepsilon}$; and every bounded θ satisfies $|g_{\chi_\theta}(x)| \le x^{1/2+\varepsilon}$ and $|L'(1, \chi_\theta)/L(1, \chi_\theta)| \le C \log x$.

The shares produce the weights of a level: averaging the characters of A/A_0 over a set $\mathcal{T}$ with these shares gives 1 for the trivial character and $(3o+1)^{-1}$ for the others (Proposition 6.3.1). In Chapter 6 the set $\mathcal{T}$ is the set of top places. The rest of this chapter shows how the pointwise clause is obtained.

5.4. Averaging over the auxiliary prime

The pointwise bounds for the twists are obtained by choosing q and d at random. The objects of the averaging are the following.

Definition 5.4.1 (head sums, prime sets and boxes). Let $Z, y > 0$, let $q \geq 1$ be an integer and let $\kappa\colon \mathbb{Z}_{\geq 0} \to \mathbb{C}$.

(1) The *head sum* is
$$S(Z, y, q, \kappa) = \sum_{1 \leq n \leq Z} \Lambda(n) \min\Big(1, \frac{y}{n}\Big) \kappa(n) \Big(\frac{n}{q}\Big),$$
with $(\frac{n}{q})$ the Jacobi symbol.

(2) For integers m_0, a_0 and $U > 0$, $P(m_0, a_0, U)$ is the set of primes q with $U^{1/2} < q \leq U$ and $q \equiv a_0 \pmod{m_0}$.

(3) For $\varepsilon, Y, U > 0$ the pair (m_0, a_0) is *admitted* if $m_0 > 0$, $8 \mid m_0$, $\gcd(a_0, m_0) = 1$, $\frac{1}{4}(\log Y)^2 \log m_0 \leq \log U$ and $Y^{\varepsilon/6} \leq \log U \leq Y^{\varepsilon/5} (\log Y)^2$.

(4) For $C, Z \in \mathbb{R}$, a finite set P of integers and $I \subset \mathbb{R}$, the *mean-square bound* $\mathrm{MeanSq}(C, Z, P, I)$ holds if
$$\sum_{q \in P} |S(Z, y, q, \kappa)|^2 \leq C\, y \log y\, |P|$$
for every $y \in I$ and every κ with $|\kappa(n)| \leq 1$ for all n.

(5) A Dirichlet character χ modulo N has a *zero-free box of height* Q if $L(s, \chi) \neq 0$ whenever $\operatorname{Re} s > 1 - 1/400$ and $|\operatorname{Im} s| \leq Q$.

(6) For $Q, C, \tau, Y, X \in \mathbb{R}$, a Dirichlet character χ modulo N *has a box* $\mathrm{Box}(Q, C, \tau, Y, X)$ if $\chi \neq 1$, $N \leq Q$, χ has a zero-free box of height Q, $|L'(1, \chi)/L(1, \chi)| \leq C \log Y$, and $|g_\chi(y)| \leq y^{1/2+\tau}$ for every $y \in [Y, X]$.

The head sum is a twisted prime sum truncated at Z, with the weight $\min(1, y/n)$ in place of $\max(y/n - 1, 0)$. For $\chi(n) = \kappa(n)(\frac{n}{q})$ nonprincipal and $y \leq Z$, it is related to g_χ by an exact identity. Write $\sum_{n \leq y} \Lambda(n)\chi(n)(y/n - 1) = y \sum_{n \leq y} \Lambda(n)\chi(n)/n - \sum_{n \leq y} \Lambda(n)\chi(n)$, and use $-L'(1, \chi)/L(1, \chi) = \sum_n \Lambda(n)\chi(n)/n$, a convergent series by the prime number theorem in progressions. Then

$$g_\chi(y) = -S(Z, y, q, \kappa) - y \sum_{n > Z} \frac{\Lambda(n)\chi(n)}{n}. \tag{5.4.1}$$

So $g_\chi(y)$ is small when the head sum is small and the tail is small, and the tail is controlled by the zeros of $L(s, \chi)$ near $\operatorname{Re} s = 1$.

Definition 5.4.2 (good pairs). Let $M \geq 1$ be an integer, $\varepsilon > 0$ and $Y \leq X$.

(i) The *box heights* are $T_j = 8M \exp(Y^{\varepsilon/5}) \exp(Y^{\varepsilon/5} \log^2 Y)^j$ for $j \in \{0, 1\}$.

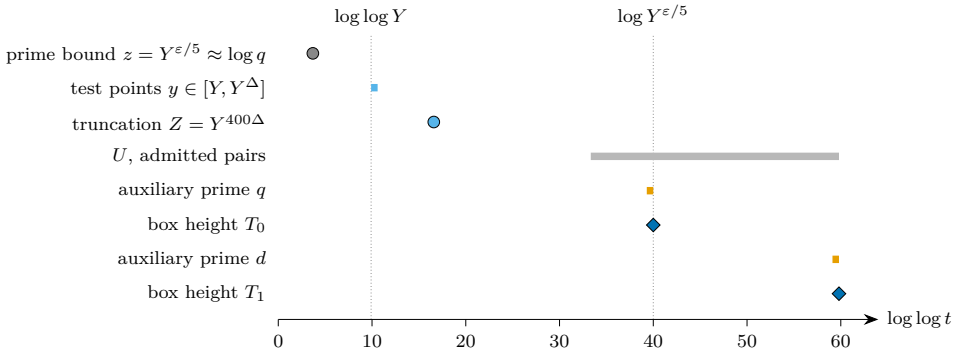


FIGURE 5.2. The sizes of this chapter on one log log axis, to scale, for $\varepsilon = 1/100$, $\Delta = 2$, $M = 1$ and $\log Y = 20000$; a quantity of size t is plotted at $\log\log t$, and a bar is the range allowed by the definitions. The test points $y \in [Y, Y^{\Delta}]$ and the truncation $Z = Y^{400\Delta}$ are of polynomial size in Y (Proposition 5.4.3); the auxiliary primes q, d of a window pair (Definition 5.3.1), the box heights T_0, T_1 (Definition 5.4.2) and the range of U for admitted pairs (Definition 5.4.1) have logarithm about a power of Y. Here $\log y/\log\log q$ lies between 500 and about 1020, so y is polylogarithmic in q. The instance is illustrative: the results need Y large, with no explicit threshold.

(ii) A window pair (q, d) for (ε, Y) with $\gcd(M, qd) = 1$ is *good* for (M, ε, Y, X) if it satisfies clauses (1) and (2) below, where $\chi_q = (\cdot/q)$ and $\chi_d = (\cdot/d)$.

(1) For every character ξ modulo M, every $(i, j) \neq (0, 0)$ and the primitive character χ inducing $\xi\chi_q^i\chi_d^j$ (with respect to the modulus $2Mqd$), every $y \in [Y, X]$ satisfies $|g_\chi(y)| \leq y^{1/2+\varepsilon/2}$ and $|L'(1,\chi)/L(1,\chi)| \leq 4\log y$.

(2) For every real primitive character ϱ with conductor dividing $8M$, every $(i, j) \neq (0, 0)$ and the primitive character χ inducing $\varrho\chi_q^i\chi_d^j$, χ has a zero-free box of height T_j.

The letter T is the usual one for the height of a box of zeros; the letter H is reserved for the groups H_ν of Definition 5.1.1 and H of Definition 6.1.1. Since $Y^{\varepsilon/5}/2 < \log q \leq Y^{\varepsilon/5}$, the height T_0 lies in $[8Mq, 8Mq^2)$, and similarly $T_1 \in [8Mqd, 8Mq^2d^2)$. So the zero-free boxes in clause (2) reach a fixed power of the conductor of χ, whose q- and d-parts are at most q and d.

The averaging rests on four estimates. The first is a mean-square bound over primes in a progression.

PROPOSITION 5.4.3 (Alpöge). *There is $C > 0$ such that for all $\varepsilon \in (0, 1]$ and $\Delta \geq 1$ there is Y_0 with the following property. Let*

$Y \ge Y_0$ and $U > 0$, and let (m_0, a_0) be admitted for (ε, Y, U), with every prime divisor p of m_0 satisfying $p \le Y^{\varepsilon/5}$ or $p > Y^{400\Delta}$. Then $\mathrm{MeanSq}(C, Y^{400\Delta}, P(m_0, a_0, U), [Y, Y^{\Delta}])$ *holds.*

Idea of proof. This is our account of the standard method. Put $Z = Y^{400\Delta}$ and $w(n) = \Lambda(n)\min(1, y/n)$, and expand the square:

$$\sum_{q\in P} |S(Z, y, q, \kappa)|^2 = \sum_{n,n'\le Z} w(n)w(n')\kappa(n)\overline{\kappa(n')} \sum_{q\in P} \Big(\frac{nn'}{q}\Big).$$

The diagonal terms, with nn' a square, contribute $\ll |P| \sum_n w(n)^2 \ll |P|\, y \log y$. For nn' not a square, quadratic reciprocity makes $q \mapsto (\frac{nn'}{q})$ a character of modulus $\ll Z^2$. The sum over $q \in P(m_0, a_0, U)$ is then a sum over primes in a progression to a modulus $\ll Z^2 m_0$. Its logarithm is $o(\log U)$, since $\log Z^2 \ll \Delta \log Y$, $\log U \ge Y^{\varepsilon/6}$ and $\log m_0 \le 4 \log U/(\log Y)^2$. It need not be polylogarithmic in U, because of the factor m_0. In this range the prime number theorem in progressions, with the Deuring–Heilbronn phenomenon for a possible exceptional zero (Remark 5.5.2), gives cancellation, unless the character is determined by the progression. The hypothesis on m_0 says that m_0 has no prime factor in $(Y^{\varepsilon/5}, Z]$. The terms $n = p^r$ whose symbol is constant on the progression ($p \le Y^{\varepsilon/5}$, or r even) contribute $O(y^{1/2} + Y^{\varepsilon/5} \log Y)$ to the head sum, which is within the bound. This is a large-sieve type argument. □

Remark 5.4.4 (a heuristic). For $n \le (\log q)^{O(1)}$ the values (n/q) behave like independent signs as q runs over primes, by the prime number theorem in progressions to the modulus $\prod_{p\le n} p$, which is small on the logarithmic scale of q. This is why the head sums are small on average. In this step no circularity with the Riemann hypothesis appears: the moduli of the progressions are far smaller than q.

The second converts the mean square into a bound for all but a small proportion of the moduli, uniformly in y.

Proposition 5.4.5 (Alpöge). *Let $C \ge 0$, $\tau > 0$ and $r_0 > 0$. There is Y_0 such that for all $Y \ge Y_0$, all $X, Z \in \mathbb{R}$, all finite sets P of natural numbers, all subsets $B \subseteq P$ and all κ with $|\kappa(n)| \le 1$ for every n the following holds. If* $\mathrm{MeanSq}(C, Z, P, [Y, X])$ *holds and every $q \in B$ has some $y \in [Y, X]$ with $|S(Z, y, q, \kappa)| > y^{1/2+\tau}/2$, then $|B| \le r_0|P|$.*

Idea of proof. At a single y, Chebyshev's inequality applied to the mean square bounds the number of q with $|S(Z, y, q, \kappa)| > y^{1/2+\tau}/2$ by $4C \log y\, y^{-2\tau}|P|$. A bad modulus needs a large head sum only at *some* $y \in [Y, X]$, so one must also pass from single points to intervals. The function $y \mapsto S(Z, y, q, \kappa)$ is piecewise linear, with derivative $\sum_{y<n\le Z} \Lambda(n)\kappa(n)(\frac{n}{q})/n$ off the integers. On each dyadic block $[W, 2W]$ a Sobolev-type inequality bounds $\sup |S|^2$ by its mean over the block plus an integral of $|S|\,|\partial_y S|$, both of which are controlled by mean squares. Summing over the blocks,

the bounds $\ll \log W\, W^{-2\tau}$ form a convergent series whose sum tends to 0 as $Y \to \infty$. This account is not a complete proof. □

The third counts the moduli for which a box fails.

PROPOSITION 5.4.6 (Alpöge). *There is C with the following property. Let $Q \geq 3$ and r_1 be real, and let B be a finite set of primes such that each $q \in B$ has $q > r_1$ and admits a positive integer $r \leq r_1$ with $rq \leq Q$ and a primitive character modulo rq without a zero-free box of height Q. Then $|B| \leq CQ^{1/4}$.*

IDEA OF PROOF. The input is a zero-density estimate for boxes near $\operatorname{Re} s = 1$ [**MV07**]: the number of primitive characters of conductor at most Q with a zero in $\operatorname{Re} s > 1 - 1/400$, $|\operatorname{Im} s| \leq Q$, is a small power of Q. Since $q > r_1 \geq r$, distinct $q \in B$ give characters of distinct conductors rq, so $|B|$ is at most that count. □

The fourth turns a bound for the head sum into a box.

PROPOSITION 5.4.7 (Alpöge). *There is $C > 0$ such that for all $\varepsilon \in (0, 1/4]$, $\tau > 0$ and $\Delta \geq 1$ there is Y_0 with the following property. Let $Y \geq Y_0$, $Q \in \mathbb{R}$, $N \geq 1$, let χ be a Dirichlet character modulo N, $q \geq 0$ an integer and $\kappa\colon \mathbb{Z}_{\geq 0} \to \mathbb{C}$. Suppose that $Y^{\varepsilon/6} \leq \log Q \leq Y^{\varepsilon/4}$; that $\chi \neq 1$, $N \leq Q$ and χ has a zero-free box of height Q; that $\chi(n) = \kappa(n)(\frac{n}{q})$ for every $n \geq 0$; and that $|S(Y^{400\Delta}, y, q, \kappa)| \leq y^{1/2+\tau}/2$ for every $y \in [Y, Y^{\Delta}]$. Then χ has a box $\operatorname{Box}(Q, C, \tau, Y, Y^{\Delta})$.*

IDEA OF PROOF. By the identity (5.4.1) with $Z = Y^{400\Delta}$, it suffices to show that the tail $y \sum_{n>Z} \Lambda(n)\chi(n)/n$ is at most $y^{1/2+\tau}/2$ for $y \leq Y^{\Delta}$, and to bound $L'(1,\chi)/L(1,\chi) = -\sum_{n\leq Z} \Lambda(n)\chi(n)/n - \sum_{n>Z} \Lambda(n)\chi(n)/n$. Both follow from a truncated explicit formula for the tail, truncated at height Q. Since $L(s,\chi)$ has no zero with $\operatorname{Re} s > 1 - 1/400$ and $|\operatorname{Im} s| \leq Q$, the contour can be moved to $\operatorname{Re} s = 1 - 1/400$ below height Q, at a cost of $Z^{-1/400} = Y^{-\Delta}$ times factors polynomial in $\log Q \leq Y^{\varepsilon/4}$; the truncation error is controlled because Q is large compared with $\log Z$. The head of $L'/L(1,\chi)$ is at most $\log Z \ll \Delta \log Y$. Our route therefore gives C depending on Δ; the statement has C uniform in Δ, and we do not reconstruct the sharper bound on $L'(1,\chi)/L(1,\chi)$ that this requires. This is of Littlewood type [**Lit28**], with the zero-free box in place of the generalised Riemann hypothesis. Littlewood's bound is for $L(1,\chi)$, and the bound here is for $L'(1,\chi)/L(1,\chi)$. □

With these four estimates good pairs exist at every large scale (Figure 5.3).

THEOREM 5.4.8 (Alpöge). *Let $M > 0$ be an integer, $0 < \varepsilon < 1/52$ and $\Delta \geq 1$. There is Y_1 such that for every $Y \geq Y_1$ some window pair for (ε, Y) is good for $(M, \varepsilon, Y, Y^{\Delta})$.*

PROOF (OUTLINE). Most pairs are good. For $y \in [Y, Y^{\Delta}]$ and $\log q \approx Y^{\varepsilon/5}$ the test point y is polylogarithmic in the conductor. By (5.4.1) with $Z = Y^{400\Delta}$, $g_\chi(y)$ is minus the head sum $S(Z, y, q, \kappa)$ minus a tail. For the twists with $(i, j) = (1, 0)$ one takes $\kappa = \xi$ and averages over q. The head sums have mean square $\ll y \log y$ over the primes q in a progression (Proposition 5.4.3), so all but a proportion r_0 of them satisfy the bound for every y in the range (Proposition 5.4.5). The tail is negligible when $L(s, \chi)$ has a zero-free box, which fails for $\ll Q^{1/4}$ primes (Proposition 5.4.6), a negligible proportion. Proposition 5.4.7 then gives the bounds of clause (1) for most q, and the box count gives clause (2) for the real twists. One chooses q in the good majority, and then d in the same way with q fixed, taking $\kappa = \xi\chi_q^i$ for the twists with $j = 1$, in the class $d \equiv -1$ modulo $q\prod_{3\le p\le Y^{\varepsilon/5}} p$, so that (q, d) is a window pair (Definition 5.3.1). The constant 4 in clause (1) uses that the constant C of Proposition 5.4.7 does not depend on Δ. □

Good pairs give the pointwise clause of Definition 5.3.3 at the scale x, with $C = 4$.

PROPOSITION 5.4.9. *Let ψ be a primitive character of conductor N_ψ and $\mathcal{H}$ a host with twist form for ψ, with $N_\psi \mid Q_S$ and every prime divisor of Q_S dividing $2N_\psi$. Let $\varepsilon > 0$, $x \ge 1$ and $Y \le x$, and suppose that the auxiliary primes (q, d) of $\mathcal{H}$ are good for $(N_\psi, \varepsilon, Y, x)$. Then every θ that is bounded for* $\mathrm{Kept}(\mathscr{L}_\psi)$ *satisfies $|g_{\chi_\theta}(x)| \le x^{1/2+\varepsilon}$ and $|L'(1, \chi_\theta)/L(1, \chi_\theta)| \le 4\log x$.*

PROOF. Since $\mathcal{H}$ has twist form for ψ (Definition 5.2.2), there are $\xi \in \langle\psi\rangle$ and $i, j \in \{0, 1\}$ with $\theta(\iota) = (-1)^j$ and $\theta(a) = \xi(a)\chi_q(a)^i\chi_d(a)^j$ for every integer a prime to $N_{\mathcal{H}}$. By Proposition 5.2.3, $(i, j) \ne (0, 0)$. An integer a prime to $2N_\psi qd$ is prime to $N_{\mathcal{H}} = Q_S qd$, since every prime divisor of Q_S divides $2N_\psi$. So $\chi_\theta(a) = \theta(a) = \xi(a)\chi_q(a)^i\chi_d(a)^j$ for such a, and χ_θ is the primitive character inducing $\xi\chi_q^i\chi_d^j$ with respect to $2N_\psi qd$. Clause (1) of Definition 5.4.2 with $M = N_\psi$, at $y = x \in [Y, x]$, gives $|g_{\chi_\theta}(x)| \le x^{1/2+\varepsilon/2} \le x^{1/2+\varepsilon}$, as $x \ge 1$, and the bound for $L'(1, \chi_\theta)/L(1, \chi_\theta)$. □

5.5. The strength of the analytic input

REMARK 5.5.1 (the analytic input is not of Riemann type). The argument does not need, for every character of the host field that is not kept, a pointwise bound for g_χ of Riemann-hypothesis quality at all large y. The analytic input on the twists is the explicit-formula condition of Definition 5.3.3, whose pointwise clause is required at the single scale x, which is polylogarithmic in the conductor. It is not assumed: it follows from Theorem 5.4.8 and Proposition 5.4.9, through the supply of hosts in Theorem 12.1.2. In terms of zeros, it gives pointwise bounds at polylogarithmic scales and zero-freeness of the real twists in the box $\mathrm{Re}\, s > 1 - 1/400$, $|\mathrm{Im}\, s| \le T_j$, whose height is a power of the conductor. Neither statement has the strength of the Riemann hypothesis near $\mathrm{Re}\, s = 1/2$.

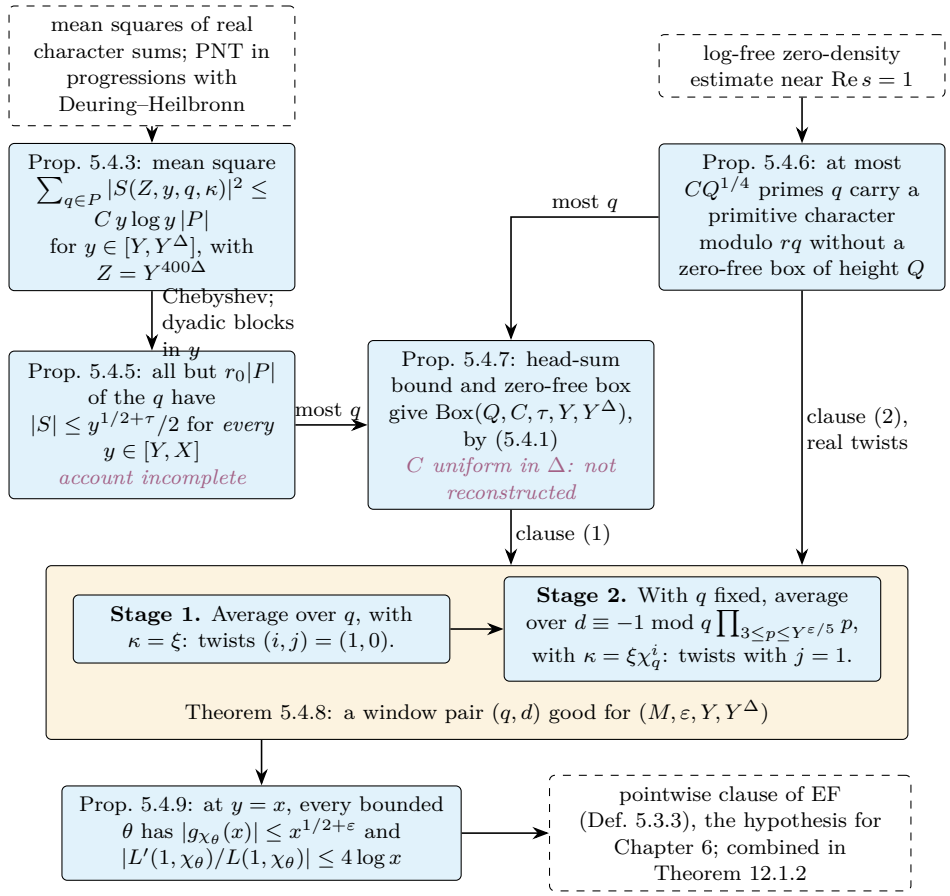


FIGURE 5.3. How the four averaging estimates of this chapter give a good window pair (schematic). Solid boxes are results of the chapter, dashed boxes are classical inputs (Remark 5.5.2) and the place where the output is used. Stage 2 repeats the averaging with d in place of q. The text gives only an incomplete account of Proposition 5.4.5, and does not reconstruct the uniformity in Δ of the constant of Proposition 5.4.7, on which the constant 4 of clause (1) depends.

REMARK 5.5.2 (classical inputs). The classical inputs to the averaging are of the following kinds.

- The zero-free region with at most one exceptional zero [**Pag35**].
- Siegel's theorem [**Sie35**], with its Siegel–Walfisz form in progressions [**Wal36**].
- The Deuring–Heilbronn phenomenon in the Linnik range [**Lin44**].
- A log-free zero-density estimate [**Gal70**], behind Proposition 5.4.6.
- Mean squares of real character sums [**Mon71, HB95**], behind Proposition 5.4.3.

- Mertens-type bounds.

None of Propositions 5.4.3–5.4.7 assumes a hypothesis of Riemann type.

CHAPTER 6

The balance identity and Fourier analysis on the Galois group

This chapter proves the balance identity, Proposition 6.1.4, and outlines the proof of the core Chebyshev bound, Theorem 6.3.4. It also treats the averaging results (Proposition 6.3.1, Proposition 6.3.2 and Theorem 6.3.3) that identify the function extracted from the balance identity with the weighted Chebyshev function $G_{\mathscr{L}}$ of a level (Definition 4.4.1). In Chapter 3 the balance identity organises the finite-side calculation. The core bound is half of the analytic core of the proof of Theorem 3.1.1: with the gap bound, Theorem 9.2.1 of Chapter 9, it gives the Chebyshev inequality on windows, Theorem 9.4.1 (Remark 3.12.2).

Chebyshev obtained bounds of the right order for $\mathrm{Ch}(x)$ from Legendre's formula for the prime factorisation of factorials. The argument of these notes uses a variant of Legendre's count in which only the primes inert in an imaginary quadratic field $K = \mathbb{Q}(\sqrt{-d})$ appear, weighted by a function on a CM type. This chapter shows that such a count equals an explicit main term plus a combination of twisted Chebyshev functions, and that averaging over suitable translates isolates exactly the weighted Chebyshev function $G_{\mathscr{L}}$ of a level, with its positive weights.

The identity (Proposition 6.1.4) is finite Fourier analysis on the Galois group $\mathfrak{G}$ of the host field: orthogonality of characters sorts the prime powers, and the terms linear in the variable assemble into the main term. Averaging over shifts and over top places with prescribed shares (Proposition 6.3.1) then extracts the characters trivial on a cyclic subgroup, with weights 1 and $(3o+1)^{-1}$. The chapter ends with the core bound (Theorem 6.3.4): an upper bound for the prime sum by its main term, up to the error bracket, gives an upper bound $x^{1/2+O(\varepsilon)}$ for $G_{\mathscr{L}}(x)$. Chapters 7 to 9 supply such an upper bound for the prime sum.

6.1. The additive model and arithmetic data

Definition 6.1.1 (the additive model). Let $V_4 = (\mathbb{Z}/2\mathbb{Z})^2$, let H be a finite abelian group of order n, and let $\mathfrak{G} = V_4 \times H$, written additively. Write an element as $g = ((g_1, g_2), h)$ with $g_1, g_2 \in \mathbb{Z}/2\mathbb{Z}$ and $h \in H$.

(1) The *CM type* is the subgroup $\Phi = \{g \in \mathfrak{G} : g_1 = 0\}$, and $f = |\Phi| = 2n$.
(2) For $\tau \in H$, the *partner map* is $\mathrm{pt}_\tau \colon \mathfrak{G} \to \mathfrak{G}$, $\mathrm{pt}_\tau(g) = g + ((1,1), \tau)$.

(3) The characters χ_K and ω of $\mathfrak{G}$ are $\chi_K(g) = (-1)^{g_1}$ and $\omega(g) = (-1)^{g_2}$, and $s_0 = ((1,0),0)$.

For the rest of this chapter H is a finite abelian group of order n and $\mathfrak{G} = V_4 \times H$, with the CM type Φ, the number f, the partner maps pt_τ, the characters χ_K and ω and the element s_0 of Definition 6.1.1. We write θ for a character of $\mathfrak{G}$ and θ_0 for the trivial character. For a host (Definition 5.1.2) we take $H = A$. The multiplicative model $\{\pm 1\}^2 \times A$ of Definition 5.1.1 is matched with the additive model coordinate by coordinate: writing $\bar{s} \in \mathbb{Z}/2\mathbb{Z}$ for the class with $s = (-1)^{\bar{s}}$, the element $((s_1, s_2), h)$ goes to $((\bar{s}_1, \bar{s}_2), h)$. Thus g_1 records the action on $\sqrt{-d}$, so $\Phi = \mathrm{Gal}(E/K)$ is the CM type induced from K and χ_K is the quadratic character of K. Under the dictionary, $\iota \mapsto s_0$, $c \mapsto ((1,1),0)$, and the character ω_q of Definition 5.1.2 becomes ω; the conditions $\theta(\iota) = \pm 1$ of Definition 5.2.1 become $\theta(s_0) = \pm 1$.

Definition 6.1.2 (arithmetic data). Let H be a finite abelian group and $\mathfrak{G} = V_4 \times H$, with its CM type Φ and element s_0 (Definition 6.1.1).

(1) *Arithmetic data* D for $\mathfrak{G}$ consist of natural numbers q_1 and d and, for each prime p, a subgroup $I_p \subset \mathfrak{G}$ (the *inertia* at p) and an element $\mathrm{Frob}_p \in \mathfrak{G}$.

(2) For arithmetic data D, a character θ of $\mathfrak{G}$ and an integer $n \geq 1$, the *averaged value* of θ at n is

$$\theta_D(n) = \prod_{p \mid n} \frac{1}{|I_p|} \sum_{z \in I_p} \theta(v_p(n)\,\mathrm{Frob}_p + z).$$

(3) Let ℓ be a function from the characters of $\mathfrak{G}$ to $\mathbb{C}$. The *arithmetic Chebyshev function* of θ is $G^{\mathrm{ar}}_\theta(y) = G_\zeta(y)$ if $\theta = \theta_0$, and otherwise

$$G^{\mathrm{ar}}_\theta(y) = \sum_{1 \leq n \leq y} \Lambda(n)\,\theta_D(n)\Big(\frac{y}{n} - 1\Big) + y\,\ell(\theta) \qquad (y > 0).$$

(4) A prime p is *inert* for D if $p \notin \{2, q_1, d\}$ and $\mathrm{Frob}_p \notin \Phi$.

(5) Let S be a finite set of primes. The data D are *split outside* d *on* S if d is prime, $2, q_1 \in S$, $d \notin S$, $I_p = 0$ for every prime $p \notin S \cup \{d\}$, $I_p \subset \Phi$ and $\mathrm{Frob}_p \in \Phi$ for every $p \in S$, and $I_d = \{0, s_0\}$.

(6) The data D satisfy the *host facts* if q_1 and d are distinct odd primes and D is split outside d on $\{2, q_1\}$.

(7) Let $\mathcal{H}$ be a host (Definition 5.1.2) with modulus $N = N_{\mathcal{H}}$, auxiliary primes q, d and Artin map $\mathrm{Art}\colon (\mathbb{Z}/N\mathbb{Z})^\times \to \mathfrak{G}$, with $H = A$. The *arithmetic data of* $\mathcal{H}$ have $q_1 = q$ and the given d. For a prime $p \nmid N$, $I_p = 0$ and $\mathrm{Frob}_p = \mathrm{Art}(p)$. For a prime $p \mid N$, write $(\mathbb{Z}/N\mathbb{Z})^\times = (\mathbb{Z}/p^v\mathbb{Z})^\times \times (\mathbb{Z}/N'\mathbb{Z})^\times$ with $v = v_p(N)$ and $p \nmid N'$, let I_p be the image under Art of the first factor (the inertia at p), and let Frob_p be the image of the class that is 1 in the first factor and p in the second.

For the arithmetic data of a host, $\theta_D(n)$ is the value at n of the primitive Dirichlet character χ_θ attached to θ: averaging over the inertia group at p

passes from θ composed with the Artin map to the primitive character. For a host we take

$$\ell(\theta) = \frac{L'(1,\chi_\theta)}{L(1,\chi_\theta)} \qquad (\theta \neq \theta_0),$$

which is defined because $\chi_\theta \neq 1$ for $\theta \neq \theta_0$, so that $G_\theta^{\mathrm{ar}} = g_{\chi_\theta}$ (Definition 4.2.1). The ties of Definition 9.1.4 impose this choice. For the identity below, ℓ is arbitrary.

Definition 6.1.3 (prime sum, main term and defect). Let D be arithmetic data for $\mathfrak{G} = V_4 \times H$ (Definition 6.1.2), let ℓ be a function from the characters of $\mathfrak{G}$ to $\mathbb{C}$, let $N \geq 1$ be an integer, $a > 0$ and $\tau \in H$, and let $w\colon \mathfrak{G} \to \mathbb{R}$ be positive on Φ.

(1) The *partner weight* is $m_\tau(g) = w(\mathrm{pt}_\tau(g))$ for $g \in \mathfrak{G}$.

(2) The *prime sum* is

$$\mathrm{PS}_{\tau,w}(a) = \sum_{p \leq N \text{ inert}} 2\log p \sum_{\substack{1 \leq i \leq N \\ i\,\mathrm{Frob}_p \notin \Phi}} \Big(\frac{a}{p^i} - m_\tau(i\,\mathrm{Frob}_p)\Big)_+ + \log d \sum_{1 \leq i \leq N} \sum_{\substack{g \notin \Phi \\ g - i\,\mathrm{Frob}_d \in I_d}} \Big(\frac{a}{d^i} - m_\tau(g)\Big)_+.$$

(3) The *main term* is

$$\mathrm{MT}_w(a) = a\Big(\log a - 1 - \gamma + \ell(\chi_K) - \frac{1}{f}\sum_{\rho \in \Phi} \log w(\rho)\Big).$$

(4) The *defect* is

$$\mathrm{DF}_{\tau,w}(a) = \frac{1}{f} \sum_{g \notin \Phi} \sum_{\theta} \overline{\theta(g)}\, m_\tau(g)\, G_\theta^{\mathrm{ar}}(a/m_\tau(g)),$$

the inner sum over all characters θ of $\mathfrak{G}$.

The prime sum is a sum over the primes $p \leq N$ inert for D and over the exponents $i \leq N$ with $i\,\mathrm{Frob}_p \notin \Phi$, plus a term for the prime d; in the case of a host these are the primes inert in K and, for such a prime, the odd exponents. The weights $m_\tau(g)$ at the nodes $g \notin \Phi$ are the values of w on Φ, moved by the partner map (Figure 6.1).

Proposition 6.1.4 (balance identity). *Let D be arithmetic data for $\mathfrak{G} = V_4 \times H$ that are split outside d on a finite set S of primes, let ℓ be a function from the characters of $\mathfrak{G}$ to $\mathbb{C}$, let $a > 0$ and $\tau \in H$, let $w\colon \mathfrak{G} \to \mathbb{R}$ be positive on Φ, and let N be an integer with $a/w(\rho) \leq N$ for every $\rho \in \Phi$. Then*

$$\mathrm{PS}_{\tau,w}(a) - \mathrm{MT}_w(a) = \mathrm{DF}_{\tau,w}(a).$$

Written out, the identity says that the Legendre-type count $\mathrm{PS}_{\tau,w}(a)$ over inert primes equals $a\big(\log a - 1 - \gamma + \ell(\chi_K) - f^{-1}\sum_{\rho\in\Phi}\log w(\rho)\big)$ plus a

combination, with coefficients $\overline{\theta(g)}\, m_\tau(g)/f$, of the values $G_\theta^{\mathrm{ar}}(a/m_\tau(g))$ over all characters θ and all $g \notin \Phi$.

PROOF. We expand $\mathrm{DF}_{\tau,w}(a)$ by orthogonality. Its prime-power terms reassemble into $\mathrm{PS}_{\tau,w}(a)$, and its terms linear in a reassemble into $\mathrm{MT}_w(a)$.

Translation by $((1,1),\tau)$ changes g_1, so pt_τ maps $\mathfrak{G} \smallsetminus \Phi$ bijectively onto Φ. Hence for $g \notin \Phi$ the weight $m_g := m_\tau(g)$ is positive and $y_g := a/m_g \le N$. Also $|\mathfrak{G} \smallsetminus \Phi| = 4n - 2n = f$.

Next we split G^{ar}. Since $\theta_{0,D}(n) = 1$ for every n, we may write, for every θ and $y > 0$,

$$G_\theta^{\mathrm{ar}}(y) = \sum_{n \le y} \Lambda(n)\, \theta_D(n) \Big(\frac{y}{n} - 1 \Big) - \kappa_\theta(y),$$

with $\kappa_{\theta_0}(y) = y \log y - (1+\gamma)y$ and $\kappa_\theta(y) = -y\, \ell(\theta)$ for $\theta \neq \theta_0$.

Consider the prime powers. If $\Lambda(n) \neq 0$, then $n = p^i$ with $i \ge 1$, and $\theta_D(p^i) = |I_p|^{-1} \sum_{z \in I_p} \theta(i\,\mathrm{Frob}_p + z)$. By orthogonality, $\sum_\theta \overline{\theta(g)}\theta(h) = |\mathfrak{G}|$ if $g = h$ and 0 otherwise, so

$$\sum_\theta \overline{\theta(g)}\, \theta_D(p^i) = \begin{cases} |\mathfrak{G}|/|I_p| & \text{if } g \in i\,\mathrm{Frob}_p + I_p, \\ 0 & \text{otherwise.} \end{cases}$$

Also $m_g(y_g/p^i - 1) = a/p^i - m_g$, which is nonnegative exactly when $p^i \le y_g$; and $p^i \le y_g \le N$ forces $p \le N$ and $i \le N$. So the prime-power part of $\mathrm{DF}_{\tau,w}(a)$ equals

$$\frac{1}{f} \sum_{g \notin \Phi} \sum_{\substack{p \le N \\ 1 \le i \le N}} \log p \cdot \frac{|\mathfrak{G}|}{|I_p|} \cdot \mathbf{1}[g \in i\,\mathrm{Frob}_p + I_p] \cdot \Big(\frac{a}{p^i} - m_g \Big)_+ .$$

We sort the primes. If $p \in S$, then $i\,\mathrm{Frob}_p + I_p \subset \Phi$ contains no $g \notin \Phi$, so p contributes nothing. Also p is not inert, since $\mathrm{Frob}_p \in \Phi$. If $p \notin S \cup \{d\}$, then $I_p = 0$ and the coefficient is $|\mathfrak{G}|/f = 2$. The only g is $i\,\mathrm{Frob}_p$, and $i\,\mathrm{Frob}_p \notin \Phi$ for some i only if $\mathrm{Frob}_p \notin \Phi$, that is, only if p is inert (recall $2, q_1 \in S$ and $p \neq d$). These primes give the first sum in $\mathrm{PS}_{\tau,w}(a)$. If $p = d$, then $|I_d| = 2$ and the coefficient is $(4n/2)/f = 1$, which gives the second sum. The three cases are disjoint, because $2, q_1 \in S$ and $d \notin S$. Hence the prime-power part of $\mathrm{DF}_{\tau,w}(a)$ is $\mathrm{PS}_{\tau,w}(a)$.

It remains to treat the linear terms. The rest of $\mathrm{DF}_{\tau,w}(a)$ is $-f^{-1} \sum_{g \notin \Phi} \sum_\theta \overline{\theta(g)}\, m_g\, \kappa_\theta(y_g)$. For θ_0 it equals

$$-\frac{1}{f} \sum_{g \notin \Phi} \Big(a \log \frac{a}{m_g} - (1+\gamma)a \Big) = -a(\log a - 1 - \gamma) + \frac{a}{f} \sum_{\rho \in \Phi} \log w(\rho),$$

by the bijection pt_τ. For $\theta \neq \theta_0$ it equals $f^{-1} a\, \ell(\theta) \sum_{g \notin \Phi} \overline{\theta(g)}$. Since $\sum_{g \in \mathfrak{G}} \overline{\theta(g)} = 0$, the last sum is $-\sum_{g \in \Phi} \overline{\theta(g)}$, which is $-f$ if θ is trivial on Φ and 0 otherwise. The only nontrivial character trivial on the index-two subgroup Φ is χ_K, so these terms contribute $-a\, \ell(\chi_K)$. Adding up, $\mathrm{DF}_{\tau,w}(a) = \mathrm{PS}_{\tau,w}(a) - \mathrm{MT}_w(a)$. □

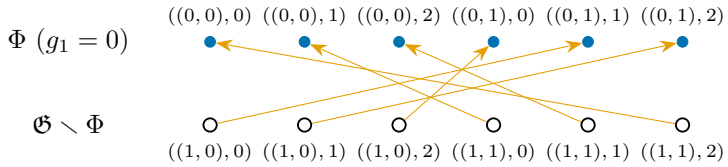


FIGURE 6.1. The partner map $\mathrm{pt}_\tau(g) = g + ((1,1),\tau)$ on $\mathfrak{G} = V_4 \times H$, here with $H = \mathbb{Z}/3$ and $\tau = 1$. It changes the coordinate g_1, so it maps $\mathfrak{G} \setminus \Phi$ bijectively onto Φ. Hence the partner weights $m_\tau(g)$ at the nodes outside Φ run over the weights $w(\rho)$, $\rho \in \Phi$. This is why $f^{-1}\sum_{\rho\in\Phi}\log w(\rho)$ appears in the main term MT_w.

REMARK 6.1.5 (Legendre's formula). Legendre's formula gives $\log\lfloor a\rfloor! = \sum_p \log p \sum_{i\ge1}\lfloor a/p^i\rfloor$ [**Leg08**]. The prime sum $\mathrm{PS}_{\tau,w}(a)$ has the smoothed shape $\sum_p \log p \sum_i (a/p^i - w)_+$ of this count, restricted to the primes inert in K (for CM by K, the supersingular primes) and to the exponents i with $i\,\mathrm{Frob}_p \notin \Phi$, which for an inert prime are the odd ones. Through Proposition 6.1.4, an inequality $\mathrm{PS} \le \mathrm{MT}+(\text{small})$ is therefore a Chebyshev-type upper bound [**Tch52**] for a combination of the functions G_θ^{ar}.

EXAMPLE 6.1.6 (inert primes for $\mathbb{Q}(\sqrt{-7})$). Let $K = \mathbb{Q}(\sqrt{-7})$ and $\chi = \chi_{-7}$. The simplest instance of the shape of Proposition 6.1.4, outside the formalism of arithmetic data, takes the weight $w \equiv 1$:

$$A(a) = \sum_{n\le a} \Lambda(n)(1-\chi(n))\Big(\frac{a}{n} - 1\Big), \qquad M(a) = a\Big(\log a - 1 - \gamma + \frac{L'(1,\chi)}{L(1,\chi)}\Big).$$

For a prime p inert in K, that is, $p \bmod 7 \in \{3,5,6\}$, we have $1-\chi(p^i) = 2$ for odd i and 0 for even i. For a split prime, $1-\chi(p^i) = 0$, and $1-\chi(7^i) = 1$. So $A(a)$ is the Legendre-type count over the inert primes and odd powers, at weight $2\log p$, plus the ramified prime 7 at weight $\log 7$. Directly from Definition 4.2.1,

$$A(a) - M(a) = G_\zeta(a) - g_\chi(a),$$

a main term plus a combination of two Chebyshev functions, which is the shape of the balance identity. Here $L'(1,\chi)/L(1,\chi) = 0.0156357\ldots$, and

a	$A(a)$	$M(a)$	$(A(a) - M(a))/\sqrt{a}$
100	305.081	304.359	0.0722
200	748.989	747.347	0.1161
1000	5346.323	5346.175	0.0047

Figure 6.3 plots the normalised difference for $10 \le a \le 10^5$; it stays small. Figure 6.2 compares, for $a = 200$, the per-prime terms of Legendre's count with those of the inert count.

The computation illustrates the shape of the main term only. It does not test the inequality of Theorem 6.3.4, whose content lies in the size of the error.

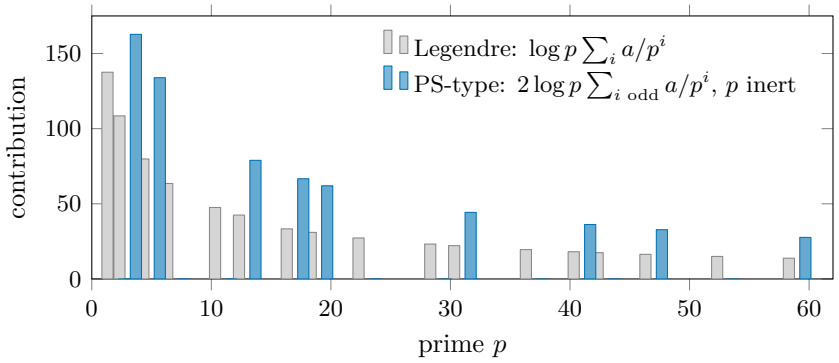


FIGURE 6.2. For $a = 200$ and $K = \mathbb{Q}(\sqrt{-7})$, per prime $p \le 60$: the smoothed Legendre count $\log p \sum_{p^i \le a} a/p^i$ of $\log\lfloor a \rfloor!$, and the count of the type of the prime sum PS (Definition 6.1.3), which keeps only the primes inert in K (those with $p \bmod 7 \in \{3,5,6\}$) and the odd powers, at weight $2\log p$. The count uses no weights and no ramified term at d.

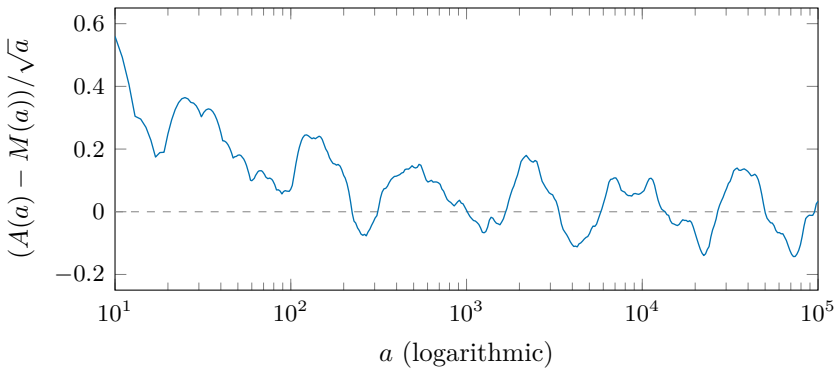


FIGURE 6.3. For $K = \mathbb{Q}(\sqrt{-7})$ and weight $w \equiv 1$: $A(a) = \sum_{n \le a} \Lambda(n)(1-\chi_{-7}(n))(a/n-1)$, the Legendre-type sum over the primes inert in K (odd powers) and the ramified prime, against $M(a) = a(\log a - 1 - \gamma + L'(1,\chi_{-7})/L(1,\chi_{-7}))$, where $L'(1,\chi_{-7})/L(1,\chi_{-7}) = 0.015635\ldots$. The normalised difference stays small. The computation illustrates the shape of the main term MT. It does not test the inequality of Theorem 6.3.4.

6.2. Thick and thin weights

The weights w in the prime sum will come from the geometry of Chapter 7, one weight function for each shift u_0 and each lift index $j < m$. They are of two kinds. At the *thick* nodes they are close to a large number b, so the corresponding arithmetic Chebyshev functions are evaluated near $a_\tau/b \approx x$. At the *thin* nodes they are at most $2b'$, far below b, so the functions are evaluated at arguments of size at least $X^\flat/2$, where the prime number theorem makes them negligible.

Definition 6.2.1 (thick and thin weights). Let D be arithmetic data for $\mathfrak{G} = V_4 \times H$, ℓ a function from the characters of $\mathfrak{G}$ to $\mathbb{C}$, $N \geq 1$ and $m \geq 1$ integers, $A_0 \subset H$ a subgroup, and $\mathcal{T}, \mathcal{T}_2 \subset H$ finite sets. Let $a\colon H \to \mathbb{R}$, written $\tau \mapsto a_\tau$, and $w\colon H \times \{0, \dots, m-1\} \times \mathfrak{G} \to \mathbb{R}$; write $w(u_0, j)$ for the function $\rho \mapsto w(u_0, j, \rho)$ on $\mathfrak{G}$. Let $x, b, b', \eta, X^\flat, \vartheta$ be real numbers.

(1) A node $\rho \in \Phi$ is *thick at* $u_0 \in H$ if $\rho = ((0,1), t + u_0)$ for some $t \in \mathcal{T}_2$, and *thin at* u_0 otherwise.

(2) The *size conditions* hold if $\eta \geq 0$, $X^\flat > 0$, $b' > 0$, $b'X^\flat \leq xb$ and $\vartheta \geq 0$; if $|a_\tau/b - x| \leq 1$ for every $\tau \in \mathcal{T}$; and if for every $u_0 \in A_0$, $j < m$ and $\rho \in \Phi$:
   (a) $w(u_0, j, \rho) > 0$;
   (b) $b \leq w(u_0, j, \rho) \leq b + \eta + 1$ if ρ is thick at u_0;
   (c) $w(u_0, j, \rho) \leq 2b'$ if ρ is thin at u_0;
   (d) $a_\tau / w(u_0, j, \rho) \leq N$ for every $\tau \in \mathcal{T}$.

(3) The *thin-scale bound* holds if $|G_\theta^{\mathrm{ar}}(y)| \leq \vartheta y$ for every character θ of $\mathfrak{G}$ and every $y \geq X^\flat/3$.

(4) For $u_0 \in H$ such that $w(u_0, j)$ is positive on Φ for every $j < m$, the *gap* at u_0 is

$$\mathrm{gap}(u_0) = \sum_{\tau \in \mathcal{T}} \sum_{j<m} \Big(\mathrm{PS}_{\tau, w(u_0,j)}(a_\tau) - \mathrm{Re}\, \mathrm{MT}_{w(u_0,j)}(a_\tau) \Big),$$

with PS and MT the prime sum and the main term of Definition 6.1.3 for D, ℓ and N.

At the sizes used later, $b = \exp(x^\varepsilon)$, $a_\tau \approx xb$, $X^\flat = \exp(x^{\varepsilon/3})$ and $b' = a/X^\flat$ (Definition 5.3.1). A thick weight therefore tests G_θ^{ar} at $a_\tau/w \approx x$, and a thin weight tests it at $a_\tau/w \geq X^\flat/2$ approximately, where the thin-scale bound, an estimate of prime number theorem type, applies (Figure 6.4). The bound is required from $X^\flat/3$ on because the exact lower bound is $a_\tau/w \geq (x-1)b/(2b')$, and the factor $(x-1)/x$ that this loses against $X^\flat/2$ is absorbed by passing from $1/2$ to $1/3$.

Definition 6.2.2 (kept characters and extracted functions). Let D be arithmetic data for $\mathfrak{G} = V_4 \times H$, let ℓ be a function from the characters of $\mathfrak{G}$ to $\mathbb{C}$, let $A_0 \subset H$ be a subgroup, $\mathcal{T} \subset H$ a finite set and $x > 0$. Let $\Psi(A_0)$ be the set of characters of $\mathfrak{G}$ trivial on $0 \times A_0$, let Ξ be a set of characters of $\mathfrak{G}$ (the *kept set*), and let $\varkappa, M, C_o, C_E, \varepsilon$ be real numbers.

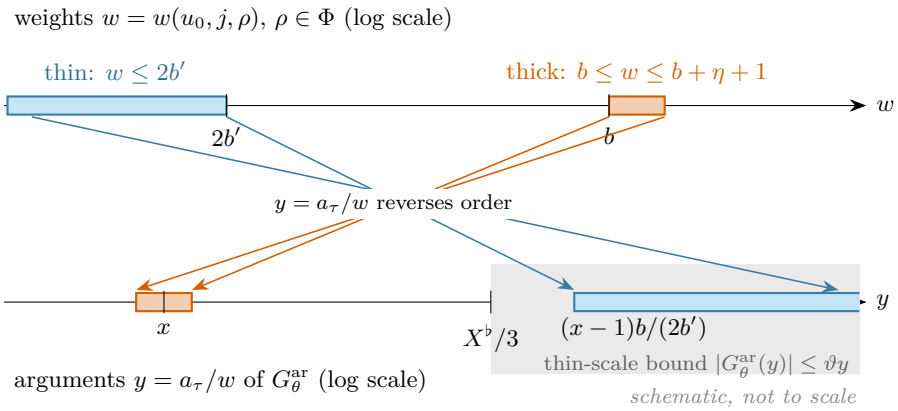


FIGURE 6.4. Schematic, not to scale: where the two kinds of weight test the functions G_θ^{ar} (Definition 6.2.1). Since $|a_\tau/b - x| \le 1$, a thick weight gives an argument a_τ/w near x, which the proof moves to exactly x. A thin weight gives $a_\tau/w \ge (x-1)b/(2b') \ge \frac{x-1}{x} \cdot \frac{X^\flat}{2}$, by $b'X^\flat \le xb$, and the thin-scale bound is assumed from $X^\flat/3$ on (grey). At the sizes used later, $b = \exp(x^\varepsilon)$ and $X^\flat = \exp(x^{\varepsilon/3})$, so the true distances are far larger than drawn.

(1) The *weight* of a character θ is $\mathrm{wt}_A(\theta) = 1$ if θ is trivial on $0 \times H$ and $\mathrm{wt}_A(\theta) = \varkappa$ otherwise.
(2) The *extracted function* and the *rest* are

$$G_A(y) = \sum_{\theta \in \Xi} \mathrm{wt}_A(\theta)\, \mathrm{Re}\, G_\theta^{\mathrm{ar}}(y), \qquad \Sigma_A(y) = \sum_{\theta \in \Psi(A_0) \smallsetminus \Xi} |G_\theta^{\mathrm{ar}}(y)|.$$

(3) The *character conditions* hold if: the trivial character lies in Ξ; every $\theta \in \Xi$ is trivial on $0 \times A_0$ and satisfies $\theta(s_0) = 1$; $|\Psi(A_0)| \le M$; for every $\theta \in \Xi$ nontrivial on $0 \times H$,

$$\Big| \sum_{\tau \in \mathcal{T}} \theta(0,\tau) - \varkappa |\mathcal{T}| \Big| \le M x^{-1+\varepsilon} |\mathcal{T}|;$$

$|\ell(\theta)| \le C_o$ for every nontrivial $\theta \in \Xi$; and $|\ell(\theta)| \le C_E \log x$ for every $\theta \in \Psi(A_0) \smallsetminus \Xi$.

DEFINITION 6.2.3 (rest sizes, costs and the error bracket). Let $\varepsilon, x, b, k, a_E, \eta, \vartheta, f$ be real numbers and n, n_T, n_{T_2}, m natural numbers with $n \ge 1$.

(1) The *rest-size conditions* hold if $a_E \le 2xb$, $x^{-\varepsilon/8} n/2 \le n_T$, $x^{-\varepsilon} \le n_{T_2}/n$, $x^{1/2-\varepsilon} \le m$ and $k \le n^2 x^{1/2-\varepsilon}$.
(2) The *cost conditions* hold if $2f(x+1)\vartheta \le 1/x$ and $6f(x+1)(\eta+1) \le b/x$.

(3) The *error bracket* is

$$\mathrm{Err} = n_T a_E\Big(1+\frac{m^2}{k}\Big) + m\Big(2nmbx^{\varepsilon} + \frac{n_{T_2}kbx^{\varepsilon}}{n^2x^{-\varepsilon}}\Big).$$

Proposition 9.3.1 shows that, under the rest-size conditions, $xm \le k$ and $m \le x^{1/2+\varepsilon/2}$, the bracket divided by $n_T mb$ is of size $x^{1/2+O(\varepsilon)}$. The square root in the final bound comes from that computation.

6.3. Averaging over shifts and top places

The kept characters must be weighted by the weights of a level. The weights come from the distribution of the top places among the cosets of A_0 for a frame ϕ (Definition 5.1.1).

Proposition 6.3.1 (shares give the weights). *Let ϕ be a frame with group A, core A_0 and index $o = [A : A_0]$, let δ be real, let $\mathcal{T} \subset A$ be a finite set with shares up to δ (Definition 5.3.3), and let θ be a character of A trivial on A_0. Put $\varkappa_\theta = 1$ if θ is trivial and $\varkappa_\theta = (3o+1)^{-1}$ otherwise. Then*

$$\Big|\sum_{\tau\in\mathcal{T}} \theta(\tau) - \varkappa_\theta|\mathcal{T}|\Big| \le o\,\delta\,|\mathcal{T}|.$$

Proof. Choose representatives $r_1 = 0, r_2, \dots, r_o$ of the cosets of A_0 in A. Since θ is constant on cosets,

$$\sum_{\tau\in\mathcal{T}} \theta(\tau) = \sum_{k=1}^{o} \theta(r_k)\,\#\{\tau\in\mathcal{T} : \tau - r_k \in A_0\}.$$

By the shares hypothesis and $|\theta| = 1$, this differs from $|\mathcal{T}|\sum_k \theta(r_k)\lambda_\phi(r_k)$ by at most $o\,\delta\,|\mathcal{T}|$. Now $\sum_k \theta(r_k)\lambda_\phi(r_k) = \big(4 + 3\sum_{k\ge2}\theta(r_k)\big)/(3o+1)$. If θ is trivial, this is $(4+3(o-1))/(3o+1) = 1$. Otherwise θ is a nontrivial character of A/A_0, so $\sum_{k=1}^{o}\theta(r_k) = 0$ and $\sum_{k\ge2}\theta(r_k) = -1$, which gives $(4-3)/(3o+1)$. □

With $\delta = x^{-1+\varepsilon}$, Proposition 6.3.1 is the share clause of the character conditions (Definition 6.2.2) with $\varkappa = (3o+1)^{-1}$ and any $M \ge o$. The weights of a level (Definition 4.4.1) come from this proposition, and so does their positivity, which the windowed Landau criterion needs (Remark 4.3.7). Figure 6.5 shows the shares and the averages for $o = 3$.

Proposition 6.3.2 (the extracted function is $G_{\mathscr{L}}$). *Let H be a finite abelian group of order n, $A_0 \subset H$ a subgroup of order n_0, D arithmetic data for $\mathfrak{G} = V_4 \times H$, ℓ a function from the characters of $\mathfrak{G}$ to $\mathbb{C}$, and Ξ a set of characters of $\mathfrak{G}$. Let $\mathscr{L}$ be a level with no exceptional character and with $o(\mathscr{L}) = o = [H : A_0]$, let $\varkappa = (3o+1)^{-1}$, and form $\Psi(A_0)$, G_A and Σ_A (Definition 6.2.2). Suppose that Ξ contains the trivial character and that each nontrivial $\theta \in \Xi$ is nontrivial on $0 \times H$ and is given a primitive character χ_θ, such that $\theta \mapsto \chi_\theta$ is a bijection from the nontrivial elements of Ξ onto $\mathrm{Kept}(\mathscr{L}) \smallsetminus \{1\}$ and $G_\theta^{\mathrm{ar}}(y) = g_{\chi_\theta}(y)$ for $y \ge 1$.*

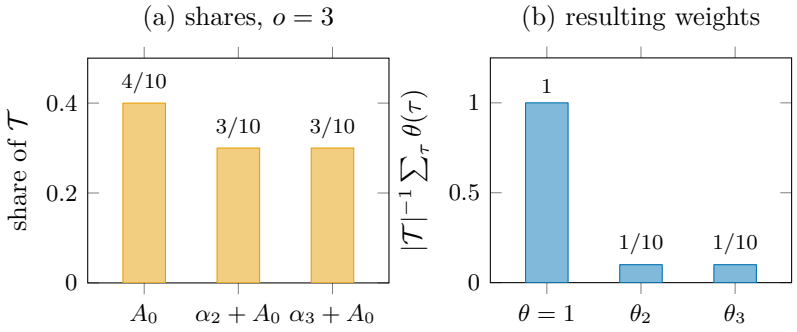


FIGURE 6.5. (a) The shares λ_ϕ of Definition 5.3.3: $4/(3o+1)$ on A_0 and $3/(3o+1)$ on each other coset, here with $o = 3$. (b) The averages over $\mathcal{T}$ of the characters of A/A_0, for top places with these shares: 1 for the trivial character and $(3o+1)^{-1}$ for each nontrivial one, up to $o\,\delta$ (Proposition 6.3.1). These are the positive weights of a level (Definition 4.4.1).

(a) $G_A(y) = G_{\mathscr{L}}(y)$ *for every* $y \geq 1$.
(b) Let $x > 0$ *and* $\varepsilon \in \mathbb{R}$. *If* $|G_\theta^{\mathrm{ar}}(x)| \leq x^{1/2+\varepsilon}$ *for every* $\theta \in \Psi(A_0) \smallsetminus \Xi$, *then* $\Sigma_A(x) \leq 4o\, x^{1/2+\varepsilon}$.

PROOF OF (A). The trivial character lies in Ξ, has weight 1 and contributes $\operatorname{Re} G_{\theta_0}^{\mathrm{ar}} = G_\zeta$. Every other $\theta \in \Xi$ is nontrivial on $0 \times H$, so it has weight $\varkappa$ and contributes $\varkappa \operatorname{Re} g_{\chi_\theta}$. Since $\mathscr{L}$ has no exceptional character, all its weights are $(3o+1)^{-1} = \varkappa$ (Definition 4.4.1), and the bijection matches these terms with the terms of $G_{\mathscr{L}}$. □

PROOF OF (B). The set $\Psi(A_0)$ is the dual of $\mathfrak{G}/(0 \times A_0)$, which has order $4n/n_0 = 4o$. So $\Sigma_A(x)$ has at most $4o$ terms, each at most $x^{1/2+\varepsilon}$. □

The hypothesis that Ξ contains the trivial character is the first of the character conditions; without it G_A would lack the term G_ζ.

THEOREM 6.3.3 (Alpöge). *Let* $\mathscr{L}$ *be a level,* ϕ *a frame, and take a host for* ϕ *(Definition 5.1.2) with its arithmetic data* D *for* $\mathfrak{G} = V_4 \times A$. *Let* $\mathcal{T} \subset A$ *be finite and* $x > 0$, ε *and* C *real, and suppose that the host satisfies the explicit-formula condition* $\mathrm{EF}(\mathscr{L}, \mathcal{T}, x, \varepsilon, C)$ *(Definition 5.3.3). Let* ℓ *be a function from the characters of* $\mathfrak{G}$ *to* $\mathbb{C}$ *such that* $\ell(\theta) = L'(1,\chi)/L(1,\chi)$ *whenever* $\theta \neq \theta_0$ *and* χ *is a primitive Dirichlet character with* $\chi(n) = \theta_D(n)$ *for all* $n \geq 1$. *Let* Ξ *be the set of characters* θ *of* $\mathfrak{G}$ *for which some primitive Dirichlet character* χ *with* $\chi(n) = \theta_D(n)$ *for all* $n \geq 1$ *lies in* $\mathrm{Kept}(\mathscr{L})$, *let* A_0 *be the core of* ϕ, *and let* G_A *and* Σ_A *be the extracted function and the rest (Definition 6.2.2) for* D, ℓ, A_0, Ξ *and* $\varkappa = (3o(\mathscr{L})+1)^{-1}$. *Then:*

(a) $G_A(y) = G_{\mathscr{L}}(y)$ *for every* $y > 0$;
(b) $\Sigma_A(x) \leq 4o(\mathscr{L})\, x^{1/2+\varepsilon}$.

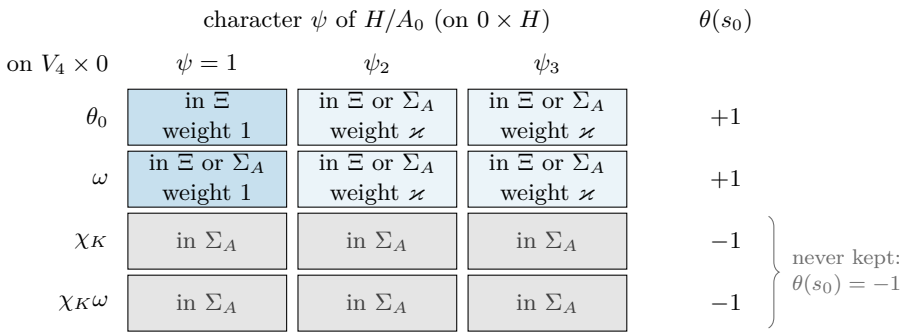


FIGURE 6.6. The set $\Psi(A_0)$ of characters of $\mathfrak{G} = V_4 \times H$ trivial on $0 \times A_0$, drawn for $o = [H : A_0] = 3$. Since $\mathfrak{G}/(0 \times A_0) \cong V_4 \times H/A_0$, each such θ is a character of V_4 (row) times a character ψ of H/A_0 (column), so $|\Psi(A_0)| = 4o$. The value $\theta(s_0)$ depends only on the row, and the character conditions (Definition 6.2.2) put only characters with $\theta(s_0) = 1$ in the kept set Ξ. The weight $\mathrm{wt}_A(\theta)$ is 1 in the column $\psi = 1$ and $\varkappa$ elsewhere. Every character of $\Psi(A_0)$ not in Ξ enters the rest Σ_A. For a host, $\omega \in \Xi$ exactly when the level has an exceptional character, and which of the other light cells lie in Ξ is fixed by $\mathrm{Kept}(\mathscr{L})$ (Theorem 6.3.3).

Theorem 6.3.3 says that for a host the hypotheses of Proposition 6.3.2 are met, so that the extracted function of the host is the weighted Chebyshev function of the level and the rest is of square-root size.

IDEA OF PROOF. We indicate why the statement holds, from the definitions. Both conclusions are term-by-term comparisons: steps (1)–(3) give (a), and step (4) gives (b).

(1) The condition on ℓ, together with the identification of $\theta_D(n)$ with the value at n of the primitive character attached to θ, gives $G_\theta^{\mathrm{ar}} = g_{\chi_\theta}$ for $\theta \neq \theta_0$.
(2) The Artin map of a host is surjective, so θ is determined by χ_θ. The explicit-formula condition asks that every kept character be χ_θ for some $\theta \in X_+$; hence $\theta \mapsto \chi_\theta$ is a bijection from Ξ onto $\mathrm{Kept}(\mathscr{L})$, and $\Xi \subset X_+ \subset \Psi(A_0)$.
(3) The characters of $\mathfrak{G}$ trivial on $0 \times A$ are θ_0, ω, χ_K and $\chi_K\omega$, and only θ_0 and ω take the value 1 at s_0 (Figure 6.6). So the kept characters of weight 1 are θ_0 and, when the level has an exceptional character, ω, with $\chi_\omega = \chi_{\omega_q} = \psi_{\mathrm{exc}}$ because the host keeps the exceptional character. This matches the weights of Definition 4.4.1, and $G_A = G_{\mathscr{L}}$.
(4) A character trivial on A_0 and not kept is bounded (Definition 5.2.1), so the explicit-formula condition gives $|g_{\chi_\theta}(x)| \leq x^{1/2+\varepsilon}$ for it.

Since $o(\phi) = o(\mathscr{L})$, the count of Proposition 6.3.2(b) gives $\Sigma_A(x) \le 4o(\mathscr{L})\, x^{1/2+\varepsilon}$.

□

The core bound converts an upper bound for the gap into an upper bound for the extracted function. By Proposition 6.1.4, for weights positive on Φ and N large enough,

$$\operatorname{gap}(u_0) = \sum_{\tau \in \mathcal{T}} \sum_{j<m} \operatorname{Re} \mathrm{DF}_{\tau, w(u_0,j)}(a_\tau),$$

since PS is real. So an upper bound for the gap is an upper bound for a weighted combination of the functions G_θ^{ar}, evaluated at the points $a_\tau / w(u_0, j, \rho)$.

THEOREM 6.3.4 (Alpöge; core Chebyshev bound). *Let $C \ge 0$, $\varepsilon > 0$, $\varkappa$ with $|\varkappa| \le 1$, and C_E, C_o, M be real numbers. There is x_1 such that the following holds for every $x \ge x_1$. Let H be a finite abelian group, $\mathfrak{G} = V_4 \times H$, m a natural number, D arithmetic data for $\mathfrak{G}$ that are split outside d on a finite set of primes, ℓ a function from the characters of $\mathfrak{G}$ to $\mathbb{C}$, $N \ge 1$ an integer, $A_0 \subset H$ a subgroup, $\mathcal{T}, \mathcal{T}_2 \subset H$ finite sets, $a\colon H \to \mathbb{R}$, $w\colon H \times \{0, \dots, m-1\} \times \mathfrak{G} \to \mathbb{R}$, $b, b', \eta, X^\flat, \vartheta, k, a_E$ real numbers and Ξ a set of characters of $\mathfrak{G}$. Let G_A and Σ_A be the extracted function and the rest (Definition 6.2.2) for D, ℓ, A_0, Ξ and $\varkappa$. Suppose that*

- *(i) $b > 0$, $a_E \ge 0$, $k > 0$, $m \le x^{1/2+\varepsilon/2}$ and $xm \le k$;*
- *(ii) the size conditions and the thin-scale bound of Definition 6.2.1 hold, and $\mathcal{T}_2 \subset A_0$;*
- *(iii) the cost conditions hold with $f = |\Phi|$, and the rest-size conditions hold with $n = |H|$, $n_T = |\mathcal{T}|$ and $n_{T_2} = |\mathcal{T}_2|$ (Definition 6.2.3);*
- *(iv) the character conditions of Definition 6.2.2 hold with the given $\varkappa, M, C_o, C_E, \varepsilon$ and x;*
- *(v) $\operatorname{gap}(u_0) \le x^{C\varepsilon} \operatorname{Err}$ for every $u_0 \in A_0$, with Err the error bracket of Definition 6.2.3 for these $n, n_T, n_{T_2}, m, b, k, a_E$.*

Then $G_A(x) \le x^{1/2+(C+3)\varepsilon} + \Sigma_A(x)$.

PROOF (OUTLINE). The argument has five elementary steps (Figure 6.7). Its only analytic input is the classical prime number theorem behind the thin-scale bound.

(1) Write $\operatorname{gap}(u_0)$ through the balance identity, as displayed above. The thin nodes contribute values of G_θ^{ar} at arguments at least $X^\flat/3$, which the thin-scale bound and the cost conditions make negligible.

(2) At the thick nodes the weight is within $\eta+1$ of b, and $|a_\tau/b - x| \le 1$. Moving each argument a_τ/w to exactly x costs $O(\log x + |\ell(\theta)|)$ per term, by the Lipschitz behaviour of G_θ^{ar}; the character conditions bound $|\ell(\theta)|$.

(3) Averaging over the shifts $u_0 \in A_0$ projects onto the $4o$ characters of $\mathfrak{G}$ trivial on $0 \times A_0$. The kept characters enter with the right sign:

a node $g \notin \Phi$ whose partner $\mathrm{pt}_\tau(g) = ((0,1), t+u_0)$ is thick satisfies $g = s_0+((0,0), t+u_0-\tau)$, and $\theta(s_0) = 1$ for $\theta \in \Xi$ (Definition 6.2.2). So the coefficient $\overline{\theta(g)}$ of G_θ^{ar} in the gap is $\overline{\theta((0,0), t+u_0-\tau)}$, and after the averages over u_0 and τ the kept characters carry the positive weights that Remark 4.3.7 requires. This is why an *upper* bound for the gap gives an *upper* bound for G_A.

(4) Summing over $\tau \in \mathcal{T}$ produces the weights 1 and $\varkappa$, by the share clause of the character conditions (compare Proposition 6.3.1); the characters outside Ξ are collected in $\Sigma_A(x)$.

(5) Divide by $|\mathcal{T}|\, m\, b\, \delta_2/2$, where $\delta_2 = |\mathcal{T}_2|/n$ is the proportion of thick shifts. By Proposition 9.3.1 and the rest-size conditions, the bracket $x^{C\varepsilon}\mathrm{Err}$ then becomes at most $x^{1/2+(C+3)\varepsilon}$ for large x.

□

Theorem 6.3.4, combined with Theorem 6.3.3, turns a gap bound into the bound $G_{\mathscr{L}}(x) \le x^{1/2+(C+3)\varepsilon} + 4o\, x^{1/2+\varepsilon}$. The gap bound is Theorem 9.2.1, and the combination is Theorem 9.4.1.

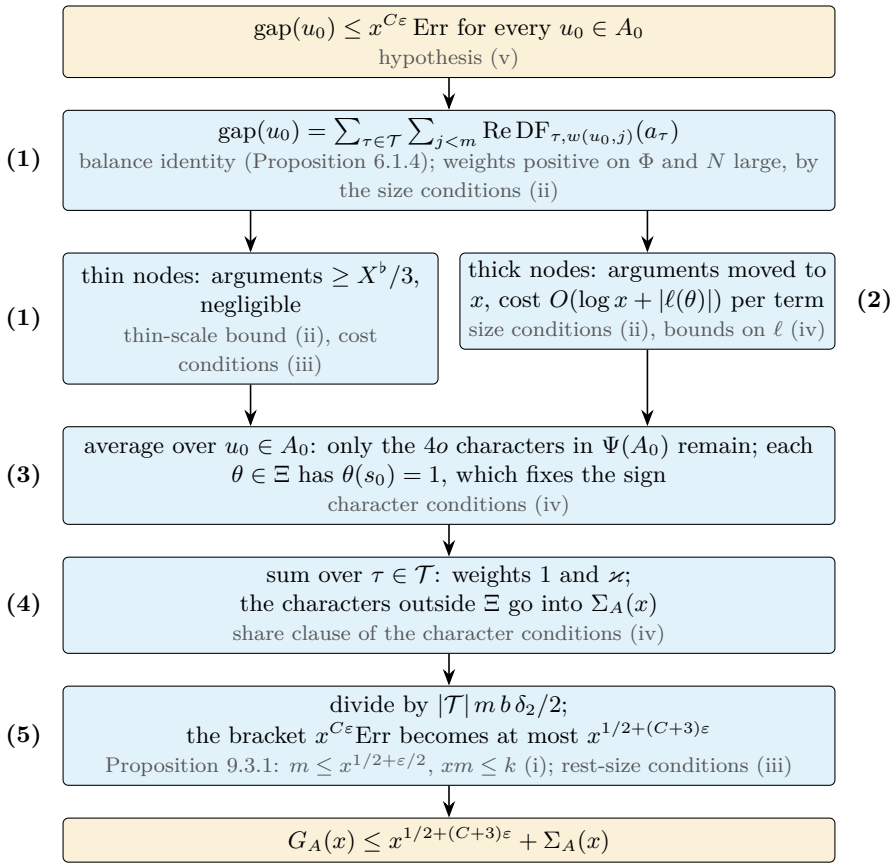


FIGURE 6.7. Schematic of the proof outline of Theorem 6.3.4. Bold numbers are the steps of the outline. The grey line in each box names the result or the hypotheses, numbered (i)–(v) as in the theorem, that the outline cites for the step. Combined with Theorem 6.3.3, the output becomes $G_{\mathscr{L}}(x) \le x^{1/2+(C+3)\varepsilon} + 4o\,x^{1/2+\varepsilon}$.

CHAPTER 7

Hermitian lattices, the theta tower and slope records

This chapter sets up the geometric side of the argument: the tower of hermitian spaces and the theta chain over a host field, hermitian lattices and their slopes, packages, and the slope record, a list of real numbers and inequalities through which the geometry enters the analytic chapters. Its main theorem is the frame theorem, Theorem 7.4.2, a slope comparison between many source lattices and one target lattice. The slope clause of the size data (Definition 7.3.2(ix)) has the shape of its conclusion. The upper and lower links (Definitions 7.5.1 and 7.6.1) record the archimedean and finite sides. In Chapter 3 the frame theorem is one of the inputs that Remark 3.12.2 lists as substantive. The slope record and the links are used in Chapters 8, 9 and 12, where packages are constructed and tied to slope records.

Chebyshev's proof of $\mathrm{Ch}(x) \asymp x$ compares two evaluations of $\log n!$: the archimedean one, by Stirling's formula, and the p-adic ones, by Legendre's formula. Alpöge's argument replaces $n!$ by the periods of theta lifts of algebraic automorphic forms of huge weight $a \approx xb$. On the archimedean side the periods carry Stirling's term $a \log a - a$ and the a-th power of a CM period Ω, whose logarithm contains $L'(1, \chi_{-d})/L(1, \chi_{-d})$. On the finite side they carry a Legendre-type gain at every prime inert in K. A slope inequality for the period map between two hermitian lattices links the two sides, and a chain of $m \approx \sqrt{x}$ lifts makes the total error of relative size about $x^{-1/2}$.

The geometric objects enter the analytic argument of Chapters 6 and 9 only through the slope record. The lattices reappear in Chapter 8, where the arithmetic geometry is concentrated.

7.1. The tower of unitary groups and the theta chain

Definition 7.1.1 (ranks). Let a frame be given (Definition 5.1.1), whose group has order ord.

(1) A *ranks datum* for the frame is a triple $\mathbf{k} = (k, m, t)$ of integers with k even, $m \geq 2$ even and t odd, such that

$$k \geq m+3, \quad t \geq k+m+4, \quad t > 3m, \quad m \leq 2\,\mathrm{ord} \leq k, \quad m^2 \leq k+m(t+1).$$

(2) The *derived quantities* of $\mathbf{k}$ are $n = \mathrm{ord}$, $f = 2n$, $r = k + m - 1$, $n_V = r + t$, $r_S = m/2 - 1$ and $N(j) = k + j(t + 1)$ for $j \geq 0$. The *signature at level* j is $(p_j, q_j) = (k + j - 1, m - j)$.

The last two conditions of (1) read $m \leq f \leq k$ and $m^2 \leq N(m)$. Since $k \geq m+3$, the difference $p_j - q_j = k-1-m+2j$ is positive for every $j \geq 0$.

Definition 7.1.2 (the spaces of the tower). Let E be a host field (Definition 5.1.2) with maximal totally real subfield F, and let $\mathbf{k} = (k, m, t)$ be a ranks datum with derived quantities $r, n_V, r_S, N(j)$ (Definition 7.1.1). The *spaces* of $\mathbf{k}$ over E/F consist of the following hermitian spaces over E/F, each with a unimodular base lattice.

(1) For $0 \leq j \leq m$, a definite space X_j of rank $N(j)$.
(2) A space X' of rank $r + mt$, an isometric embedding $X' \hookrightarrow X_m$, and an anisotropic vector $z \in X_m$ orthogonal to its image.
(3) For $0 \leq j \leq m$, a space Y_j of rank $r + jt$, with isometric embeddings $X_j \hookrightarrow Y_j$ for $j < m$, an isometry $X' \cong Y_m$, and a decomposition $Y_0 = X_0 \oplus \langle -T \rangle \oplus \mathbb{H}^{r_S}$, where $\mathbb{H}$ is a hyperbolic plane and T is a prime.
(4) For $1 \leq j \leq m$, a space W_j of rank r and signature $(p_j, q_j) = (k + j - 1, m - j)$ at every archimedean place, locally isometric to Y_0 at every finite place.
(5) For $1 \leq j \leq m$, a space V_j of rank n_V carrying a lattice with an orthonormal basis up to sign.

The hermitian forms, the lattices and the isometries are part of the data. We do not write them out. Nothing below depends on more than the ranks and signatures listed. The spaces X_j are definite, so their unitary groups $U(X_j)$ carry algebraic automorphic forms in the sense of Gross [**Gro99**]. The spaces W_j are indefinite for $j < m$, and their unitary groups carry the Shimura varieties of Chapter 8.

Remark 7.1.3 (the top of the tower). The embeddings $X_j \hookrightarrow Y_j$ of Definition 7.1.2(3) stop below the top of the tower: there is no embedding $X_m \hookrightarrow Y_m$. The ranks satisfy

$$\mathrm{rk}\, Y_j - \mathrm{rk}\, X_j = (r + jt) - N(j) = m - 1 - j.$$

An isometric embedding of a nondegenerate space is injective, so such an embedding exists only for $j \leq m - 1$; at $j = m - 1$ the two ranks are equal. At the top the direction reverses: $Y_m \cong X'$ embeds in X_m, whose rank $N(m) = \mathrm{rk}\, Y_m + 1$ is one larger. By the rank count, X_m is the orthogonal sum of the image of X' and the line spanned by the anisotropic vector z.

Figure 7.1 shows the spaces for small ranks.

Definition 7.1.4 (forms, theta kernels and the column; schematic). Let $\mathbf{k} = (k, m, t)$ be a ranks datum and fix spaces of $\mathbf{k}$ over E/F (Definition 7.1.2). For $0 \leq i \leq m$ let cls_i be the class set of the genus of a unimodular lattice in X_i, weighted by the masses $1/\#\,\mathrm{Stab}$.

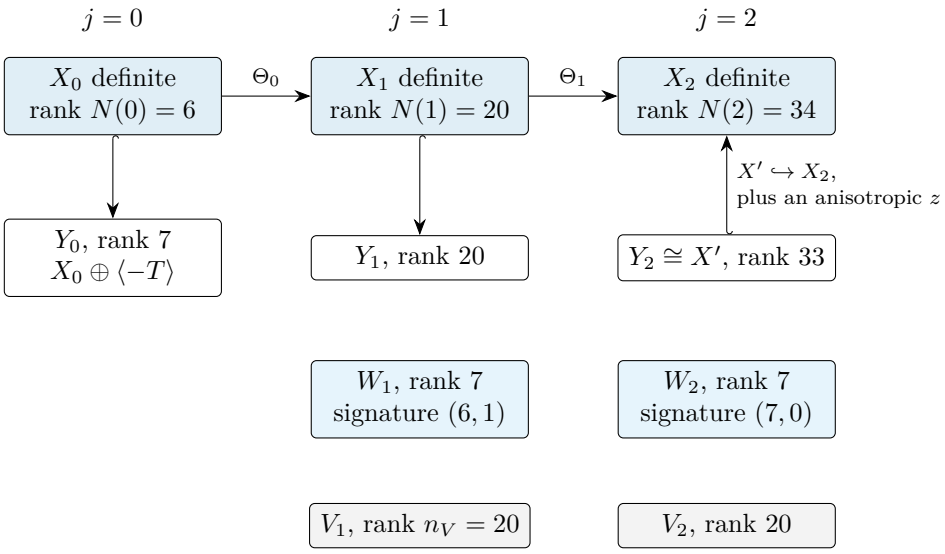


FIGURE 7.1. The spaces of Definition 7.1.2 for the small ranks $(k, m, t) = (6, 2, 13)$. They satisfy the conditions of Definition 7.1.1: k and m are even, t is odd, $m + 3 \le k$, $k+m+4 \le t$, $3m < t$, $m^2 \le k+m(t+1)$, and $m \le 2\,\mathrm{ord} \le k$ for $\mathrm{ord} \in \{1, 2, 3\}$. Here $r = k+m-1 = 7$, $N(j) = k+j(t+1)$, $\mathrm{rk}\, Y_j = r + jt$ and $n_V = r + t$. The space W_j has signature $(k + j - 1, m - j)$ and is definite at $j = m$; Y_0 has signature $(6, 1)$. The embeddings $X_j \hookrightarrow Y_j$ exist only for $j < m$ (Remark 7.1.3); at $j = m$ the space $X' \cong Y_m$ embeds in X_m, whose rank is one larger. The arrows Θ_j are the theta kernels of Definition 7.1.4.

(1) A *form at level* i is a function $\mathrm{cls}_i \times Y_i \to \mathbb{C}$.
(2) The *theta kernel* Θ_i, for $0 \le i < m$, is the linear map from forms at level i to forms at level $i + 1$ given by

$$(\Theta_i \phi)(\Lambda') = \sum_{\Lambda \in \mathrm{cls}_i} \frac{\upsilon_i(\Lambda)\, \upsilon_i'(\Lambda')}{\#\,\mathrm{Stab}(\Lambda)}\, K_i(\Lambda, \Lambda')(\phi(\Lambda)) \qquad (\Lambda' \in \mathrm{cls}_{i+1}),$$

where υ_i, υ_i' are functions of absolute value 1 on the class sets (the splitting characters), and $K_i(\Lambda, \Lambda')$ is a linear map from functions on Y_i to functions on Y_{i+1}, the theta kernel between the lattices Λ and Λ', composed on both sides with multiplication by fixed nonvanishing functions (the archimedean normalisations).
(3) The *column* is the m-fold composite $\mathrm{col} = \Theta_{m-1} \circ \cdots \circ \Theta_0$, the *theta chain* from $U(X_0)$ to $U(X_m)$.
(4) Let L_z be the map that evaluates a form at level m at the anisotropic vector $z \in X_m$ of the spaces and then applies a fixed linear lowering map to a space of row vectors, and let $\langle\ ,\ \rangle$ be a

fixed bilinear pairing of row vectors with test forms. The *period* of a form ϕ at level m against a test form F is $P(\phi, F) = \langle \mathrm{L}_z\phi, F\rangle$. The *period map* of a form ϕ at level 0 is $F \mapsto P(\operatorname{col}\phi, F)$.

Thus Θ_i lifts forms on $U(X_i)$ to forms on $U(X_{i+1})$, and the period map pairs the lift of a form on $U(X_0)$ to the top of the tower with test forms. In terms of Arthur parameters each step of the chain may be pictured as $\pi \mapsto \pi \boxplus \mu[t+1]$. This picture is only a guide. The kernels K_i, the lowering map and the pairing are not written out (Section 4.1); the argument uses the column only through the period maps B_u of Definition 7.2.2.

7.2. Hermitian lattices and slopes

Definition 7.2.1 (hermitian lattices, slopes and heights). Let L be a number field.

(1) A *hermitian vector bundle* $\overline{\mathcal{E}}$ over $\operatorname{Spec}\mathcal{O}_L$ is a finitely generated projective $\mathcal{O}_L$-module $\mathcal{E}$ with a hermitian norm on $\mathcal{E}\otimes_\sigma \mathbb{C}$ for each embedding $\sigma\colon L \hookrightarrow \mathbb{C}$, the norms at σ and $\bar\sigma$ being conjugate.

(2) A *hermitian lattice* over $\mathcal{O}_L$ is a finite-dimensional L-vector space with a basis, a full-rank $\mathcal{O}_L$-submodule, and a positive hermitian Gram matrix at each infinite place. It is a hermitian vector bundle over $\operatorname{Spec}\mathcal{O}_L$.

(3) The *normalised Arakelov degree* is $\widehat{\deg}_n = \widehat{\deg}/[L:\mathbb{Q}]$. The *slope* of $\overline{\mathcal{E}} \neq 0$ is $\hat\mu(\overline{\mathcal{E}}) = \widehat{\deg}_n(\overline{\mathcal{E}})/\operatorname{rk}\mathcal{E}$, and its *maximal* and *minimal slopes* $\hat\mu_{\max}(\overline{\mathcal{E}})$ and $\hat\mu_{\min}(\overline{\mathcal{E}})$ are those of its Harder–Narasimhan filtration [**Stu76, Gra84**]. For a hermitian lattice N we write $\mu(N)$, $\mu_{\max}(N)$, $\mu_{\min}(N)$, and $N^\vee$ for the dual lattice.

(4) For hermitian vector bundles $\overline{\mathcal{E}}$, $\overline{\mathcal{F}}$ and an L-linear map $\varphi\colon \mathcal{E}_L \to \mathcal{F}_L$, the *height* of φ is

$$h(\overline{\mathcal{E}}, \overline{\mathcal{F}}, \varphi) = \frac{1}{[L:\mathbb{Q}]}\Big(\sum_{\mathfrak{p}} \log\|\varphi\|_{\mathfrak{p}} + \sum_{\sigma\colon L\hookrightarrow\mathbb{C}} \log\|\varphi\|_\sigma\Big),$$

the first sum over the finite places of L. For hermitian lattices N, R and $\varphi \neq 0$ we write $\operatorname{ht}(\varphi)$.

Definition 7.2.2 (period lattice data). Let E be a host field and $L \supset E$ a CM field, and let $\mathcal{U}$ be a finite nonempty set, whose elements we call *words*.

(1) *Period lattice data* over L indexed by $\mathcal{U}$ consist of a hermitian lattice R over $\mathcal{O}_L$ (the *test forms*); for each $u \in \mathcal{U}$ a hermitian lattice N_u (the *ramified forms*), an L-linear map $B_u\colon N_u \otimes L \to R^\vee \otimes L$ (the *period map*) and a real number h_u; and a real number Δ_R (the *spread*). They are required to satisfy $B_u(N_u) \subset R^\vee$, and $\operatorname{ht}(B_u) \le h_u$ whenever $B_u \neq 0$ (Definition 7.2.1(4)).

(2) The *frame slope* of the data is

$$s = \frac{\sum_{u\in\mathcal{U}} \operatorname{rk}(N_u)\big(\mu(N_u) - h_u\big)}{\sum_{u\in\mathcal{U}} \operatorname{rk}(N_u)}.$$

The maps B_u are not required to be injective.

DEFINITION 7.2.3 (packages; schematic). A *package* $\mathcal{P}$ consists of the following data.

(1) A ranks datum $\mathbf{k} = (k, m, t)$ (Definition 7.1.1) and a host (Definition 5.1.2) with $\sqrt{-q} \in E$, together with: a *shift* $u_0 \in A$; size bounds; small primes split in K; an ideal $\mathfrak{b}$ with $\mathfrak{d}_{E/K} = \mathfrak{b}\mathfrak{b}^c$; the norm primes ϖ_0, ϖ_1 of the host (Definition 5.3.2); an ideal of $\mathcal{O}_K$ supported on split primes; the top places $\mathcal{T}$ with top sizes a_τ ($\tau \in \mathcal{T}$); the set $\mathcal{T}_2$; weights $w(j, \rho) > 0$ for the lift indices $1 \le j \le m$ and $\rho \in \Phi$; and a finite set $\mathcal{U}$ of words, each with a level prime $T(u)$.
(2) Spaces of $\mathbf{k}$ over E/F (Definition 7.1.2).
(3) Coefficient data: a CM field $L \supset E$, archimedean places, local twist types, finite-dimensional Hilbert spaces with lattices and idempotents, spaces of forms on $U(W_j)$ with integral structures, theta maps from the definite groups $U(V_j)$, eigenvalues, nonnegative congruence invariants, and nonnegative rational numbers $\mathrm{len}(l, e)$ (the lengths used in Chapter 10).
(4) For each word u, a column (Definition 7.1.4) with its period map.
(5) Period lattice data over L indexed by $\mathcal{U}$ (Definition 7.2.2).
(6) Form lattices, with isometric embeddings of the lattices N_u into them.
(7) Local p-adic and Galois-side data, a theta jet, and compatibility conditions among all of these data.

The definition is schematic (Section 4.1): the compatibility conditions among these data are omitted. Informally, a package is the automorphic input at one scale: a host field, the tower of hermitian spaces over it, forms on the unitary groups of the tower with their integral structures and theta lifts, and the lattices and period maps through which the lifts are measured. A package contains no inequality of the kind collected in the slope record (Definition 7.3.2). The links of Definitions 7.5.1 and 7.6.1 attach the inequalities to a package.

DEFINITION 7.2.4 (size data; schematic). Let $\mathcal{P}$ be a package, $x > 1$ and $\varepsilon > 0$. A *size datum* for $\mathcal{P}$ at scale x and ε consists of a scale (x, ε, J) (Definition 5.3.1) and of local sizes of $\mathcal{P}$ at the places of a finite set of places of its coefficient field, the *support*.

The local sizes are real numbers attached to the places $\mathfrak{p}$ of the support and vanishing off it: for each lift index j a *gain* $\mathrm{gain}_j(\mathfrak{p})$, a *field term* $F_j(\mathfrak{p})$ and, for each word u, a *congruence term* $\mathrm{Cg}_{u,j}(\mathfrak{p})$; and a *loss* $\mathrm{ta}_\mathfrak{p}$ and an *orbit term* $\mathrm{io}_\mathfrak{p}$. A local quantity y at $\mathfrak{p}$ contributes the global size

$$\mathrm{sz}_\mathfrak{p}(y) = 2e_\mathfrak{p} \frac{\log N\mathfrak{p}}{[L : \mathbb{Q}]} y,$$

where L is the coefficient field of $\mathcal{P}$ and $e_{\mathfrak{p}}$ the ramification index of $\mathfrak{p}$ over $\mathbb{Q}$. The *field term* and the *finite-side gain* of the size datum are

$$\mathrm{Fo} = \sum_j \sum_{\mathfrak{p}} \mathrm{sz}_{\mathfrak{p}}(F_j(\mathfrak{p})), \qquad \mathrm{Ga} = \sum_{\mathfrak{p}} \mathrm{sz}_{\mathfrak{p}}\Big(\sum_j \mathrm{gain}_j(\mathfrak{p}) - \mathrm{ta}_{\mathfrak{p}}\Big).$$

The constraints that a size datum imposes on its local sizes are omitted (Section 4.1).

The comparison of slopes that the argument rests on goes back to Bost.

THEOREM 7.2.5 (Bost [**Bos01**]). *Let L be a number field, let $\overline{\mathcal{E}}$ and $\overline{\mathcal{F}}$ be hermitian vector bundles over $\operatorname{Spec}\mathcal{O}_L$, and let $\varphi\colon \mathcal{E}_L \to \mathcal{F}_L$ be an injective L-linear map. Then*

$$\widehat{\deg}_n(\overline{\mathcal{E}}) \le \operatorname{rk}\mathcal{E}\,(\hat{\mu}_{\max}(\overline{\mathcal{F}}) + h(\overline{\mathcal{E}}, \overline{\mathcal{F}}, \varphi)).$$

Equivalently, if $\mathcal{E} \neq 0$, then $\hat{\mu}(\overline{\mathcal{E}}) \le \hat{\mu}_{\max}(\overline{\mathcal{F}}) + h(\overline{\mathcal{E}}, \overline{\mathcal{F}}, \varphi)$. See [**Bos01**, §4.1.4, Prop. 4.5, (4.10)–(4.11)]. The inequality originates in [**Bos96**, §4.3.1, Prop. 4.3(1)], where φ is a morphism of $\mathcal{O}_L$-modules and only archimedean norms appear; see also [**Che10**].

REMARK 7.2.6 (the period map as the auxiliary map). In transcendence proofs, the map φ of Theorem 7.2.5 evaluates small sections that one first constructs. Here its role is played by the period map $B_u\colon N_u \to R^\vee$ of Definition 7.2.2. No small sections are constructed. Instead one shows that the period map between two given lattices is integral and of bounded height, and applies a slope inequality to it.

7.3. The slope record

The geometric objects enter the analytic argument only through the following list of real numbers and inequalities. Definitions 7.3.1–7.3.3 are shared with Chapters 6 and 9.

DEFINITION 7.3.1 (slope records). Let m and f be positive integers.

(1) A *slope record* consists of finite sets U_0 (the *shifts*), $\mathcal{U}(u_0) \neq \emptyset$ for $u_0 \in U_0$ (the *words* at u_0) and $\mathcal{T}$ (the *top places*), together with the following real numbers:
   (a) $x, \varepsilon, a, b, b', k, \Upsilon, \Upsilon_a, B_0, N_c, \gamma^*, \ell_K, v, \lg$ and $\bar{\beta}_0, \bar{\beta}_1, \bar{\beta}_2$;
   (b) the *top sizes* a_τ for $\tau \in \mathcal{T}$, and the *weights* $w(u_0, j, \rho)$ for $u_0 \in U_0$, $0 \le j < m$ and $0 \le \rho < f$;
   (c) for $u_0 \in U_0$ and $u \in \mathcal{U}(u_0)$: $\mathrm{rk}(u)$, $T(u)$, $\mu_N(u)$, $\mu_N^{\max}(u)$, $h(u)$, $\mu_M(u)$, $\mu_M^{\max}(u)$, and $\mathrm{Cg}_j(u)$ for $0 \le j < m$;
   (d) for $u_0 \in U_0$: $\mu_R(u_0)$, $\mu_R^{\min}(u_0)$, $\mathrm{Fo}(u_0)$, $\mathrm{Ga}(u_0)$, $\mathrm{PS}(u_0)$, $\mathrm{IO}(u_0)$, $\Delta_R(u_0)$, $\beta_0(u_0)$, $\beta_1(u_0)$ and $\beta_2(u_0)$.

(2) The *top count* and the *top total* of the record are $n_T = |\mathcal{T}|$ and $\Sigma a = \sum_{\tau\in\mathcal{T}} a_\tau$.

(3) The *congruence size* of a word u is $\mathrm{Cg}(u) = \sum_{j<m} \mathrm{Cg}_j(u)$.

(4) The *frame slope* at $u_0 \in U_0$ is
$$s(u_0) = \frac{\sum_{u \in \mathcal{U}(u_0)} \mathrm{rk}(u)\big(\mu_N(u) - h(u)\big)}{\sum_{u \in \mathcal{U}(u_0)} \mathrm{rk}(u)},$$
defined when the denominator is nonzero.
(5) The *excess* at $u_0 \in U_0$ is $u_F(u_0) = \max_{u \in \mathcal{U}(u_0)}\big(\mu_N^{\max}(u) - h(u)\big) - s(u_0)$.
(6) The *period* of the record is the real number $\Omega > 0$ with $\log \Omega = \gamma^* + \log(\pi/v) - \ell_K$, defined when $v > 0$.

The record is a list of real numbers, but each entry stands for a quantity of the geometry, and the clauses of Definition 7.3.2 below compare these quantities. Once the record is tied to a family (Definition 9.1.4), the entries have the meanings listed in Table 7.1. Two conventions fix the indexing. The record index j, with $0 \le j < m$, corresponds to the lift index $j + 1$ of a package, whose weights $w(j + 1, \rho)$ are indexed by $1 \le j + 1 \le m$. The index $\rho < f$ of the record corresponds to a node of Φ through the bijection e_P of the identification clause (Definition 9.1.1). Likewise the top places of the record correspond to the top places $\mathcal{T}$ of a host through the bijection e_T of that clause, and we write $\mathcal{T}$ for both.

symbol	meaning once tied	clauses
a_τ, n_T, Σa	top sizes, the number of top places, their total	(ii)–(v)
$w(u_0, j, \rho)$	weight of the member at the shift u_0, lift index $j+1$, node $e_P(\rho)$	(v)
$\mathrm{rk}(u)$, $T(u)$	rank of the ramified lattice N_u; level prime of the word u	(ii), (iv), (v)
$\mu_N(u)$, $\mu_N^{\max}(u)$	slope and maximal slope of N_u	(vii), (ix)
$h(u)$	the height bound h_u for the period map B_u	(ii), (iv), (ix)
$\mu_M(u)$, $\mu_M^{\max}(u)$	slope and maximal slope of the form lattice of u	(vi), (vii)
$\mathrm{Cg}_j(u)$, $\mathrm{Cg}(u)$	congruence size of u at lift index $j+1$, and its sum over j	(iii), (vii)
μ_R, $\mu_R^{\min}$	slope and minimal slope of $R^\vee$, the dual of the lattice of test forms	(ii), (ix), (x)
Fo	field term of the size datum	(ii), (vii)
Ga	finite-side gain of the size datum, the Legendre-type gain at the inert places	(vii), (viii)
PS	the prime sum $\sum_\tau \sum_j \mathrm{PS}_{\tau,w}(a_\tau)$ of the arithmetic data	(viii)
IO	orbit term: the sum of the global sizes of the orbit terms $\mathrm{io}_{\mathfrak{p}}$	(vii), (viii)
γ^*, ℓ_K, v	Euler's constant, $L'(1,\chi_{-d})/L(1,\chi_{-d})$, the half root discriminant; they give Ω	(ii)
lg	$\log(\varpi_0\varpi_1)$ for the norm primes	(iii), (v)
Υ, Υ_a, B_0, N_c	size quantities of the size datum	(ii)–(vii)
Δ_R, β_1, β_2	spread and frame defects; set by no tie	(ix)–(xi)
β_0	rank defect of the member	(ix), (xi)

TABLE 7.1. The entries of a slope record (Definition 7.3.1) and their meanings once the record is tied (Definition 9.1.4), with the clauses of Definition 7.3.2 in which they occur.

DEFINITION 7.3.2 (size data of a slope record). Let a slope record be given (Definition 7.3.1), with positive integers m and f.

(1) *Size constants* are real numbers

$$c_{\mathrm{div}},\ e_{\mathrm{div}},\ c_{\mathrm{cg}},\ c_{\mathrm{ht}},\ c_{\mathrm{ct}},\ c_{\mathrm{fs}},\ c_{\mathrm{qs}},\ c_{\mathrm{loss}},\ e_{\mathrm{loss}},\ c_{\mathrm{sp}} \quad \text{and} \quad X^\sharp.$$

(2) The *losses* of the record for given size constants are

$$\mathrm{Lo}_P = c_{\mathrm{loss}}\, n_T m b\, x^{e_{\mathrm{loss}}\varepsilon},$$
$$\mathrm{Lo}_Q(u) = c_{\mathrm{qs}}\mathrm{IO}(u_0) + c_{\mathrm{qs}}\, n_T m\Big(\frac{a}{k} + b'\Big)\log T(u) + c_{\mathrm{qs}}X^\sharp + 1$$

for $u_0 \in U_0$ and $u \in \mathcal{U}(u_0)$.

(3) The record *satisfies the size data* for given size constants if for every $u_0 \in U_0$ and all $u, u' \in \mathcal{U}(u_0)$ the following clauses hold, where s, u_F, μ_R, $\mu_R^{\min}$, Fo, Ga, PS, IO, Δ_R and $\beta_0, \beta_1, \beta_2$ are taken at u_0.

(i) (Positivity.) $x, b, k, v, f > 0$; $a, b' \geq 0$; $a_\tau \geq 0$ for every τ; $w(u_0, j, \rho) > 0$ for all j, ρ; $\mathrm{rk}(u) > 0$; $T(u) > 0$; $\mu_N(u) \leq \mu_N^{\max}(u)$; $c_{\mathrm{qs}} \geq 0$; $\bar\beta_0 < 1$; $\bar\beta_1 \geq 0$; and $0 \leq \bar\beta_2 \leq 1/4$.

(ii) (Divisor bound.)

$$\mu_R + h(u) - \mathrm{Fo} \leq \sum_{\tau \in \mathcal{T}} \Big(m(a_\tau \log a_\tau - a_\tau - a_\tau \log \Omega) - a_\tau \log T(u) \Big) + c_{\mathrm{div}}\, n_T a \Upsilon_a + c_{\mathrm{div}}\, m(B_0 + N_c)\, x^{e_{\mathrm{div}}\varepsilon}.$$

(iii) (Congruence sizes.) $|\mathrm{Cg}(u) - m\,\Sigma a\ \mathrm{lg}\,| \leq c_{\mathrm{cg}}\, m B_0$.

(iv) (Heights.) $\big|\big(h(u) + \Sigma a \log T(u)\big) - \big(h(u') + \Sigma a \log T(u')\big)\big| \leq c_{\mathrm{ht}}\,(n_T a \Upsilon_a + m B_0)$.

(v) (Counts.)

$$\sum_{j<m} \frac{1}{f} \sum_{\rho<f} \log \frac{w(u_0, j, \rho)\, v}{\pi} + m \lg - c_{\mathrm{ct}} \Big(\Upsilon_a + \frac{m^2 \Upsilon_a}{k} \Big) \leq \log T(u) \leq c_{\mathrm{ct}}\, m \Upsilon,$$

and $|\log T(u) - \log T(u')| \leq c_{\mathrm{ct}} \log k$.

(vi) (Form slopes.) $-c_{\mathrm{fs}}\, n_T a \Upsilon_a (1 + m/k) \leq \mu_M(u)$ and $\mu_M^{\max}(u) \leq c_{\mathrm{fs}}\, n_T a \Upsilon_a$.

(vii) (Quotient slopes.) $\mu_N^{\max}(u) \leq \mu_M^{\max}(u) + \mathrm{Ga} + \mathrm{Fo} - \mathrm{Cg}(u) + c_{\mathrm{qs}}\, m B_0$ and $\mu_M(u) + \mathrm{Ga} + \mathrm{Fo} - \mathrm{Cg}(u) - \mathrm{Lo}_Q(u) \leq \mu_N(u)$.

(viii) (Losses.) $\mathrm{PS} - \mathrm{Ga} \leq \mathrm{Lo}_P$ and $\mathrm{IO} \leq \mathrm{Lo}_P$.

(ix) (Slope inequality.) If $0 \leq \beta_0 < 1$, $\beta_1 \geq 0$, $0 \leq \beta_2 \leq 1/4$, $\Delta_R > 0$ and $s - \mu_R^{\min} \leq \Delta_R$, then

$$s - \log |\mathcal{U}(u_0)| - \frac{1}{1 - \beta_0} \Big(4 u_F \log \Big(1 + \frac{\Delta_R}{u_F} \Big) + (\beta_1 + 4\sqrt{\beta_2}) \Delta_R \Big) \leq \mu_R,$$

where the term $4u_F \log(1 + \Delta_R/u_F)$ is read as 0 when $u_F = 0$.

(x) (Spread.) $s - \mu_R^{\min} \leq \Delta_R$ and $0 < \Delta_R \leq x^{c_{\mathrm{sp}}} b$.

(xi) (Defects.) $0 \leq \beta_0 \leq \bar\beta_0$, $0 \leq \beta_1 \leq \bar\beta_1$ and $0 \leq \beta_2 \leq \bar\beta_2$.

Once the record is tied to a family (Definition 9.1.4), μ_R and $\mu_R^{\min}$ are the slope and the minimal slope of $R^\vee$, $\mu_N(u)$ and $\mu_N^{\max}(u)$ are the slope and maximal slope of N_u, and $h(u) = h_u$. The numbers IO and $\mathrm{Cg}_j(u)$ become sums of local sizes over the support of a size datum, and PS becomes the sum of the prime sums $\mathrm{PS}_{\tau,w}$ of Definition 6.1.3. The spread Δ_R and the defects β_1, β_2 are set by no tie. They are constrained only by the clauses above. Figure 7.2 shows which entries of the record each clause connects. The hypotheses of clause (ix) follow from clauses (i), (x) and (xi).

Definition 7.3.3 (main term of a slope record). Let a slope record be given (Definition 7.3.1) whose top sizes a_τ and weights $w(u_0, j, \rho)$ are

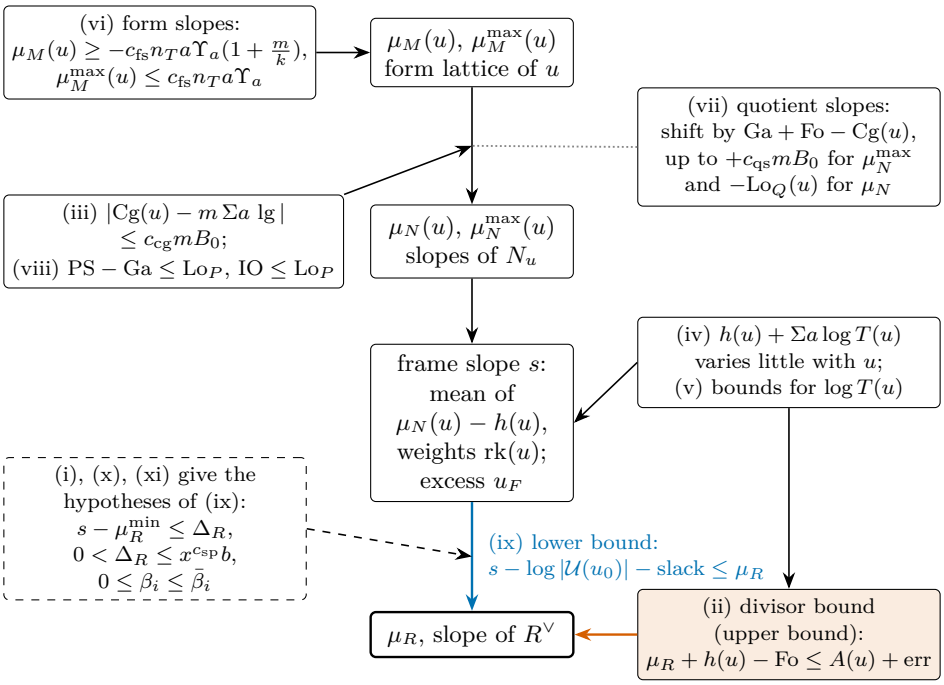


FIGURE 7.2. Schematic map of the clauses of Definition 7.3.2 at one shift u_0 and one word u. Arrows show which quantities each clause relates; they are not implications. Clause (vi) bounds the slopes of the form lattice, clause (vii) passes to the slopes of N_u, shifted by the gain Ga, the field term Fo and the congruence size $\mathrm{Cg}(u)$, and the frame slope s and the excess u_F are formed from the slopes of N_u and the heights $h(u)$ (Definition 7.3.1). Clause (ix) (blue) bounds μ_R from below by s minus the slack $\log|\mathcal{U}(u_0)| + (1-\beta_0)^{-1}\big(4u_F \log(1 + \Delta_R/u_F) + (\beta_1 + 4\sqrt{\beta_2})\Delta_R\big)$, and clause (ii) (vermillion) bounds μ_R from above; there $A(u)$ is the archimedean main term of Definition 7.5.1 and err the remaining terms of the clause. In Chapter 8 the second inequality of clause (vii) and clause (ii) give a lower bound for s, clause (ix) gives an upper bound, and μ_R and the heights cancel (Remark 9.2.2 and Figure 9.2).

positive. Its *main term* at $u_0 \in U_0$ is

$$\mathrm{Main}(u_0) = \sum_{j<m} \sum_{\tau \in \mathcal{T}} a_\tau \Big(\log a_\tau - 1 - \gamma^* + \ell_K - \frac{1}{f} \sum_{\rho<f} \log w(u_0, j, \rho) \Big).$$

7.4. The frame theorem

The slope clause (ix) has the shape of the conclusion of the following theorem. It compares a rank-weighted mean of the slopes of many source lattices with the slope of one target lattice.

DEFINITION 7.4.1 (frame operator). Let J and ι_i, for i in a finite set I, be finite sets. For each $i \in I$ let $\widetilde{B}_i$ be a complex matrix with rows indexed by J and columns by ι_i, and let c_i be a real number.

(1) The *frame operator* of the weighted family $(c_i, \widetilde{B}_i)_{i\in I}$ is $Q = \sum_{i\in I} c_i \widetilde{B}_i \widetilde{B}_i^*$, an operator on $\mathbb{C}^J$.
(2) The *Hilbert–Schmidt defect* of the family is $\sum_{i\in I} \|c_i \widetilde{B}_i^* \widetilde{B}_i - 1\|_{\mathrm{HS}}^2$, where 1 is the identity of $\mathbb{C}^{\iota_i}$ and $\|A\|_{\mathrm{HS}}^2 = \operatorname{tr}(A^*A)$.

THEOREM 7.4.2 (Alpöge; frame theorem). *Let L be a totally complex number field and $\sigma\colon L \hookrightarrow \mathbb{C}$ an embedding. Let W be an L-vector space with a basis indexed by a finite set J, with a positive definite hermitian Gram matrix at each infinite place, and let $R \subset W$ be a finitely generated $\mathcal{O}_L$-submodule spanning W. Let I be a nonempty finite set, and for $i \in I$ let V_i be an L-vector space with a basis indexed by a finite set ι_i and positive definite hermitian Gram matrices, $N_i \subset V_i$ a finitely generated $\mathcal{O}_L$-submodule spanning V_i, $B_i\colon V_i \to W$ an L-linear map, and h_i a real number; assume $\sum_i |\iota_i| > 0$. Let J_1 be a finite set, and let Y' and Y_i ($i \in I$) be invertible complex matrices indexed by J and by ι_i. For each i let $\widetilde{B}_i$ be a complex matrix with rows indexed by $J \sqcup J_1$ and columns by ι_i, whose J-rows form the matrix $Y'^{-1}[B_i]_\sigma Y_i$, where $[B_i]_\sigma$ is the matrix of B_i in the given bases, embedded by σ. Let $c_i > 0$ for $i \in I$, let Q be the frame operator of $(c_i, \widetilde{B}_i)$ (Definition 7.4.1), and let $s, u, \Delta, \beta_0, \beta_1, \beta_2$ be real numbers. Suppose:*

(a) $B_i(N_i) \subset R$ for all i, and $\operatorname{ht}(B_i) \le h_i$ whenever $B_i \neq 0$;
(b) $s\sum_i \operatorname{rk} N_i \le \sum_i \operatorname{rk} N_i\,(\mu(N_i) - h_i)$; $u \ge 0$ and $\mu_{\max}(N_i) - h_i \le s + u$ for all i; $\Delta \ge 0$ and $s - \mu_{\min}(R) \le \Delta$;
(c) $0 \le \beta_0 < 1$ and $(1-\beta_0)|J \sqcup J_1| \le |J|$;
(d) the Hilbert–Schmidt defect of $(c_i, \widetilde{B}_i)$ is at most $\beta_2 \sum_i \operatorname{rk} N_i$, with $0 \le \beta_2 \le 1/4$; and $\operatorname{tr}(Q^2) \le (1+\beta_1)(\operatorname{tr} Q)^2/|J \sqcup J_1|$, with $\beta_1 \ge 0$.

Then

$$s - \log|I| - \frac{1}{1-\beta_0}\Big(4u\log\Big(1+\frac{\Delta}{u}\Big) + (\beta_1 + 4\sqrt{\beta_2})\Delta\Big) \le \mu(R),$$

where the term $4u\log(1+\Delta/u)$ is read as 0 when $u = 0$.

So (W, R) and the (V_i, N_i) are hermitian lattices in the sense of Definition 7.2.1, and the slopes μ, $\mu_{\max}$, $\mu_{\min}$ and the heights are computed from the given Gram matrices. The matrices Y' and Y_i are arbitrary: the hypotheses (c) and (d) concern only the matrices $\widetilde{B}_i$, and the theorem holds for every choice of Y' and Y_i for which they are satisfied. A natural choice builds them from the Gram matrices at σ, so that $\widetilde{B}_i$ is the matrix of B_i

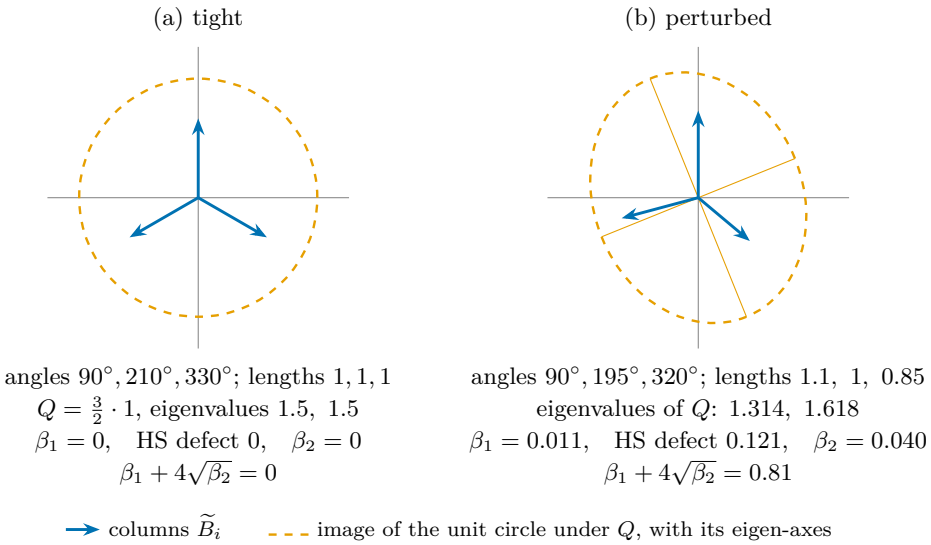


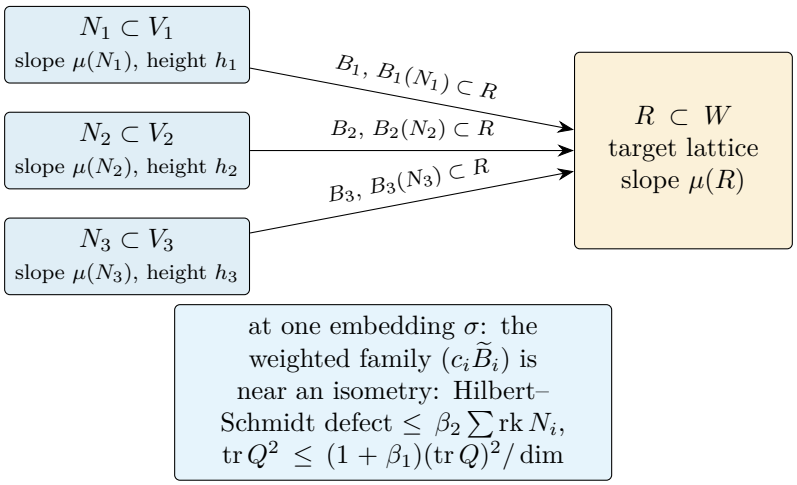
FIGURE 7.3. Two weighted families $(c_i, \widetilde{B}_i)_{i=1,2,3}$ as in Definition 7.4.1, with $J = \{1, 2\}$, no padding rows ($J_1 = \emptyset$, so $\beta_0 = 0$), $c_i = 1$, and each $\widetilde{B}_i$ a single column with real entries. The values shown are the smallest β_1 and β_2 allowed by hypothesis (d) of Theorem 7.4.2: $1 + \beta_1 = |J| \operatorname{tr}(Q^2)/(\operatorname{tr} Q)^2$, and β_2 is the Hilbert–Schmidt defect $\sum_i (\|\widetilde{B}_i\|^2 - 1)^2$ divided by $\sum_i \operatorname{rk} N_i = 3$. In (a) Q is scalar and every column has length 1. In (b) the eigenvalues of Q are close together, so β_1 is small, but $4\sqrt{\beta_2}$ is already 0.80: in this example the factor $\beta_1 + 4\sqrt{\beta_2}$ of Δ in the conclusion of the theorem comes mostly from the lengths of the columns.

in bases that are orthonormal at σ. The rows indexed by J_1 are padding rows, and β_0 bounds their proportion by (c). In the theorem s plays the role of the frame slope of Definition 7.2.2, u is the excess of the most unstable source over the mean, Δ is the spread of the mean over the target, and $\beta_0, \beta_1, \beta_2$ measure how far the normalised maps at σ are from a tight frame (Figure 7.3): β_0 the proportion of extra coordinates, β_1 the spread of the eigenvalues of Q, and β_2 the distance of each weighted $\widetilde{B}_i$ from an isometry.

IDEA OF PROOF. This is our account of the method. Alpöge's proof is not reproduced (Section 4.1). Theorem 7.2.5 is applied, as one step, to the Harder–Narasimhan pieces of the sources N_i, and bounds the slope of each piece by the slope of its image in R plus the height of B_i. A linear-algebra statement about near-tight frames then controls how much of W the images miss: if the weighted maps $c_i^{1/2} \widetilde{B}_i$ have Hilbert–Schmidt defect at most $\beta_2 \sum \operatorname{rk} N_i$ and the trace condition with β_1 holds, then the images

of subspaces of small codimension in the V_i span all of $\mathbb{C}^{J \sqcup J_1}$ except a proportion bounded by $\beta_1 + 4\sqrt{\beta_2}$ plus four times the codimension. In this route, a layer-cake summation over the slopes between s and $s+u$ produces the term $4u\log(1+\Delta/u)$, the missing proportion times the spread Δ produces the term $(\beta_1 + 4\sqrt{\beta_2})\Delta$, and the factor $(1-\beta_0)^{-1}$ comes from passing from $J \sqcup J_1$ to J. The term $\log|I|$ is the cost of averaging over the index set. □

Figure 7.4 shows the configuration.



conclusion:
$$s - \log|I| - (1-\beta_0)^{-1}\big(4u\log(1+\Delta/u) + (\beta_1 + 4\sqrt{\beta_2})\Delta\big) \leq \mu(R),$$
with s the rank-weighted mean of $\mu(N_i) - h_i$

FIGURE 7.4. The frame theorem (Theorem 7.4.2): source lattices N_i map integrally into one target lattice R. If the maps are jointly close to an isometry at one infinite place, the mean source slope s is at most the target slope plus explicit corrections. Bost's inequality (Theorem 7.2.5) compares one source with the maximal slope of the target. It is one step of the proof of the frame theorem (Remark 7.4.3).

REMARK 7.4.3 (the slope clause is not Bost's inequality). The slope input of the argument is neither Theorem 7.2.5 applied to each period map B_u (when B_u is injective), which bounds $\mu(N_u)$ by $\mu_{\max}(R^\vee) + h_u$, nor an average of these bounds. Clause (ix) of Definition 7.3.2 bounds the *average* s of $\mu(N_u) - h_u$ by the *slope* $\mu(R^\vee)$, not by the maximal slope, up to a correction controlled by the spread $s - \mu_{\min}(R^\vee) \leq \Delta_R$ and by the defects $\beta_0, \beta_1, \beta_2$. It has the form of the conclusion of Theorem 7.4.2, with target $R^\vee$, sources N_u and maps B_u. Bost's inequality is one step of the proof of that theorem.

Clause (ix) is a lower bound for $\mu(R^\vee)$. Clause (ii) is an upper bound for $\mu(R^\vee)+h(u)-\mathrm{Fo}$, so the two clauses together determine $\mu(R^\vee)$ up to the error terms. This is the cancellation of Remark 9.2.2. In the construction of tied families the spread is taken to be $\Delta_R = x^{C_H}b$, and clause (x) follows from the bound $s+\mu_{\max}(R)\le x^{C_H}b$ of Definition 7.5.1(2)(a), since $\mu_{\min}(R^\vee)=-\mu_{\max}(R)$.

7.5. The archimedean side

DEFINITION 7.5.1 (the upper link; schematic). Let $\mathcal{P}$ be a package (Definition 7.2.3) with ranks datum (k,m,t), period lattice data R, N_u, B_u, h_u and frame slope s, norm primes ϖ_0,ϖ_1 (Definition 5.3.2), top places $\mathcal{T}$ with top sizes a_τ and level primes $T(u)$, and put $\lg=\log(\varpi_0\varpi_1)$. Let a size datum for $\mathcal{P}$ be given (Definition 7.2.4), with the scale and size quantities $x,\varepsilon,a,b,\Upsilon_a,B_0,N_c$, the real threshold $X^\sharp$ and the field term Fo, and with the local sizes $\mathrm{Cg}_j(u)$ of the congruence modules of the words u at the lift indices j. Let Ω be the period with $\log\Omega=\gamma+\log(\pi/v)-L'(1,\chi_{-d})/L(1,\chi_{-d})$, where v is the half root discriminant of the host (Definition 5.1.2), and let C, C_{abs}, C_H be real numbers.

(1) The *archimedean main term* of a word u is

$$A(u)=\sum_{\tau\in\mathcal{T}}\Big(m\big(a_\tau\log a_\tau-a_\tau-a_\tau\log\Omega\big)-a_\tau\log T(u)\Big).$$

(2) The package *satisfies the upper link* with constants C, C_{abs}, C_H if the following hold.

(a) (Integrality.) $B_u(N_u)\subset R^\vee$ for every word u, R has full rank, and $s+\mu_{\max}(R)\le x^{C_H}b$; and the local sizes vanish above $X^\sharp$: at every place $\mathfrak{p}$ of residue characteristic $p>X^\sharp$ with $p\notin\{T(u),\varpi_0,\varpi_1,d\}$ and p prime to the orders of the stabilisers in the class sets, the gains and the congruence terms of u at $\mathfrak{p}$ are 0.

(b) (Bessel identity.) The periods are rational, and there are a unitary map V of the space of test forms and, for each word u, a unitary map U_u of the space of forms of u and a scalar $\alpha_u\neq 0$, such that $B_u(\phi)(F)=\alpha_u\, b_u(U_u\phi, VF)$, where b_u is the pairing of Bessel type carried by the package.

(c) (Heights.) $\mathrm{ht}(B_u)\le h_u$ whenever $B_u\neq 0$, for the lattice structures of N_u and $R^\vee$.

(d) (Divisor bound.) For every word u,

$$\mu(R^\vee)+h_u-\mathrm{Fo}\le A(u)+C|\mathcal{T}|a\Upsilon_a+Cm(B_0+N_c)x^{C_{\mathrm{abs}}\varepsilon}.$$

(e) (Congruences and heights.) For every word u and lift index j, $\big|\mathrm{Cg}_j(u)-\big(\sum_\tau a_\tau\big)\lg\big|\le CB_0$; and for all words u,u',

$$\Big|\Big(h_u+\sum_\tau a_\tau\log T(u)\Big)-\Big(h_{u'}+\sum_\tau a_\tau\log T(u')\Big)\Big|\le C\big(|\mathcal{T}|a\Upsilon_a+mB_0\big).$$

In words, the Bessel identity (2)(b) says that the periods agree, up to a nonzero scalar and unitary changes of variables, with a pairing of Bessel type. The pairings b_u are part of the schematic data of a package (Section 4.1). The divisor bound (2)(d) has the shape of clause (ii) of Definition 7.3.2, and (2)(e) has the shape of clauses (iii) and (iv). Three main terms occur and are kept apart. The main term of the balance identity for one weight is $\mathrm{MT}_w(a)$ (Definition 6.1.3). The main term of a record is $\mathrm{Main}(u_0)$ (Definition 7.3.3); under the identification clause it is the sum of the $\mathrm{Re}\,\mathrm{MT}$ over τ and j (Proposition 9.1.2). The archimedean side of the upper link is $A(u)$.

Remark 7.5.2 (Stirling and the CM period). Heuristically, the archimedean norm of a vector of weight a is governed by Stirling's formula [**NIS26**, (5.11.1)], which gives $a \log a - a$, normalised by the a-th power of a CM period Ω with

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

where v is the half root discriminant of the host field (Definition 5.1.2). This is a period of Chowla–Selberg or Kronecker-limit type for CM by K; compare [**SC67**], which gives the product formula for CM periods. The m lifts contribute m copies of this term. The quotient $L'(1, \chi_{-d})/L(1, \chi_{-d})$ enters the argument at this point.

Remark 7.5.3 (no CM heights). The argument does not use heights of CM points on a Shimura variety or a Colmez-type sum of $L'/L(1, \chi)$ over many characters [**Col93**]. Only $L'(1, \chi_{-d})/L(1, \chi_{-d})$ enters the main term. The other characters enter only through prime sums at inert primes weighted by w. No CM point is evaluated, and the slope inequality is applied to a period map. Arakelov geometry is used to bound the covolume of a lattice on the indefinite side (Theorem 8.2.3).

7.6. The finite side

Definition 7.6.1 (the lower link; schematic). Let $\mathcal{P}$ be a package with ranks datum (k, m, t), so with m lift indices, top places $\mathcal{T}$ and top sizes a_τ, and let $f = |\Phi|$ and $N(m) = k + m(t+1)$. Let a size datum for $\mathcal{P}$ be given (Definition 7.2.4), with scale x, total size a and real threshold $X^\sharp$. For each place $\mathfrak{p}$ of its support and each lift index j let $\mathrm{gain}_j(\mathfrak{p})$ and $\mathrm{ta}_\mathfrak{p}$ be the local sizes of the size datum. Let C be a constant of the link, put

$$\lambda_{\mathrm{fl}}(p) = Cf\, N(m)^2 \log_p a + Cf(\log_p a)^2$$

for a prime p, and let $\lambda_{\mathrm{lo}}(p)$ be a lower-order term of the same kind: C times size quantities of the datum times $\log_p(2a)$, plus, for inert $p < X^\sharp$, correction terms, each a multiple of $\log x$ by size quantities of the scale, whose exact form is omitted (Section 4.1).

(1) For a lift index j, $\tau \in \mathcal{T}$ and a prime p, the *Legendre sum* is
$$L_{j,\tau}(p) = \sum_{i \text{ odd}} \Big(\frac{a_\tau}{p^i} - w^{\mathrm{pt}}_{j,\tau} \Big)_+ ,$$
where $w^{\mathrm{pt}}_{j,\tau}$ is the partner weight of the package at level j (compare Definition 6.1.3(1)).

(2) The package satisfies the *Legendre-type gain condition* if at every place $\mathfrak{p}$ above a prime $p \le X^\sharp$ inert in K, and for every lift index j,
$$0 \le \mathrm{gain}_j(\mathfrak{p}) \le \frac{4|\mathcal{T}|a}{p},$$
$$\sum_j \sum_\tau \big(L_{j,\tau}(p) - \lambda_{\mathrm{lo}}(p) \big) - m\, \lambda_{\mathrm{fl}}(p) \le \sum_j \mathrm{gain}_j(\mathfrak{p}) - \mathrm{ta}_{\mathfrak{p}},$$
the inner sum over the top places attached to $\mathfrak{p}$.

(3) The package satisfies the *lower link* if it satisfies the Legendre-type gain condition and further local conditions: bounds for local lengths at places whose residue characteristic is not $T(u)$, and bounds for sums over orbits of places.

The further local conditions in (3) are part of the data and are not written out (Section 4.1). The gain condition is the finite side of the argument: each inert prime $p \le X^\sharp$ contributes the Legendre sum of a_τ/p^i over odd i, which mirrors the contribution of p to the prime sum $\mathrm{PS}_{\tau,w}$ of Definition 6.1.3.

Remark 7.6.2 (the balance of $n!$). Summing the gains over places gives approximately $\sum_{p \text{ inert}} (a/p) \log p + \cdots$, a prime sum over a Frobenius class. Set against Stirling's $a \log a - a$ (Remark 7.5.2), this is the balance between the archimedean and the finite sizes of $\log n!$ used by Chebyshev [**Tch52**]. The method is Chebyshev's factorial argument, twisted by the characters of $\mathrm{Gal}(E/\mathbb{Q})$ through CM periods and sharpened by a chain of $m \approx \sqrt{x}$ theta lifts. The chain is the multiplicity device. The role that vanishing of an auxiliary polynomial to high order plays in Stepanov's method [**Ste69**] is played here by p-adic divisibility of lifted periods at all inert p at once, forced by weights $a \ge p$.

CHAPTER 8

Families and the covolume bound

This chapter collects the packages of Chapter 7 into families (Definition 8.1.1), introduces the closed properties through which constants are fixed before a family is chosen (Definition 8.1.2), and states the covolume bound, Theorem 8.2.3, the single result through which arithmetic geometry enters the argument. The covolume bound is not proved here: its inputs are listed after the statement, and Remark 8.3.4 records the precise form in which each of them is used, including one input for which the implication from the cited theorem is not claimed. Chapter 3 uses Theorem 8.2.3 as a black box. In the proof of the main theorem it enters through the existence of admissible-family data, Theorem 12.3.4(c) of Chapter 12, and, summed over the lifts, it supplies the main term of the divisor bound of the upper link (Definition 7.5.1).

The packages of Chapter 7 are assembled into families, one package for each member at each scale. The covolume bound is a bound for the covolumes of lattices of integral forms on the indefinite unitary groups $U(W_j)$. The chapter ends with the forms in which the theorems of arithmetic geometry are used.

8.1. Families and closed properties

DEFINITION 8.1.1 (families; schematic). Let $\varepsilon > 0$.

(1) A *family at* ε, written $\mathfrak{F}$, consists of:
   (a) a threshold $x_1 > 1$ and an unbounded set $\mathrm{Win} \subset \mathbb{R}$;
   (b) for each real x a finite index set I_x, nonempty for $x \in \mathrm{Win}$, and for each $i \in I_x$ a *member*: a package $\mathcal{P}_{x,i}$ together with a size datum $Z_{x,i}$ for $\mathcal{P}_{x,i}$ at scale x and ε (Definition 7.2.4) when $x > x_1$, such that the third entry J of the scale (x, ε, J) of $Z_{x,i}$ is independent of x and i;
   (c) uniformity conditions: for each x, all members at x share their ranks (Definition 7.1.1), the primes q and d and the integer Q_S of their hosts (Definition 5.1.2), the set $\mathcal{T}_2$, the Artin data, the auxiliary primes ϖ_0, ϖ_1 and their sets of words, through fixed identifications; and for every member and every shift there is a member with that shift and the same top sizes;
   (d) parameter choices and tie data, and period lattice data (Definition 7.2.2) whose heights h_u are given by an explicit formula.

(2) A family $\mathfrak{F}$ *has windows of exponent* $\lambda > 1$ if Win has windows of exponent λ (Definition 4.4.2).

The local sizes of the size data, the remaining uniformity conditions in (1)(c), the parameter choices and the formula for the heights in (1)(d) are part of the data of a family. They are not written out here (Section 4.1). Later chapters use them through the ties of Definition 9.1.4 and the hypotheses of Definitions 8.2.1 and 12.3.2.

In plain terms, a property is *closed* if, for ε small, it holds for every family whose constants exceed thresholds chosen before the family. The next definition makes this precise.

Definition 8.1.2 (thresholds and closed properties). Let $\varepsilon > 0$. Each family $\mathfrak{F}$ at ε (Definition 8.1.1) carries finite lists of real constants $c_{\mathrm{pre}}(\mathfrak{F})$, $c_{\mathrm{out}}(\mathfrak{F})$, $c_{\mathrm{last}}(\mathfrak{F})$ and, for each real number c, a list $c_{\mathrm{mid}}(\mathfrak{F}, c)$ whose first entry is c.

(1) A *threshold* κ consists of a real number κ_{pre} and three functions κ_t, κ_{mid}, κ_{last} from finite lists of real numbers to real numbers.
(2) A family $\mathfrak{F}$ is *large for* (κ, c) if every entry of $c_{\mathrm{pre}}(\mathfrak{F})$ is at least κ_{pre}, every entry of $c_{\mathrm{mid}}(\mathfrak{F}, c)$ is at least $\kappa_{\mathrm{mid}}(c_{\mathrm{out}}(\mathfrak{F}))$, and every entry of $c_{\mathrm{last}}(\mathfrak{F})$ is at least κ_{last} of the list $c_{\mathrm{out}}(\mathfrak{F})$ followed by $c_{\mathrm{mid}}(\mathfrak{F}, c)$.
(3) Let $\mathcal{C}$ be a class of families and P a property of a family and a real number. Then P is *closed* on $\mathcal{C}$ if there is $\varepsilon_1 > 0$ such that for every $\varepsilon \in (0, \varepsilon_1]$ there is a threshold κ for which $P(\mathfrak{F}, c)$ holds for every family $\mathfrak{F}$ at ε in $\mathcal{C}$ and every real c such that $\mathfrak{F}$ is large for (κ, c).

The entries of the constant lists are constants of the family and are not listed here. In (3) the threshold is chosen *before* the family. Several closed properties can therefore be combined by combining their thresholds, and a single family that is large for the combined threshold is then supplied by the existence theorem of Chapter 12.

8.2. The covolume bound

Definition 8.2.1 (the covolume condition; schematic). Let $\mathcal{C}_{\mathrm{cov}}$ be the class of families whose packages carry the integral structures used below, let $\varepsilon > 0$, and let $\mathfrak{F}$ be a family at ε in $\mathcal{C}_{\mathrm{cov}}$ with members $(\mathcal{P}_{x,i}, Z_{x,i})$.

(1) A member $(\mathcal{P}, Z)$ satisfies the *member hypothesis* if the following two conditions hold.
  (a) (Field size.) Let $F_O \subset \mathbb{C}$ be the field generated over $\mathbb{Q}$ by the image of the coefficient ring of $\mathcal{P}$ under its fixed complex embedding, and put $P_O = (19fr^2 + 198f^2)\, x^{\varepsilon}$, where f and r are the derived quantities of the ranks of $\mathcal{P}$ (Definition 7.1.1) and (x, ε, J) is the scale of Z. Then $\log[F_O : \mathbb{Q}] \le P_O$ and $\log|\operatorname{disc} F_O|/[F_O : \mathbb{Q}] \le P_O$.
  (b) (Local types at the norm primes.) For $k = 0, 1$, write $G_k = \mathrm{GL}_r(\mathbb{F}_{\varpi_k})$. There is a labelling of the irreducible characters

of G_k by polynomials over $\mathbb{F}_{\varpi_k}$, of the kind used in Green's parametrization (Theorem 12.2.1) and built on the character of $\mathbb{F}_{\varpi_k}^{\times}$ that $\mathcal{P}$ fixes at ϖ_k, such that for every polynomial P the idempotent of $\mathbb{C}[G_k]$ that $\mathcal{P}$ attaches to P is the sum of the central idempotents of the irreducible characters with label P.

(2) A *point* of a member is a choice of the auxiliary integral data (lattices, integral structures and constants) at which the lattices M_j below are formed, and the *point conditions* are compatibility conditions on such a choice; both are schematic. A point $\mathfrak{z}$ carries a positive integer deg, the degree by which covolumes are normalised.

(3) Let $\mathfrak{z}$ be a point of a member $\mathcal{P}$, with weights $w(j,\rho)$ and period Ω, and let $C \geq 0$ and $x > 1$. For a lift index $1 \leq j \leq m$, let M_j be the lattice of integral forms on the indefinite unitary group $U(W_j)$ attached to $\mathfrak{z}$, and let $(p_j, q_j) = (k+j-1, m-j)$ be the signature of W_j (Definition 7.1.2). The point satisfies the *covolume inequality* with constant C at x if for every lift index j

$$\frac{\log \operatorname{covol}(M_j)}{\deg} \leq \sum_{\rho} (p_j - q_j)\, w(j,\rho) \log \Omega + C\, B_0 \log\log x,$$

where $\log \operatorname{covol}(M_j)$ is the normalised logarithmic covolume of M_j, deg is the degree of the point, the sum runs over $\rho \in \Phi$, and B_0 is the size quantity of the member (Definition 7.2.4).

(4) The *covolume condition* is the statement that the following property of a family $\mathfrak{F}$ and a real number c is closed on $\mathcal{C}_{\mathrm{cov}}$ (Definition 8.1.2): there is $C_0 \geq 0$ such that for all large x, every index i for which $(\mathcal{P}_{x,i}, Z_{x,i})$ satisfies the member hypothesis, every point $\mathfrak{z}$ of $\mathcal{P}_{x,i}$ that satisfies the point conditions, and every $C \geq C_0$, the point $\mathfrak{z}$ satisfies the covolume inequality with constant C at x.

The class $\mathcal{C}_{\mathrm{cov}}$ is fixed once and for all. It is distinct from the classes of admissible families of Chapter 12 (see the paragraph after Definition 12.3.2). The class $\mathcal{C}_{\mathrm{cov}}$ and the points of a member with their point conditions are schematic (Section 4.1). The period Ω is that of the slope record (Definition 7.3.1(6)), with $\gamma^* = \gamma$ and $\ell_K = L'(1,\chi_{-d})/L(1,\chi_{-d})$ (Remark 7.5.2). The lattices M_j lie on the indefinite side, one for each lift index. They are not the test lattice R of Definition 7.2.2.

Remark 8.2.2 (the member hypothesis is not a tie). Chapter 9 ties a family at a scale to a slope record (Definition 9.1.4). The member hypothesis of Definition 8.2.1(1) is not a condition on the slope record of a member. The member hypothesis is a bound on the size of the field F_O together with a condition on the local types of the package at the two norm primes. Both conditions involve only the package and its size datum. They involve no point of the member and no slope record. The covolume bound is independent of the analytic data of Chapters 6 and 9.

THEOREM 8.2.3 (Alpöge; covolume bound). *There is $\varepsilon_1 > 0$ such that for every $\varepsilon \in (0, \varepsilon_1]$ there is a threshold κ with the following property. Let $\mathfrak{F}$ be a family at ε in $\mathcal{C}_{\mathrm{cov}}$ and c a real number such that $\mathfrak{F}$ is large for (κ, c). Then there is $C_0 \geq 0$ such that for all large x, every member $(\mathcal{P}_{x,i}, Z_{x,i})$ satisfying the member hypothesis, every point of $\mathcal{P}_{x,i}$ satisfying the point conditions and every $C \geq C_0$, the point satisfies the covolume inequality with constant C at x (Definition 8.2.1(3)).*

Written out, the conclusion is that for every lift index j

$$\frac{\log \mathrm{covol}(M_j)}{\deg} \leq \sum_{\rho \in \Phi} (p_j - q_j)\, w(j,\rho) \log \Omega + C\, B_0 \log\log x,$$

where M_j, (p_j, q_j), deg, $w(j,\rho)$, Ω and B_0 are the lattice, signature, degree, weights, period and size quantity attached to the point and the member by Definition 8.2.1(3). So Theorem 8.2.3 says that the covolume condition holds (Figure 8.1 shows the order of its quantifiers).

Informal statement. For the lattice M_j of integral cusp forms of the weights $w(j,\rho)$ on the Shimura variety of $U(W_j)$, at the level prescribed by the package (Theorem 12.3.4(b)), the normalised logarithmic covolume is at most $\sum_{\rho\in\Phi}(p_j - q_j)\, w(j,\rho) \log \Omega$, up to an error $O(B_0 \log\log x)$, for all large x and uniformly over the members of a family whose field of coefficients is not too large.

ON THE PROOF. We do not reproduce the proof. The bound rests on the following inputs, all of which are cited theorems of arithmetic geometry in the forms described in Remark 8.3.4 and Figure 8.3.

(1) Toroidal compactifications of the unitary Shimura varieties of the groups $U(W_j)$ and of Siegel modular varieties [**FC90, Lan13**], and resolution of singularities in characteristic zero [**Hir64**].
(2) The arithmetic Hilbert–Samuel theorem [**GS92, AB95**] for a smooth metric, applied once, as an upper bound for a counting volume on the projective image of a map given by cusp forms with a Fubini–Study metric. Also Moriwaki's bigness theorem [**Mor09**].
(3) A comparison of the Fubini–Study metric with the Petersson metric through the Bergman kernel and the comparison of sup and L^2 norms of cusp forms. Geometric degrees from a Chern–Weil statement for good metrics of Mumford type [**Mum77**].
(4) Kawamata–Viehweg vanishing [**Kaw82, Vie82**] and Nadel vanishing [**Nad90**].
(5) Integral models of the Shimura varieties, through Grothendieck–Messing deformation theory [**Mes72**].

□

Figure 8.2 shows the coefficient of the main term along the tower, and how the bound feeds the divisor bound of a slope record.

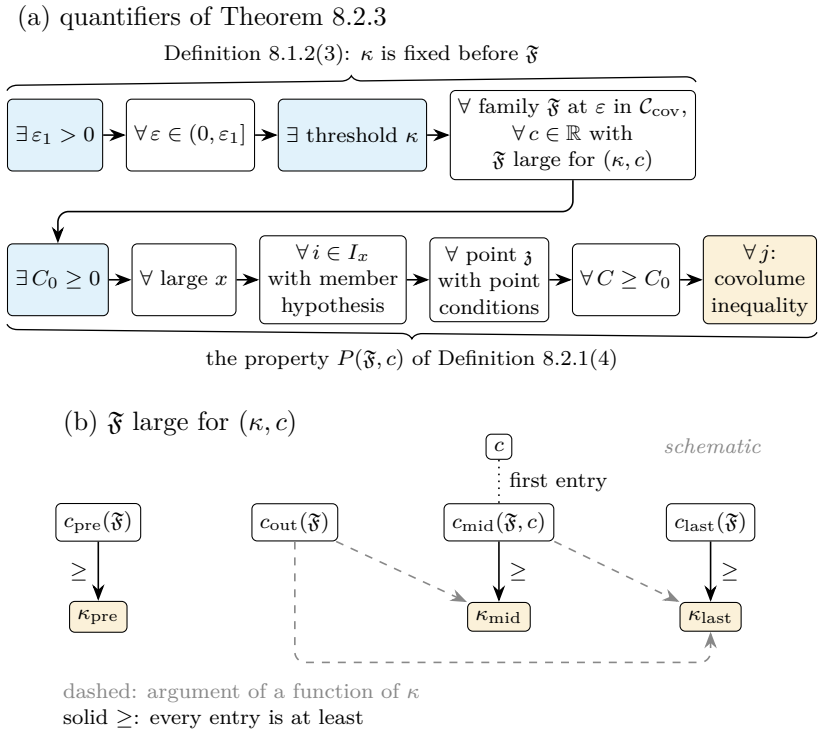


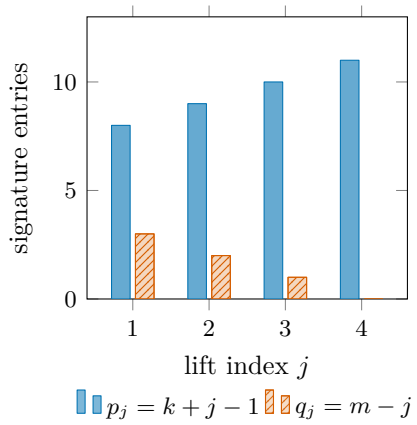
FIGURE 8.1. Schematic. (a) The order in which Theorem 8.2.3 fixes its objects; shaded boxes are existential, white boxes universal, the last box is the conclusion, and each object may depend on every object before it in the chain. The threshold κ is chosen before the family, while C_0 and the range of large x may depend on $\mathfrak{F}$ and c. (b) The comparisons in "$\mathfrak{F}$ is large for (κ, c)" (Definition 8.1.2). The list $c_{\mathrm{out}}(\mathfrak{F})$ is not bounded by anything; it only enters as an argument of κ_{mid} and, followed by $c_{\mathrm{mid}}(\mathfrak{F}, c)$, of κ_{last}. The function κ_t of a threshold does not enter this condition and is not drawn.

Summed over j, the main terms of Theorem 8.2.3 give the term $-ma_\tau \log \Omega$ of the divisor bound of the upper link (Definition 7.5.1(2)(d)): the lattice of rows is built from the M_j.

8.3. How arithmetic geometry enters

REMARK 8.3.1 (one output of the arithmetic geometry). The theorems of arithmetic geometry used by the argument (resolution of singularities, arithmetic Hilbert–Samuel, Moriwaki's bigness theorem, toroidal compactifications, Kawamata–Viehweg and Nadel vanishing, Grothendieck–Messing theory) do not enter the argument at separate steps. All of them enter

(a) W_j: (p_j, q_j), small example $k = 8$, $m = 4$



(b) the inputs of Theorem 8.2.3

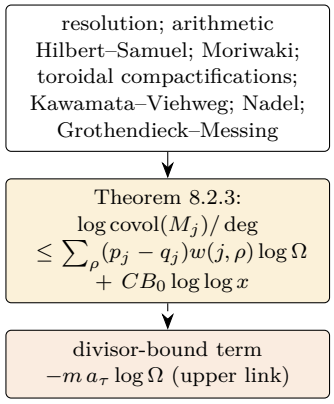


FIGURE 8.2. (a) The signature $(p_j, q_j) = (k + j - 1, m - j)$ of the space W_j for $k = 8$ and $m = 4$; t plays no role. Since $k \geq m + 3$, the coefficient $p_j - q_j = k - 1 - m + 2j$ of the main term of Theorem 8.2.3 is positive for every j and grows with j. (b) The theorems of arithmetic geometry feed only the covolume bound (Remark 8.3.1). The dashed arrow: the main term $\log \Omega$ of the covolume bound corresponds to the term $-ma_\tau \log \Omega$ of the divisor bound, clause (ii) of Definition 7.3.2.

through the one covolume bound, Theorem 8.2.3. Nothing in the argument bounds heights of CM points (Remark 7.5.3).

REMARK 8.3.2 (smooth metrics suffice). The Hodge bundle with its Petersson metric has logarithmic singularities along the boundary of a toroidal compactification [**BGKK05**]. The argument does not need an arithmetic Hilbert–Samuel theorem for such log-singular metrics. The Hilbert–Samuel statements used require smooth metrics, and four devices avoid the singularity. First, only cusp forms with L^2 Petersson norms enter. Second, cusp forms are identified with sections of an adjoint-type bundle on a smooth projective resolution. Third, for the height bound the Petersson metric is replaced by a Fubini–Study metric on a projective image, with Bergman-kernel and sup-versus-L^2 comparisons. Fourth, geometric degrees come from a Chern–Weil statement for good metrics of Mumford type [**Mum77**]. The comparisons of the second and third devices cost at most the unit error $C\, B_0 \log\log x$ of each lift in Theorem 8.2.3.

REMARK 8.3.3 (the heights need no further geometry). The heights h_u of the period lattice data (Definition 7.2.2) need no arithmetic geometry beyond the covolume bound. The height used depends only on a number

input	sketch	restriction in the form used	implied by the cited theorem?
(1) near-general			
(a) resolution of singularities	(1)	characteristic zero	yes
(b) arithmetic Hilbert–Samuel	(2)	nef $\overline{L}$ with smooth metric, self-intersection a as hypothesis	via Moriwaki for nef $\overline{L}$; Gillet–Soulé and Abbes–Bouche treat ample
(2) special or weakened			
(a) Moriwaki's bigness theorem	(2)	intersection pairing specified by axioms	**not claimed**: the pairing is not checked to be Moriwaki's arithmetic one
(b) toroidal compactifications	(1)	unitary of signature (p, q); separately, integrally, Siegel of level $n \geq 3$ odd over $\mathbb{Z}[\zeta_n, 1/n]$	yes
(c) Kawamata–Viehweg	(4)	nef and big; no $\mathbb{Q}$-divisors	yes
(d) Nadel	(4)	*some* ideal sheaf, not identified with the multiplier ideal	yes
(e) Grothendieck–Messing	(5)	nilpotent divided powers, p nilpotent	yes

FIGURE 8.3. The arithmetic-geometry inputs of Theorem 8.2.3, labelled as in Remark 8.3.4. "Sketch" is the item of the list after Theorem 8.2.3 that names the input; the inputs of its item (3), the metric comparisons and the Chern–Weil statement, are not listed in Remark 8.3.4. The last column records the two qualifications at the end of that remark: for (2)(a) the implication is not claimed.

computed from the rows of the package. All the geometric content of the argument is in the covolume bound, Theorem 8.2.3.

REMARK 8.3.4 (the forms of the arithmetic-geometry inputs). Each input of Theorem 8.2.3 is used in a specific formulation, which need not have the generality of the theorem it is named after. The formulations range from near-general to special or weakened.

(1) Near-general.
   (a) Resolution of singularities in characteristic zero [**Hir64**]: a quasi-compact, separated, reduced scheme locally of finite type over a field of characteristic zero admits a sequence of blow-ups to a smooth scheme that is an isomorphism over the smooth locus and whose exceptional locus has simple normal crossings.

(b) Arithmetic Hilbert–Samuel [**GS92, AB95**]: for a nef hermitian line bundle $\overline{L}$ with smooth metric on an arithmetic variety of dimension $D+1$, whose self-intersection number a enters as a hypothesis, and any hermitian line bundle $\overline{N}$, the logarithm of the number of small sections of $\overline{L}^{\otimes n} \otimes \overline{N}$ is $a\, n^{D+1}/(D+1)! + o(n^{D+1})$.

(2) Special or weakened.

(a) Moriwaki's bigness theorem [**Mor09**]: a lower bound for the arithmetic volume by the self-intersection number, stated on multi-projective coordinate rings, for an intersection pairing that is specified by axioms.

(b) Toroidal compactifications [**FC90, Lan13**]: for unitary Shimura varieties of signature (p, q) attached to a CM field, a smooth model with simple normal crossings boundary and a logarithmic Kodaira–Spencer isomorphism; and, in a separate integral statement, for Siegel modular varieties of level $n \geq 3$ odd over $\mathbb{Z}[\zeta_n, 1/n]$, a smooth projective compactification with a semi-abelian extension of the universal abelian scheme and an ample modification of a power of the Hodge bundle.

(c) Kawamata–Viehweg vanishing [**Kaw82, Vie82**] for a nef and big line bundle on a smooth projective variety over an algebraically closed field of characteristic zero, without $\mathbb{Q}$-divisors.

(d) Nadel vanishing [**Nad90**] in the form: there is *some* ideal sheaf whose cosupport is the non-klt locus and for which the twisted higher cohomology vanishes. The ideal sheaf is not identified with the multiplier ideal.

(e) Grothendieck–Messing theory [**Mes72**] over a divided-power thickening with nilpotent divided powers, along which p is nilpotent.

Each formulation is implied by the cited theorem, with two qualifications. For (1)(b) with a nef, rather than ample, line bundle the asymptotic formula is due to Moriwaki [**Mor09**], as a consequence of the continuity of arithmetic volumes; Gillet–Soulé and Abbes–Bouche treat ample bundles with positive metrics. For (2)(a) the implication needs the axiomatic intersection pairing to be the arithmetic intersection pairing of [**Mor09**]. This is not checked here, so for (2)(a) the implication is not claimed.

CHAPTER 9

The Chebyshev inequality on windows

This chapter proves the Chebyshev inequality on windows, Theorem 9.4.1: for a family that satisfies the uniform explicit-formula condition and is tied at every large scale of its window set, $G_{\mathscr{L}}(x) \le x^{1/2+(e_D+e_L+7)\varepsilon}$ on the windows. The proof combines the slope records of Chapter 7 with the balance identity and the core bound of Chapter 6, through the gap bound, Theorem 9.2.1. The proof of the gap bound is given in outline: its steps are written out, but its last step, an estimation of error terms by linear arithmetic in which the largeness of x absorbs the constants, is not. In Chapter 3 the Chebyshev inequality on windows is a black box. It is applied in the last step of the proof of the main theorem, inside Theorem 12.4.1 of Chapter 12, whose output goes into the windowed Landau criterion for levels (Theorem 4.4.3, Chapter 4).

A slope record (Definitions 7.3.1 and 7.3.2) is a list of real numbers subject to eleven inequalities. The prime sum and the gap of Chapter 6 are attached to arithmetic data. Two hypotheses connect them: an identification clause, which says that the record's prime sum, top sizes and weights are those of the arithmetic data (Definition 9.1.1), and the ties of a family (Definition 9.1.4), which say that the record's numbers are the slopes, heights and local sizes of actual packages. Under these hypotheses the eleven clauses give an upper bound for the gap (Theorem 9.2.1), the core bound of Chapter 6 turns it into an upper bound for $G_{\mathscr{L}}$, and one obtains

$$G_{\mathscr{L}}(x) \le x^{1/2+(e_D+e_L+7)\varepsilon}$$

for all large x in the window set of the family (Theorem 9.4.1).

In the computation the slope of the test lattice cancels (Remark 9.2.2). The exponent $1/2$ comes from the bookkeeping of the error bracket, in which the number m of theta lifts is balanced against x/m (Proposition 9.3.1).

9.1. Window hypotheses and ties

Definition 9.1.1 (window and identification hypotheses). Let H be a finite abelian group of order n and $\mathfrak{G} = V_4 \times H$. Let a slope record with shifts $U_0 = H$ and positive integers m and f be given (Definition 7.3.1), together with size constants (Definition 7.3.2) and real numbers c_W, e_K and C_L. In (1) the letters $x, \varepsilon, a, b, b', k, \Upsilon, \Upsilon_a, B_0, N_c, \bar{\beta}_0, \bar{\beta}_1, \bar{\beta}_2$ and n_T denote the quantities of the record, and the c's and e's with subscripts denote its size constants. Let D be arithmetic data for $\mathfrak{G}$, ℓ a function from the

characters of $\mathfrak{G}$ to $\mathbb{C}$, $N \geq 1$ an integer, $\mathcal{T}, \mathcal{T}_2 \subset H$ finite sets, $a\colon H \to \mathbb{R}$ and $w\colon H \times \{0, \dots, m-1\} \times \mathfrak{G} \to \mathbb{R}$.

(1) The *window conditions* hold if
   (a) $1 \leq x$, $0 < \varepsilon \leq 1$, $1 \leq b$, $xb \leq a$ and $1 \leq n_T$;
   (b) $m \leq x^{1/2+\varepsilon/2}$ and $xm \leq k$;
   (c) $0 \leq e_K$, $1 \leq k \leq x^{e_K}$ and $k \leq a$;
   (d) $0 \leq \Upsilon \leq 2x^{\varepsilon}$, $0 \leq \Upsilon_a \leq 2x^{\varepsilon}$, $0 \leq b'$, $kb' \leq a$ and $0 \leq X^{\sharp} \leq a$;
   (e) $0 \leq a_\tau \leq 2a$ for every top place τ of the record, and $B_0, N_c \geq 0$;
   (f) $\log|\mathcal{U}(u_0)| \leq (k-3)\log 2$ for every $u_0 \in U_0$;
   (g) $\bar{\beta}_0 \leq 1/2$, $\bar{\beta}_1 \geq 0$ and $\bar{\beta}_2 \geq 0$;
   (h) $c_{\mathrm{div}}, c_{\mathrm{cg}}, c_{\mathrm{ht}}, c_{\mathrm{ct}}, c_{\mathrm{fs}}, c_{\mathrm{qs}}, c_{\mathrm{loss}} \in [0, c_W]$ and $e_{\mathrm{div}}, e_{\mathrm{loss}}, c_{\mathrm{sp}} \geq 0$;
   (i) (*defect smallness*) $\big(\bar{\beta}_1 + 4\sqrt{\bar{\beta}_2}\big)\, x^{c_{\mathrm{sp}}} b \leq c_W\, x^{\varepsilon}\, n_T\, a$.

(2) The *identification clause* holds if there are bijections e_T from the top places of the record onto $\mathcal{T}$ and $e_P\colon \{0, \dots, f-1\} \to \Phi$ such that the record's top size at τ is $a_{e_T(\tau)}$ and its weight at (u_0, j, ρ) is $w(u_0, j, e_P(\rho))$; $\gamma^* = \gamma$; $\ell_K = \operatorname{Re}\ell(\chi_K)$; and for every $u_0 \in H$

$$\mathrm{PS}(u_0) = \sum_{\tau \in \mathcal{T}} \sum_{j<m} \mathrm{PS}_{\tau, w(u_0,j)}(a_\tau).$$

(3) The *bounds on B_0 and N_c* hold if

$$B_0 \leq 2nmbx^{\varepsilon} \qquad \text{and} \qquad N_c \leq \frac{|\mathcal{T}_2|\, kbx^{\varepsilon}}{n^2 x^{-\varepsilon}}.$$

(4) The *values at x* are bounded if $|G_\theta^{\mathrm{ar}}(x)| \leq x^{1/2+\varepsilon}$ and $|\ell(\theta)| \leq C_L \log x$ for $\theta \in \{\omega, \chi_K, \chi_K\omega\}$.

The defect smallness (1)(i) is where the size of the defects in clause (ix) of Definition 7.3.2 is fixed; in the families of Chapter 12, $\bar{\beta}_1 = 1/x$.

PROPOSITION 9.1.2 (the main terms). *Let H be a finite abelian group and $\mathfrak{G} = V_4 \times H$. Let a slope record with shifts $U_0 = H$ and positive integers m and f be given (Definition 7.3.1). Let D be arithmetic data for $\mathfrak{G}$, ℓ a function from the characters of $\mathfrak{G}$ to $\mathbb{C}$, $N \geq 1$ an integer, $\mathcal{T}, \mathcal{T}_2 \subset H$ finite sets, $a\colon H \to \mathbb{R}$ with $a_\tau > 0$ for $\tau \in \mathcal{T}$, and $w\colon H \times \{0, \dots, m-1\} \times \mathfrak{G} \to \mathbb{R}$ with $w(u_0, j, \rho) > 0$ for all $u_0 \in H$, $j < m$ and $\rho \in \Phi$. Suppose that the identification clause (Definition 9.1.1) holds, and let $u_0 \in H$. Then the main term* Main *(Definition 7.3.3), the term* MT *(Definition 6.1.3) and the gap* gap *(Definition 6.2.1) satisfy:*

(a) $\mathrm{Main}(u_0) = \sum_{\tau \in \mathcal{T}} \sum_{j<m} \operatorname{Re}\mathrm{MT}_{w(u_0,j)}(a_\tau)$;
(b) $\mathrm{gap}(u_0) = \mathrm{PS}(u_0) - \mathrm{Main}(u_0)$.

PROOF. *Proof of (a).* Since a_τ and the weights are real and positive,

$$\operatorname{Re}\mathrm{MT}_w(a) = a\Big(\log a - 1 - \gamma + \operatorname{Re}\ell(\chi_K) - \frac{1}{f}\sum_{\rho \in \Phi} \log w(\rho)\Big),$$

where $f = |\Phi|$; the bijection e_P shows that this f is the record's f. By the identification clause, $\gamma^* = \gamma$, $\ell_K = \operatorname{Re}\ell(\chi_K)$, and e_T and e_P carry the record's top sizes and weights to a_τ and $w(u_0, j, \rho)$. Summing over τ and j and comparing with Definition 7.3.3 gives the identity.
Proof of (b). This follows from the definition of $\mathrm{gap}(u_0)$, part (a) and the clause of the identification for $\mathrm{PS}(u_0)$. □

Definition 9.1.3 (uniform explicit-formula condition). Let $\mathscr{L}$ be a level, $\varepsilon > 0$ and $\mathfrak{F}$ a family at ε (Definition 8.1.1). The family satisfies the *uniform explicit-formula condition* for $\mathscr{L}$ if there are real numbers C and x_0 such that for every $x \geq x_0$ and every $i \in I_x$ the host of the package $\mathcal{P}_{x,i}$, with its top places $\mathcal{T}$, satisfies $\mathrm{EF}(\mathscr{L}, \mathcal{T}, x, \varepsilon, C)$ (Definition 5.3.3).

Definition 9.1.4 (ties). Let $\varepsilon > 0$, let $\mathfrak{F}$ be a family at ε (Definition 8.1.1), and let $e_D, e_L, c_W, c_S, e_K, C_L$ and x be real numbers. The family is *tied at* x with these constants if there are

- an index $i_0 \in I_x$, with A the group of the frame of the package $\mathcal{P}_{x,i_0}$ and $\mathfrak{G} = V_4 \times A$;
- positive integers m and f, and a map $\mathrm{mem}\colon A \to I_x$, the *choice of members*;
- a slope record with shifts $U_0 = A$ whose words $\mathcal{U}(u_0)$ are the words of the package $\mathcal{P}_{x,\mathrm{mem}(u_0)}$, and size constants for it;
- a function ℓ from the characters of $\mathfrak{G}$ to $\mathbb{C}$, an integer $N \geq 1$, finite sets $\mathcal{T}, \mathcal{T}_2 \subset A$, maps $a\colon A \to \mathbb{R}$ and $w\colon A \times \{0, \ldots, m-1\} \times \mathfrak{G} \to \mathbb{R}$, and real numbers $\eta, X^\flat, \vartheta$,

such that the following hold for every $u_0 \in A$.

(1) (Members.) The member $\mathcal{P}_{x,\mathrm{mem}(u_0)}$ has the frame, the integers m and f, and the primes q and d of $\mathcal{P}_{x,i_0}$; its host has shift u_0; and the sizes that its size datum assigns to a top place do not depend on u_0.
(2) (Global ties.) The record's a, b, b' and Υ are those of the scale of the size datum $Z_{x,\mathrm{mem}(u_0)}$; k is the rank k of the package (Definition 7.1.1); Υ_a, B_0, N_c and $X^\sharp$ are those of the size datum, and so are η and $X^\flat$; $\gamma^* = \gamma$; ℓ_K is the logarithmic derivative at 1 recorded by the size datum; v is the half root discriminant of the host (Definition 5.1.2); and $\lg = \log(\varpi_0\varpi_1)$, with ϖ_0, ϖ_1 the norm primes of the host (Definition 5.3.2).
(3) (Local ties.) For every word $u \in \mathcal{U}(u_0)$: $\mathrm{rk}(u)$ is the rank of the lattice N_u of the period lattice data (Definition 7.2.2); $T(u)$ is the level prime of u; $\mu_N(u)$ and $\mu_N^{\max}(u)$ are the slope and the maximal slope of N_u (Definition 7.2.1); $h(u) = h_u$; $\mu_M(u)$ and $\mu_M^{\max}(u)$ are the slope and the maximal slope of the form lattice of u; and

$$\mathrm{Cg}_j(u) = \sum_{\mathfrak{p}} \mathrm{sz}_{\mathfrak{p}}\big(\text{congruence invariant of } u \text{ at lift index } j+1\big),$$

the sum over the places $\mathfrak{p}$ of the support of the size datum, with $\mathrm{sz}_{\mathfrak{p}}$ its local sizes. For the shift itself, $\mu_R(u_0)$ and $\mu_R^{\min}(u_0)$ are the slope and the minimal slope of $R^\vee$; $\mathrm{Fo}(u_0)$ and $\mathrm{Ga}(u_0)$ are those of the size datum; $\mathrm{IO}(u_0)$ is a sum over the support of local quantities of the size datum; and $\beta_0(u_0)$ is the rank defect of the member.

(4) (Arithmetic ties.) D are the arithmetic data built from the Artin map of the host of $\mathcal{P}_{x,i_0}$; $\mathcal{T}$, $\mathcal{T}_2$ and a are those of the members; $w(u_0, j, \cdot)$ are the weights of $\mathcal{P}_{x,\mathrm{mem}(u_0)}$ at lift index $j+1$; and $\ell(\theta) = L'(1,\chi)/L(1,\chi)$ whenever $\theta \neq \theta_0$ and χ is a primitive Dirichlet character with $\chi(n) = \theta_D(n)$ for every $n \geq 1$.

(5) (Constants.) The record's x and ε are x and ε, $e_{\mathrm{div}} \leq e_D$, $e_{\mathrm{loss}} \leq e_L$ and $c_{\mathrm{sp}} \leq c_S$.

(6) (Hypotheses.) The record satisfies the size data (Definition 7.3.2); the window conditions with c_W and e_K, the identification clause and the bounds on B_0 and N_c (Definition 9.1.1); the size conditions of Definition 6.2.1 for every $u_0 \in A$, and the thin-scale bound; the bound on the values at x with C_L; the cost conditions with $f = |\Phi|$; and the rest-size conditions with $n = |A|$, $n_T = |\mathcal{T}|$, $n_{T_2} = |\mathcal{T}_2|$ and $a_E = a$ (Definition 6.2.3).

The record's prime sum $\mathrm{PS}(u_0)$ is fixed by the identification clause in (6). The spread Δ_R and the defects β_1, β_2 are fixed by no tie; they are constrained only by the hypotheses collected in (6). The definition names quantities of packages and of size data (shifts, level primes, form lattices, congruence invariants, local sizes, the rank defect) that belong to the data of a package (Definition 7.2.3) and of a family (Definition 8.1.1); the definitions there are schematic (Section 4.1). Through the ties, the window conditions and the rest-size conditions, a tied family has $x^{1/2-\varepsilon} \leq m \leq x^{1/2+\varepsilon/2}$ and $xm \leq k$ at every scale at which it is tied.

9.2. The gap bound

THEOREM 9.2.1 (Alpöge; gap bound). *Let H be a finite abelian group of order n and $\mathfrak{G} = V_4 \times H$. Let a slope record with shifts $U_0 = H$, positive integers m and f and size constants be given (Definitions 7.3.1 and 7.3.2); write x, ε, a, b and k for the quantities of the record, and c_{sp}, e_{div} and e_{loss} for its size constants. Let c_W and e_K be real numbers. Let D be arithmetic data for $\mathfrak{G}$, ℓ a function from the characters of $\mathfrak{G}$ to $\mathbb{C}$, $N \geq 1$ an integer, $\mathcal{T}, \mathcal{T}_2 \subset H$ finite sets, $a\colon H \to \mathbb{R}$ and $w\colon H \times \{0, \dots, m-1\} \times \mathfrak{G} \to \mathbb{R}$. Suppose that the record satisfies the size data (Definition 7.3.2), that the identification clause, the window conditions with c_W and e_K and the bounds on B_0 and N_c (Definition 9.1.1) hold, and that*

$$\log x \leq x^{\varepsilon} \qquad \text{and} \qquad 200(1+c_W)^2(1+c_{\mathrm{sp}})(1+e_K) \leq x^{\varepsilon}.$$

Then for every $u_0 \in H$

$$\mathrm{gap}(u_0) \leq x^{(e_{\mathrm{div}}+e_{\mathrm{loss}}+3)\varepsilon}\,\mathrm{Err},$$

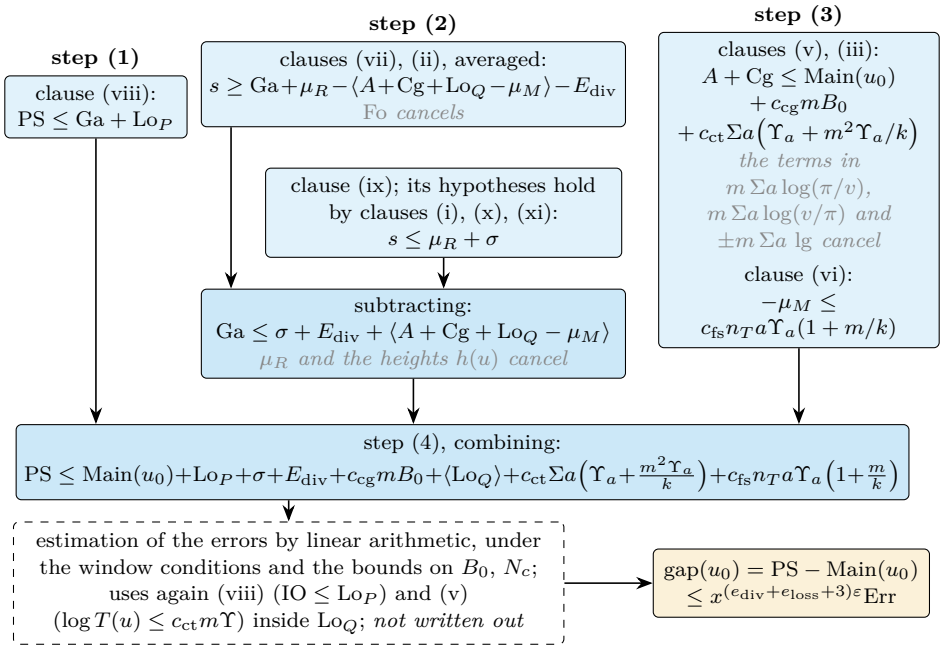


FIGURE 9.1. The outline proof of the gap bound, Theorem 9.2.1 (schematic). Roman numerals are the clauses of Definition 7.3.2; the steps are those of the proof. In step (2) the slope μ_R of the test lattice and the heights cancel (Remark 9.2.2, Figure 9.2); in step (3) what remains of the archimedean sum and of the weights in clause (v) is $\mathrm{Main}(u_0)$. Clauses (v) and (viii) are used twice. The last estimation is not written out.

with Err *the error bracket of Definition 6.2.3 for* $n = |H|$, $n_T = |\mathcal{T}|$, $n_{T_2} = |\mathcal{T}_2|$, *the given* b, k *and* m, *and* $a_E = a$.

PROOF (OUTLINE). By Proposition 9.1.2(b), $\mathrm{gap}(u_0) = \mathrm{PS}(u_0) - \mathrm{Main}(u_0)$, so it suffices to bound $\mathrm{PS}(u_0)$ by $\mathrm{Main}(u_0)$ plus errors (Figure 9.1). Fix u_0; all record quantities below are taken at u_0, and for a function X on the words we write

$$\langle X \rangle = \frac{\sum_{u \in \mathcal{U}(u_0)} \mathrm{rk}(u) X(u)}{\sum_{u \in \mathcal{U}(u_0)} \mathrm{rk}(u)},$$

so that $s = \langle \mu_N - h \rangle$. Put

$$A(u) = \sum_{\tau \in \mathcal{T}} \Big(m(a_\tau \log a_\tau - a_\tau - a_\tau \log \Omega) - a_\tau \log T(u) \Big),$$
$$E_{\mathrm{div}} = c_{\mathrm{div}} n_T a \Upsilon_a + c_{\mathrm{div}} m (B_0 + N_c) x^{e_{\mathrm{div}} \varepsilon},$$

and let $\sigma = \log |\mathcal{U}(u_0)| + (1-\beta_0)^{-1}\big(4u_F \log(1 + \Delta_R/u_F) + (\beta_1 + 4\sqrt{\beta_2})\Delta_R\big)$ be the slack of clause (ix). The clauses below are those of Definition 7.3.2.

(1) Clause (viii) gives $\mathrm{PS} \le \mathrm{Ga} + \mathrm{Lo}_P$.
(2) The second inequality of clause (vii) and clause (ii) give, for every word u, $\mu_N(u) - h(u) \ge \mathrm{Ga} + \mu_R + \mu_M(u) - \mathrm{Cg}(u) - \mathrm{Lo}_Q(u) - A(u) - E_{\mathrm{div}}$; the term Fo cancels. Averaging,

$$s \ge \mathrm{Ga} + \mu_R - \langle A + \mathrm{Cg} + \mathrm{Lo}_Q - \mu_M \rangle - E_{\mathrm{div}}. \tag{9.2.1}$$

The hypotheses of clause (ix) hold by clauses (i), (x) and (xi), so

$$s \le \mu_R + \sigma. \tag{9.2.2}$$

Subtracting, $\mathrm{Ga} \le \sigma + E_{\mathrm{div}} + \langle A + \mathrm{Cg} + \mathrm{Lo}_Q - \mu_M \rangle$, an inequality in which neither μ_R nor the heights $h(u)$ appear (Remark 9.2.2).
(3) With $\log \Omega = \gamma^* + \log(\pi/v) - \ell_K$, the lower bound for $\log T(u)$ in clause (v), multiplied by Σa, and the bound $\mathrm{Cg}(u) \le m\,\Sigma a\,\lg + c_{\mathrm{cg}} m B_0$ of clause (iii) give

$$A(u) + \mathrm{Cg}(u) \le \mathrm{Main}(u_0) + c_{\mathrm{cg}} m B_0 + c_{\mathrm{ct}}\,\Sigma a\Big(\Upsilon_a + \frac{m^2 \Upsilon_a}{k}\Big) :$$

the terms $m\,\Sigma a \log(\pi/v)$ and $m\,\Sigma a \log(v/\pi)$ cancel, the terms $\pm m\,\Sigma a\,\lg$ cancel, and what remains of the archimedean sum and of the weights in clause (v) is $\mathrm{Main}(u_0)$. Clause (vi) bounds $-\mu_M(u)$ by $c_{\mathrm{fs}} n_T a \Upsilon_a (1 + m/k)$.
(4) Combining (1)–(3),

$$\begin{aligned} \mathrm{PS} \le\ & \mathrm{Main}(u_0) + \mathrm{Lo}_P + \sigma + E_{\mathrm{div}} + c_{\mathrm{cg}} m B_0 + \langle \mathrm{Lo}_Q \rangle \\ & + c_{\mathrm{ct}} \Sigma a\Big(\Upsilon_a + \frac{m^2 \Upsilon_a}{k}\Big) + c_{\mathrm{fs}} n_T a \Upsilon_a \Big(1 + \frac{m}{k}\Big). \end{aligned}$$

The error terms on the right are then estimated under the window conditions and the bounds on B_0 and N_c (with $\mathrm{IO} \le \mathrm{Lo}_P$ from clause (viii) and $\log T(u) \le c_{\mathrm{ct}} m \Upsilon$ from clause (v) inside Lo_Q). This estimation, linear arithmetic in which the largeness of x absorbs the constants, gives the bracket $x^{(e_{\mathrm{div}} + e_{\mathrm{loss}} + 3)\varepsilon}\mathrm{Err}$.

□

Remark 9.2.2 (the test lattice cancels). The slope $\mu(R^\vee)$ of the test lattice, and with it its covolume, does not appear in the final inequality, although the periods live in the lattice $R^\vee$. The two lattice inequalities are the bounds (9.2.1) and (9.2.2) for the frame slope s:

$$\mathrm{Ga} + \mu_R - \langle A + \mathrm{Cg} + \mathrm{Lo}_Q - \mu_M \rangle - E_{\mathrm{div}} \ \le\ s \ \le\ \mu_R + \sigma.$$

The left bound comes from the divisor bound (clause (ii)) and the quotient slopes (clause (vii)), the right one from the slope inequality (clause (ix)). Both are bounds on the one number $\mu_R = \mu(R^\vee)$ in terms of s, and subtracting them removes it (Figure 9.2). The heights $h(u)$ cancel as well, since they enter s and clause (ii) with opposite signs. What survives is an inequality between the arithmetic quantity Ga, which bounds the prime sum, and the archimedean main term; one may read it as "finite-side gain at most archimedean main term".

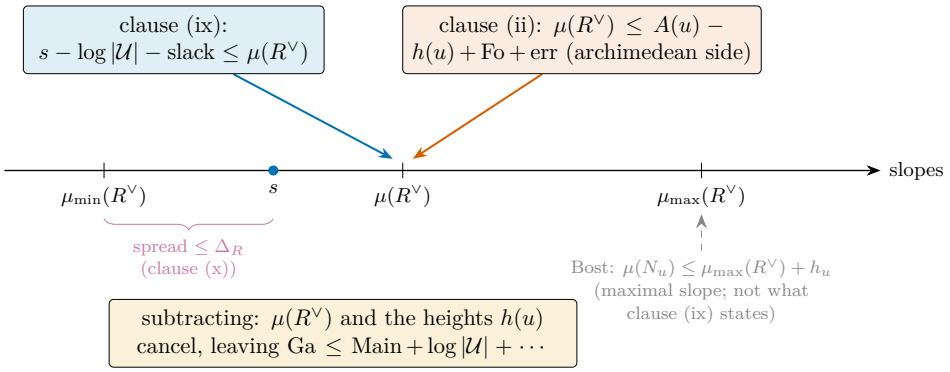


FIGURE 9.2. Clause (ii) of Definition 7.3.2 bounds the slope $\mu(R^\vee)$ of the test lattice from above; here $A(u) = \sum_\tau (m(a_\tau \log a_\tau - a_\tau - a_\tau \log \Omega) - a_\tau \log T(u))$ is the archimedean sum. Clause (ix) bounds it from below in terms of the frame slope s. Subtracting, $\mu(R^\vee)$ and the heights cancel (Remark 9.2.2). Bost's inequality would compare with $\mu_{\max}(R^\vee)$ instead (Remark 7.4.3). Positions on the line are illustrative.

9.3. The square-root bookkeeping

PROPOSITION 9.3.1 (the bookkeeping behind $m \approx \sqrt{x}$). *Let $x \ge 1$, $\varepsilon > 0$, $b > 0$, $k > 0$ and $a_E \ge 0$ be real numbers and n, n_T, n_{T_2}, m positive integers. Suppose that $a_E \le 2xb$, $n_T \ge x^{-\varepsilon/8} n/2$, $x^{1/2-\varepsilon} \le m \le x^{1/2+\varepsilon/2}$, $k \le n^2 x^{1/2-\varepsilon}$, $xm \le k$ and $n_{T_2} \le n$. Then the error bracket* Err *of Definition 6.2.3 satisfies*

$$\frac{\text{Err}}{n_T m b} \le \frac{2x}{m} + 2 + 4m\, x^{9\varepsilon/8} + 2x^{1/2+9\varepsilon/8} \le 10\, x^{1/2+13\varepsilon/8}.$$

PROOF. Divide each of the three terms of Err by $n_T m b$. Using $a_E \le 2xb$ and $xm \le k$,

$$\frac{a_E(1 + m^2/k)}{mb} \le \frac{2x}{m} + \frac{2xm}{k} \le \frac{2x}{m} + 2.$$

Using $1/n_T \le 2x^{\varepsilon/8}/n$,

$$\frac{2nmx^{\varepsilon}}{n_T} \le 4m\, x^{9\varepsilon/8}.$$

Using also $n_{T_2} \le n$ and $k \le n^2 x^{1/2-\varepsilon}$,

$$\frac{n_{T_2}\, k\, x^{2\varepsilon}}{n^2 n_T} \le 2x^{1/2+9\varepsilon/8}.$$

This is the first inequality. For the second, $m \ge x^{1/2-\varepsilon}$ gives $2x/m \le 2x^{1/2+\varepsilon}$, and $m \le x^{1/2+\varepsilon/2}$ gives $4mx^{9\varepsilon/8} \le 4x^{1/2+13\varepsilon/8}$; since $x \ge 1$, each of the four terms is at most a multiple of $x^{1/2+13\varepsilon/8}$, with multiples $2, 2, 4, 2$. □

The exponent $1/2$ is the balance point of the terms x/m and m: the largest of x/m, m and $x^{1/2}$ is smallest at $m = x^{1/2}$, and it is $x^{1/2}$ there (Figure 9.3). The constraints on m and k are hypotheses: for the families of Chapter 12 they are part of the ties (Definition 9.1.4). The proof of Theorem 6.3.4 divides by $n_T mb\,\delta_2/2$ with $\delta_2 = n_{T_2}/n$ instead of $n_T mb$. Since $x^{-\varepsilon} \le n_{T_2}/n$ by the rest-size conditions, the extra factor satisfies $2/\delta_2 \le 2x^{\varepsilon}$, which changes only the multiple of ε in the exponent.

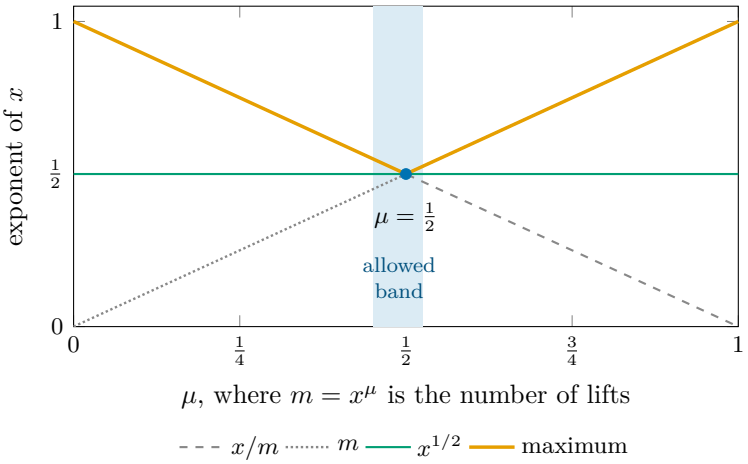


FIGURE 9.3. Exponents of x in the three terms of $\mathrm{Err}/(n_T mb) \ll (x/m + m + x^{1/2})x^{O(\varepsilon)}$ (Proposition 9.3.1), with $m = x^{\mu}$ and the factors $x^{O(\varepsilon)}$ ignored. The largest term is smallest at $m = x^{1/2}$. The shaded band is the range $x^{1/2-\varepsilon} \le m \le x^{1/2+\varepsilon/2}$, drawn at $\varepsilon = 0.05$; this range is part of the hypotheses of Theorem 9.4.1, through the ties of Definition 9.1.4, and it is met by the families of Theorem 12.3.5.

9.4. The Chebyshev inequality on windows

THEOREM 9.4.1 (Alpöge; Chebyshev inequality on windows). *Let $\mathscr{L}$ be a level, e_D and e_L real numbers and $\varepsilon > 0$. Let $\mathfrak{F}$ be a family at ε that satisfies the uniform explicit-formula condition for $\mathscr{L}$ (Definition 9.1.3). Suppose that there are real numbers c_W, c_S, e_K, C_L and x_0 such that $\mathfrak{F}$ is tied (Definition 9.1.4) with constants $e_D, e_L, c_W, c_S, e_K, C_L$ at every $x \in \mathrm{Win}$ with $x \ge x_0$. Then there is x_1 such that*

$$G_{\mathscr{L}}(x) \le x^{1/2+(e_D+e_L+7)\varepsilon}$$

for every $x \in \mathrm{Win}$ with $x \ge x_1$.

PROOF (OUTLINE). The first step discards the geometric data (Figure 9.4). At each $x \in \mathrm{Win}$ with $x \ge x_0$, the tie provides a slope record, arithmetic data of a host and the data $\ell, N, \mathcal{T}, \mathcal{T}_2, a, w, \eta, X^{\flat}, \vartheta$; dropping the

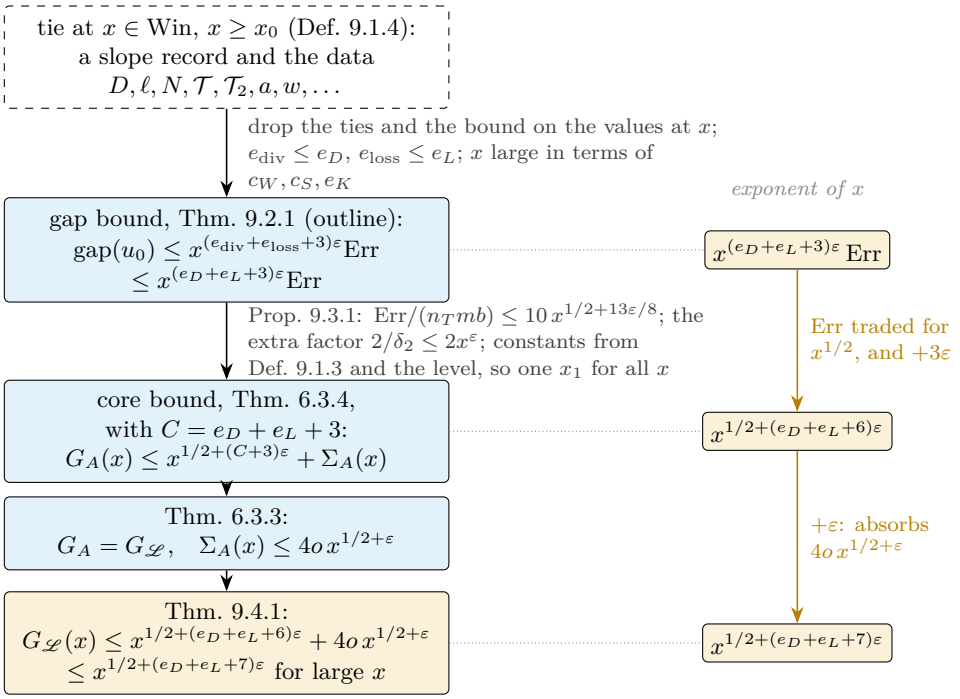


FIGURE 9.4. The proof of Theorem 9.4.1 as a chain (schematic). Left: the theorems applied, with their conclusions; beside the arrows, what each step uses. Right: the exponent of x carried along. The increments $+3\varepsilon$ and $+\varepsilon$ are read off the statement of Theorem 6.3.4 and the last step of the proof; the exponents e_D and e_L enter only through the gap bound, whose proof is given in outline.

ties themselves and the bound on the values at x leaves hypotheses on real numbers only. Since $e_{\mathrm{div}} \leq e_D$ and $e_{\mathrm{loss}} \leq e_L$, Theorem 9.2.1 bounds the gap by $x^{(e_D+e_L+3)\varepsilon}\mathrm{Err}$ at every shift, once x is large in terms of c_W, c_S, e_K. Theorem 6.3.4, with $C = e_D+e_L+3$, then gives $G_A(x) \leq x^{1/2+(e_D+e_L+6)\varepsilon}+\Sigma_A(x)$, where the constants of the character conditions come from the uniform explicit-formula condition and from the level, so that one x_1 serves for all x. Theorem 6.3.3 identifies G_A with $G_{\mathscr{L}}$ and bounds $\Sigma_A(x)$ by $4o\, x^{1/2+\varepsilon}$. So

$$G_{\mathscr{L}}(x) \leq x^{1/2+(e_D+e_L+6)\varepsilon} + 4o\, x^{1/2+\varepsilon},$$

which is at most $x^{1/2+(e_D+e_L+7)\varepsilon}$ for large x, since $e_D, e_L \geq 0$ by the window conditions. □

REMARK 9.4.2 (no vacuity). Theorem 9.4.1 would say nothing for a family whose window set has no large elements, or if the exponents e_D and e_L could grow as $\varepsilon \to 0$. Neither happens. The window set of a family is unbounded (Definition 8.1.1). The same families and the same ties occur in

the existence statement Theorem 12.3.5, where the window set is required to have windows of a fixed exponent, and the exponents e_D and e_L are fixed before ε is sent to 0 (Figure 12.2).

REMARK 9.4.3 (where the exponent 1/2 comes from). Chebyshev's identity alone gives $\mathrm{Ch}(x) \asymp x$, so a one-sided bound with square-root error must come from somewhere else. In the present argument it comes from the way the errors of the slope record depend on the number m of theta lifts. The main term $\mathrm{Main}(u_0)$ is a sum of m copies of an archimedean term, one for each lift, and so is of size about $m\, n_T a \log a$. The error bracket Err has two kinds of term. The first, $n_T a_E(1 + m^2/k)$, does not grow with m once $k \geq xm$; it collects errors such as $c_{\mathrm{div}} n_T a \Upsilon_a$ in the divisor bound (clause (ii)), which do not depend on m. The other kind, of the form $m(B_0 + N_c)$, grows like m^2 through $B_0 \leq 2nmbx^{\varepsilon}$, while its part mN_c gives, through the bound on N_c and $k \leq n^2 x^{1/2-\varepsilon}$, the term of size $x^{1/2}$ of Proposition 9.3.1. After division by $n_T mb$ the first kind and the B_0 part become the terms x/m and m of Proposition 9.3.1, and their balance at $m \approx x^{1/2}$ is the exponent 1/2. If the errors of the first kind grew linearly in m, the first term would be of size x, the size given by Chebyshev's identity.

The inequality that the argument proves is, through Proposition 6.1.4 and Remark 6.1.5, a one-sided bound for a Legendre-type sum over the primes inert in K, with the exact main term and an error $x^{1/2+O(\varepsilon)}$ after normalisation. The main term enters with coefficient 1 through an identity (Proposition 6.1.4) and through the archimedean side of clause (ii); no sieve is used.

CHAPTER 10

Tate–Shafarevich groups of Tate twists

This chapter states a uniform bound for the Tate–Shafarevich groups of the Tate twists $\mathbb{Q}_l/\mathbb{Z}_l(-e)$ over the host fields, the Sha bound on windows (Theorem 10.5.1), and the results from which it is put together: a case split of the eigenspaces, proved in full (Proposition 10.2.3), a bound for the minus part (Theorem 10.3.1), bounds for the plus part (Theorems 10.4.2, 10.4.3 and 10.4.9), two auxiliary theorems (Theorems 10.4.6 and 10.4.7) and a pigeonhole theorem that produces the windows (Theorem 10.4.8). Most of these theorems are stated with an outline or a description of the route of proof only; each proof environment says which. In Chapter 3 the Sha bound on windows is a black box: it is the second step of the proof of the main theorem, its automorphy input is Theorem 11.1.4 of Chapter 11 (Remark 11.3.1), and its output is the hypothesis of the supply of hosts, Theorem 12.1.2 of Chapter 12.

The congruence defects at the twisted places of the families of Chapter 12 are bounded by lengths of Selmer groups, and these in turn by a bound on the Tate–Shafarevich groups of the Tate twists $\mathbb{Q}_l/\mathbb{Z}_l(-e)$ over the host fields. The bound is needed uniformly in the auxiliary pair (q, d), and it is generous in size: its total is allowed to be $\exp(x^{\varepsilon}/2)$, while the classical heuristics predict a size polynomial in x. The difficulty is uniformity, and above all the eigenspaces of the wrong parity, where no L-value controls the group. For a bounded number of exceptional Tate twists the only bound available is doubly exponential, and a pigeonhole argument then gives the bound only on windows of scales. This is where the windows of Chapter 4 come from.

The logical structure of the chapter is shown in Figure 10.1. No result in it carries an unproved hypothesis in its statement. The proofs are Alpöge's claims (Remark 1.7.1). The automorphy that the plus-part bounds use is Theorem 11.1.4(b), proved in Chapter 11, and the output feeds the supply of hosts in Chapter 12.

Throughout this chapter l denotes a prime number, the coefficient prime; it is kept apart from the function ℓ of Chapter 6.

10.1. The groups and the bound

Definition 10.1.1 (Tate–Shafarevich groups of Tate twists). Let E be a number field with an algebraic closure $\overline{E}$, let l be a prime, and let $e \in \mathbb{Z}$. Let S_l be the set of places of E above l, and let G_{E,S_l} be the Galois group

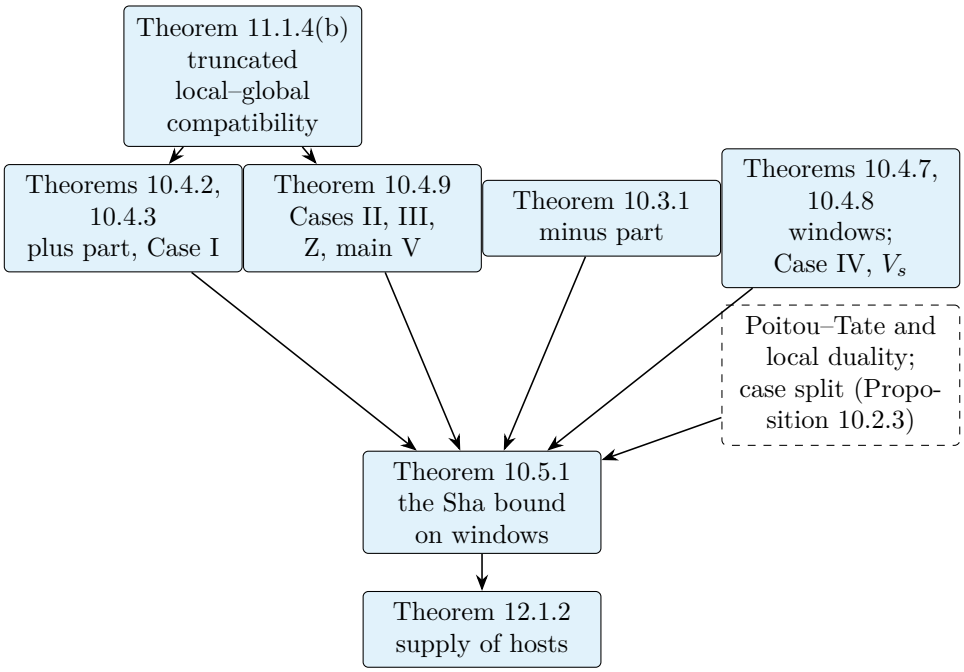


FIGURE 10.1. Dependencies in Chapter 10. Shaded boxes are results stated in this chapter; the white box collects classical inputs. An arrow points from a result to a result whose proof uses it.

of the maximal extension of E in $\overline{E}$ unramified outside S_l and the infinite places. Choose a decomposition group at each place of S_l; the group defined below, and its length, do not depend on these choices or on the choice of $\overline{E}$, up to isomorphism.

(1) The *Tate–Shafarevich group* of the twist e is
$$Ш_l(E,e) = Ш^1(G_{E,S_l}, \mathbb{Q}_l/\mathbb{Z}_l(-e)),$$
the group of classes in $H^1(G_{E,S_l}, \mathbb{Q}_l/\mathbb{Z}_l(-e))$ whose restrictions to the chosen decomposition groups at S_l vanish.
(2) Its *length* $\mathrm{len}_l(e) \in \mathbb{Z}_{\geq 0} \cup \{\infty\}$ is the length of $Ш_l(E,e)$ as a $\mathbb{Z}_l$-module.

DEFINITION 10.1.2 (Sha budgets and Sha bounds on windows). Let $t \geq 1$ be an integer and $B \geq 0$ a real number, and let ψ be a primitive Dirichlet character.

(1) A number field E has *Sha budget* (t, B) if $\mathrm{len}_p(e) < \infty$ for every prime p and every $0 < e < t$, and
$$\sum_{l \in S} \log l \sum_{1 \leq e < t} \mathrm{len}_l(e) \leq B$$

for every finite set S of primes.

(2) Let $C_t > 0$. The character ψ has *Sha bounds on windows with constant* C_t if the following holds for every $\varepsilon \in (0, 1/52)$ and every function ν on $[1, \infty)$ such that $\nu(x)$ is adapted to (ε, x) for every $x \geq 1$ (Definition 5.3.1). There are $J \geq 1$ and X_0 such that for every $X \geq X_0$ and every window pair (q, d) for $(\varepsilon, X^{1/J})$ (Definition 5.3.1) there is a real Y with $X^{1/J} \leq Y$ and $Y^2 \leq X$ such that, for every $x \in [Y, Y^2]$, every ψ-field for $(\psi, \nu(x), q, d)$ (Definition 5.1.3) has Sha budget $(\lfloor x^{C_t} \rfloor, \exp(x^{\varepsilon})^{1/2})$.

(3) The *Sha bound on windows* is the statement that every primitive Dirichlet character has Sha bounds on windows with constant 401.

In words: on some window $[Y, Y^2]$ inside every range $[X^{1/J}, X]$, the CM fields of degree about $x^{2+8\varepsilon}$ attached to ψ satisfy $\sum_l \log l \sum_{1 \leq e < x^{401}} \mathrm{len}_l(e) \leq \exp(x^{\varepsilon}/2)$. The window $[Y, Y^2]$ may depend on X and on the pair (q, d), but the exponent J may not.

Remark 10.1.3 (duality and K-theory). By Poitou–Tate duality [**Tat63, Poi67**] the group $Ш_l(E, e)$ is dual to $Ш^2(G_{E,S_l}, \mathbb{Z}_l(e+1))$, which is a subgroup of $H^2(G_{E,S_l}, \mathbb{Z}_l(e+1))$. For odd l the Quillen–Lichtenbaum comparison identifies this H^2 with the l-part of $K_{2e}(\mathcal{O}_E)$; the comparison rests on the theorem of Voevodsky and Rost (the Bloch–Kato conjecture, proved for odd l in [**Voe11**]). So the Sha bound on windows is, in classical terms, a bound on the subgroup $Ш^2$ of the l-part of $K_{2e}(\mathcal{O}_E)$, the kernel of localisation at the places above l, uniform in l; it is not a bound on the whole l-part. Figure 10.2 sets out these identifications.

$$Ш^1\big(G_{E,S_l}, \mathbb{Q}_l/\mathbb{Z}_l(-e)\big) \xrightarrow[\text{Poitou–Tate}]{\cong} Ш^2\big(G_{E,S_l}, \mathbb{Z}_l(e+1)\big)^{\vee}$$

$$Ш^2\big(G_{E,S_l}, \mathbb{Z}_l(e+1)\big) \hookrightarrow H^2\big(G_{E,S_l}, \mathbb{Z}_l(e+1)\big) \xrightarrow[\text{QL, odd } l]{\cong} K_{2e}(\mathcal{O}_E) \otimes \mathbb{Z}_l$$

The quantity bounded by the Sha bound on windows:
$\sum_l \log l \sum_{1 \leq e < x^{401}} \mathrm{len}_l(e) \leq e^{x^{\varepsilon}/2}$ on a window. Heuristic size:
$\log |K_{2e}(\mathcal{O}_E)| \approx e[E : \mathbb{Q}] \log(e \,\mathrm{rd}_E)$.

Figure 10.2. The classical dictionary for the groups bounded by the Sha bound on windows (Definition 10.1.2): Poitou–Tate duality and, dashed, the Quillen–Lichtenbaum identification for odd l, which rests on the theorem of Voevodsky and Rost (Remark 10.1.3). The size estimate is heuristic.

Remark 10.1.4 (the size of the budget). Heuristically $\log |K_{2e}(\mathcal{O}_E)| \approx e[E : \mathbb{Q}] \log(e \,\mathrm{rd}_E)$, where rd_E is the root discriminant. With $[E : \mathbb{Q}] \approx x^{2+8\varepsilon}$ and $e < x^{401}$ the total is polynomial in x, far below the allowed $\exp(x^{\varepsilon}/2)$. The bound therefore has a wide margin in size. It must hold uniformly in

(q, d), and for that the eigenspaces of the wrong parity must be controlled. This step can be compared with Vandiver's conjecture (Remark 10.4.10).

10.2. Eigenspaces and the case split

The host fields are abelian over $\mathbb{Q}$, so their Tate–Shafarevich groups split into eigenspaces for the characters of the Galois group, and complex conjugation acts on each eigenspace by a sign. The bound is proved eigenspace by eigenspace, and the eigenspaces fall into a minus part and six kinds of plus part.

As elsewhere in these notes, ψ is the primitive character from which a host is built. In this section and the next, η denotes a Dirichlet character of the host modulus whose eigenspace is considered, an *eigen-character*.

Definition 10.2.1 (abelian data and their fields). Let $N_M \geq 1$ be an integer.

(1) An *abelian datum* M of modulus N_M is a subgroup $H_M \leq (\mathbb{Z}/N_M\mathbb{Z})^\times$. Its *degree* is the index $[(\mathbb{Z}/N_M\mathbb{Z})^\times : H_M]$.
(2) A *prime pair* for M is a pair of primes (q, d) with $q \equiv 3 \pmod 8$, $d \equiv 7 \pmod{16}$ and $\gcd(qd, N_M) = 1$.
(3) Let $\nu \geq 0$ be an integer and (q, d) a pair of primes. A number field E is a *field for* (M, ν, q, d) if there are an integer N divisible by $2^{\nu+2}$, N_M, q and d and an embedding $E \hookrightarrow \mathbb{Q}(\zeta_N)$ whose image is the fixed field of the classes $a \in (\mathbb{Z}/N\mathbb{Z})^\times$ with $a \equiv \pm 1 \pmod{2^{\nu+2}}$, $a \bmod N_M \in H_M$, $(a/q) = 1$ and $(a/d) = 1$.
(4) The *conductor* of an abelian number field E is the least $N \geq 1$ such that E embeds into $\mathbb{Q}(\zeta_N)$.

For a primitive character ψ of conductor N_ψ and the datum $M = (N_\psi, \ker\psi)$, the fields for (M, ν, q, d) are the ψ-fields of Definition 5.1.3.

Definition 10.2.2 (eigenspaces and cases). Let E be a CM field, l a prime and $e \in \mathbb{Z}$.

(1) The *plus part* and the *minus part* of $Ш_l(E, e)$ are its eigenspaces for complex conjugation with eigenvalues 1 and -1.
(2) A character η of $\mathrm{Gal}(E/\mathbb{Q})$ is *critical* for e if $\eta(-1) = (-1)^{e+1}$, and *plus* for e if $\eta(-1) = (-1)^e$.
(3) A *host datum* $\mathfrak{h}$ consists of a primitive Dirichlet character ψ of conductor N_ψ and order o_ψ, an integer $\nu \geq 2$ and a prime pair (q, d) for $(N_\psi, \ker\psi)$. Its *modulus* is $N_\mathfrak{h} = \mathrm{lcm}(2^{\nu+2}, N_\psi)\, qd$, and $H_\mathfrak{h} \leq (\mathbb{Z}/N_\mathfrak{h}\mathbb{Z})^\times$ is the subgroup of classes u with $u \equiv \pm 1 \pmod{2^{\nu+2}}$, $\psi(u) = 1$, and u a square modulo q and modulo d.
(4) Let p_K be the least prime with $p_K \equiv 3 \pmod 4$, $p_K \geq 7$ and $p_K \nmid N_\psi$; let h^* be the class number of $\mathbb{Q}(\sqrt{-p_K})$; and let p_1 be the least prime that is $\equiv 1$ modulo the odd part of o_ψ and prime to $2N_\psi p_K$. The *bad set* B_ψ is the set of odd primes dividing $o_\psi N_\psi p_K h^* p_1$, together with 3.

(5) The host datum $\mathfrak{h}$ is *large for* an integer $t \geq 1$ if $2t+1 < q$, $2t+1 < d$, and each of N_ψ, o_ψ, p_K, h^*, p_1 is less than q and less than d.

(6) Let $N \geq 1$ and let r be a prime. Write $(\mathbb{Z}/N\mathbb{Z})^\times = (\mathbb{Z}/r^{v_r(N)}\mathbb{Z})^\times \times (\mathbb{Z}/N'\mathbb{Z})^\times$ with $r \nmid N'$, and let pr_r be the endomorphism of $(\mathbb{Z}/N\mathbb{Z})^\times$ that is the identity on the first factor and trivial on the second. For a character θ modulo N put $\theta|_r = \theta \circ \mathrm{pr}_r$. For a character η modulo N, let $\eta^{(l')}$ be the prime-to-l part of η (the component of order prime to l in the decomposition of η into its l-part and its prime-to-l part), and put
$$\chi'_l(\eta) = \eta^{(l')} \cdot (\eta^{(l')}|_l)^{-1}.$$

(7) Let $\mathfrak{h}$ be a host datum, $t \geq 1$, η a character modulo $N_\mathfrak{h}$ trivial on $H_\mathfrak{h}$, l an odd prime and $0 < e < t$. The triple (l, e, η) is in
   (a) *Case I* if $l \notin B_\psi$ and $(l-1) \nmid 2e$;
   (b) *Case Z* if $l \notin B_\psi$, $(l-1) \mid 2e$ and $\eta = 1$;
   (c) *Case II* if $e \geq 2$, $\chi'_l(\eta)$ has even order, and $l \in B_\psi$ or $(l-1) \mid 2e$;
   (d) *Case III* if $l \notin B_\psi$, $(l-1) \mid 2e$, and $\eta \neq 1$ has odd order;
   (e) *Case IV* if $l \in B_\psi$, $e \geq 2$ and $\chi'_l(\eta)$ has odd order;
   (f) *Case V* if $l \in B_\psi$ and $e = 1$.

The projection pr_r keeps the component in the factor $(\mathbb{Z}/r^{v_r(N)}\mathbb{Z})^\times$ of the Chinese remainder decomposition. It is not the projection onto the Sylow r-subgroup: the factor $(\mathbb{Z}/r^{v_r(N)}\mathbb{Z})^\times$ has order prime to r in general. So $\chi'_l(\eta)$ is the prime-to-l part of η with its component at the modulus $l^{v_l(N)}$ removed.

Proposition 10.2.3. *Let $\mathfrak{h}$ be a host datum with primitive character ψ, let $t \geq 1$ be an integer for which $\mathfrak{h}$ is large, let η be a character modulo $N_\mathfrak{h}$ trivial on $H_\mathfrak{h}$, let l be an odd prime and let $0 < e < t$.*

(a) If $l \notin B_\psi$, then l does not divide the order of η and $\eta^{(l')} = \eta$; if in addition $(l-1) \mid 2e$, then $e \geq 2$, $l \nmid N_\mathfrak{h}$ and $\chi'_l(\eta) = \eta$. The triple (l, e, η) is in Case I, Z, II or III.

(b) If $l \in B_\psi$, the triple (l, e, η) is in Case II, IV or V.

The prime $l = 2$ lies outside the case split, and the minus part is treated separately (Theorem 10.3.1). The 2-parts of the lengths are bounded by the first clause of Theorem 10.4.8. Figure 10.3 shows, for one ψ, where the cases lie in the plane of (l, e), and Figure 10.4 shows the case split with the result that bounds each case.

Proof of (a). Let l be an odd prime not in B_ψ. Since $3 \in B_\psi$, we have $l \geq 5$, and l divides none of o_ψ, N_ψ.

We first show that the order of η divides $2^{\nu+1} o_\psi$. The group $H_\mathfrak{h}$ is the intersection of the kernels of four homomorphisms on $(\mathbb{Z}/N_\mathfrak{h}\mathbb{Z})^\times$, namely reduction modulo $2^{\nu+2}$ followed by the quotient by $\{\pm 1\}$, the character ψ, and the quadratic residue symbols modulo q and d. Hence $(\mathbb{Z}/N_\mathfrak{h}\mathbb{Z})^\times / H_\mathfrak{h}$

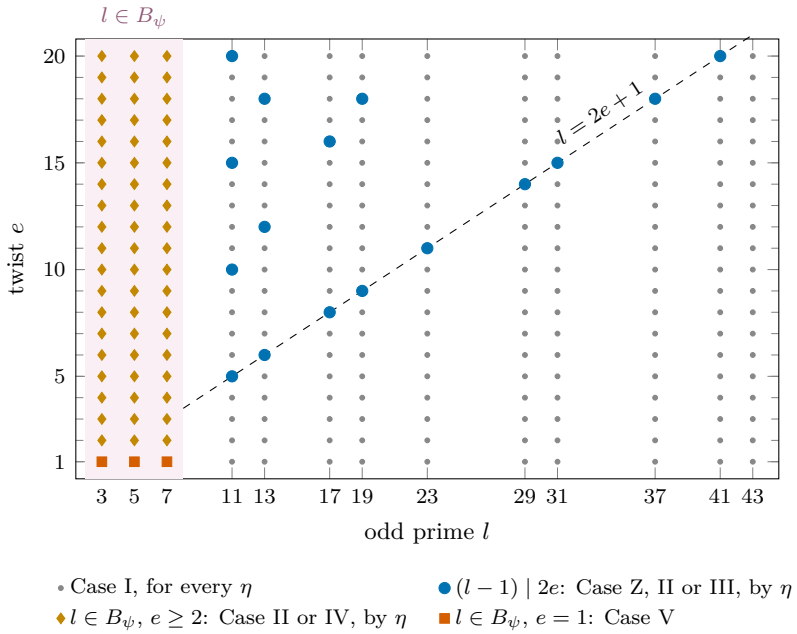


FIGURE 10.3. The case split of item (7) of Definition 10.2.2 in the plane of (l, e), for ψ the quadratic character of conductor 5 and $1 \leq e \leq 20 < t$. Here $p_K = 7$, $h^* = 1$ and $p_1 = 3$, so $B_\psi = \{3, 5, 7\}$. Outside the columns $l \in B_\psi$ the case of (l, e, η) is Case I for every η unless $(l-1) \mid 2e$; these exceptional points lie on or above the dashed line, so $l \leq 2e + 1 < 2t$, as in the proof of Proposition 10.2.3. The figure records only which cases a point allows; which characters η occur depends on the host datum.

embeds into a product of groups of exponents dividing 2^ν, o_ψ, 2 and 2, and its exponent divides $\operatorname{lcm}(2^\nu, o_\psi, 2)$, hence $2^{\nu+1} o_\psi$. A character trivial on $H_\mathfrak{h}$ factors through this quotient. As l is odd and $l \nmid o_\psi$, l does not divide the order of η, so the l-part of η is trivial and $\eta^{(l')} = \eta$.

Suppose now that $(l-1) \mid 2e$. Since $l \geq 5$ and $e > 0$, we get $4 \leq l-1 \leq 2e$, so $e \geq 2$; and $l - 1 \leq 2e < 2t$, so $l \leq 2t < q$ and $l < d$ because $\mathfrak{h}$ is large for t. The odd prime divisors of $N_\mathfrak{h}$ divide $N_\psi q d$, and l divides none of N_ψ, q, d; hence $l \nmid N_\mathfrak{h}$. Then the factor $(\mathbb{Z}/l^{v_l(N_\mathfrak{h})}\mathbb{Z})^\times$ is trivial, pr_l is trivial, $\eta^{(l')}|_l = 1$ and $\chi'_l(\eta) = \eta^{(l')} = \eta$.

Now sort the triple. If $(l-1) \nmid 2e$, it is in Case I. Otherwise $(l-1) \mid 2e$. If $\eta = 1$, the triple is in Case Z. If $\eta \neq 1$ has odd order, it is in Case III. If η has even order, then $\chi'_l(\eta) = \eta$ has even order and $e \geq 2$, so it is in Case II. □

PROOF OF (B). Let $l \in B_\psi$. If $e = 1$ the triple is in Case V. Otherwise $e \geq 2$, and the triple is in Case II if $\chi'_l(\eta)$ has even order and in Case IV if it has odd order. □

REMARK 10.2.4 (the parity test is on $\chi'_l(\eta)$). Cases II and IV test the parity of the order of $\chi'_l(\eta)$. They do not test the order of η or of its l-part. For odd l the l-part of η has odd order, so a test on it could never hold in Case II, and the list would miss the triples in the proof of Proposition 10.2.3 (a) with $(l-1) \mid 2e$ and η of even order. Only Case III tests the order of η itself.

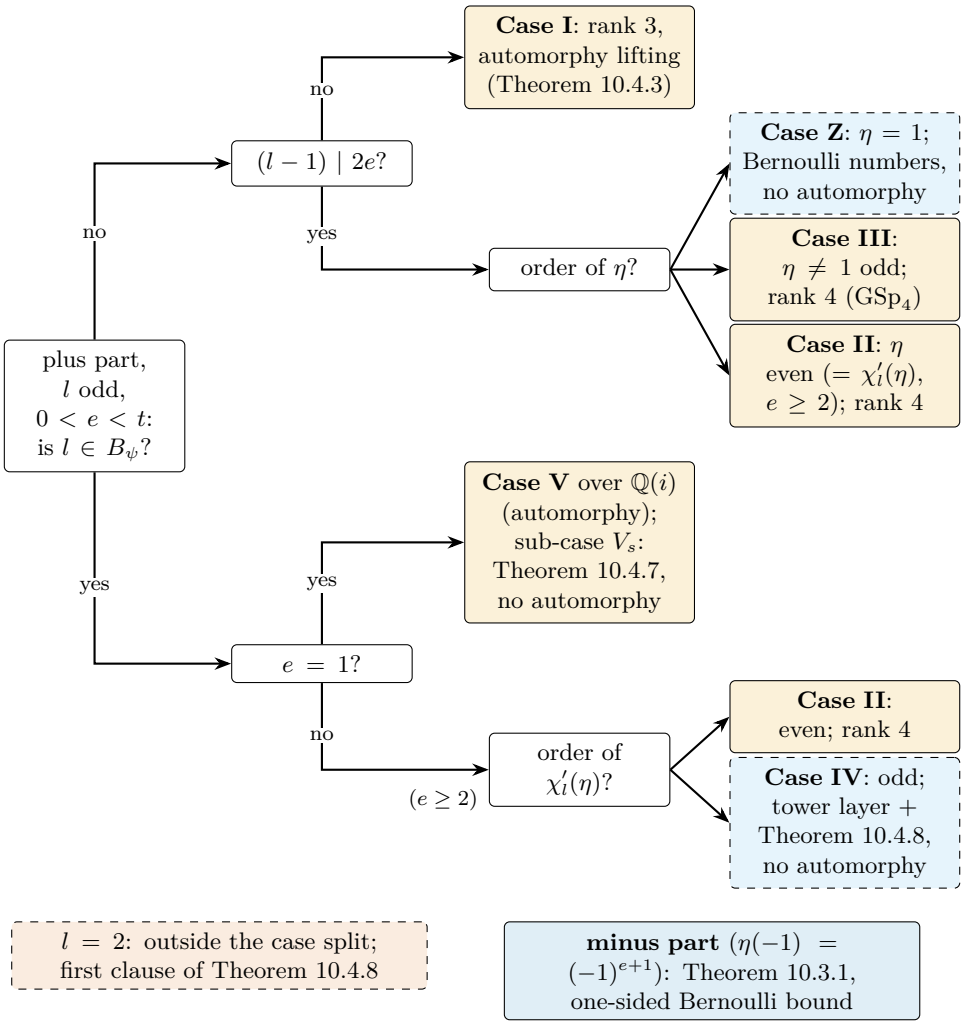
EXAMPLE 10.2.5 (the projection is not the Sylow projection). Let $N = 63$ and $l = 3$, so that $(\mathbb{Z}/63\mathbb{Z})^\times \cong (\mathbb{Z}/9\mathbb{Z})^\times \times (\mathbb{Z}/7\mathbb{Z})^\times \cong C_6 \times C_6$, and write the characters as $\alpha^a\beta^b$, with α a generator of the characters of the first factor and β of the second. Take $\eta = \alpha$, of order 6. Its prime-to-3 part is $\eta^{(l')} = \alpha^3$, a character of the factor $(\mathbb{Z}/9\mathbb{Z})^\times$, so $\eta^{(l')}|_3 = \alpha^3$ and $\chi'_3(\eta) = 1$ has odd order, while η has even order. Projecting to the Sylow 3-subgroup instead would return $\eta^{(l')} = \alpha^3$, of even order. Table 10.1 lists further characters; the host conditions are ignored here.

η	ord	parity	$\eta^{(l')}$	$\eta^{(l')}\|_l$	$\chi'_l(\eta)$	ord	parity	"Sylow" χ'_l	parity
α	6	even	α^3	α^3	1	1	odd	α^3	even
$\alpha\beta$	6	even	$\alpha^3\beta^3$	α^3	β^3	2	even	$\alpha^3\beta^3$	even
$\alpha^2\beta^3$	6	even	β^3	1	β^3	2	even	β^3	even
$\alpha^3\beta$	6	even	$\alpha^3\beta^3$	α^3	β^3	2	even	$\alpha^3\beta^3$	even
β^2	3	odd	1	1	1	1	odd	1	odd
$\alpha^3\beta^3$	2	even	$\alpha^3\beta^3$	α^3	β^3	2	even	$\alpha^3\beta^3$	even
$\alpha^2\beta^2$	3	odd	1	1	1	1	odd	1	odd

TABLE 10.1. For $N = 63$ and $l = 3$: $(\mathbb{Z}/63)^\times \cong (\mathbb{Z}/9)^\times \times (\mathbb{Z}/7)^\times \cong C_6 \times C_6$, with characters $\alpha^a\beta^b$. Here $\eta^{(l')}$ is the prime-to-3 part of η, $\eta^{(l')}|_l$ its composite with the projection to the factor $(\mathbb{Z}/9)^\times$, and $\chi'_l(\eta) = \eta^{(l')}(\eta^{(l')}|_l)^{-1}$. In the first row η has even order and $\chi'_l(\eta)$ is trivial, so the parity test of Cases II and IV differs from a test on η. The last two columns show the result of projecting to the Sylow 3-subgroup instead, which returns $\eta^{(l')}$. The example ignores the host conditions.

10.3. The minus part

THEOREM 10.3.1 (Alpöge). *There is a constant C with the following property. Let M be an abelian datum of degree o, let $\nu \geq 2$ and $t \geq 2$ be integers, let (q, d) be a prime pair for M, and let E be a CM field that is a field for (M, ν, q, d), of degree $d_E = [E : \mathbb{Q}]$ and conductor N. For an*



Solid boxes: an automorphy statement is used. Dashed boxes: no automorphic input. Cases II, III, V and Z: Theorem 10.4.9. $\chi'_l(\eta) = \eta^{(l')}(\eta^{(l')}|_l)^{-1}$, $\eta^{(l')}$ the prime-to-l part of η; for $l \notin B_\psi$ and $(l-1) \mid 2e$ one has $\chi'_l(\eta) = \eta$ (Example 10.2.5).

FIGURE 10.4. The case split of Definition 10.2.2 for the plus part (l an odd prime, $0 < e < t$, the host large, η trivial on $H_{\mathfrak{h}}$), with the result that bounds each case. The cases are exhaustive (Proposition 10.2.3). The prime $l = 2$ lies outside the split, and the minus part is bounded separately (Theorem 10.3.1).

odd prime l and $0 < e < t$ let $\mathrm{len}_l^-(e)$ be the length of the minus part of $Ш_l(E, e)$. Then $\mathrm{len}_l^-(e) < \infty$ for every odd prime l and every $0 < e < t$, and for every finite set S of odd primes

$$\sum_{l \in S} \log l \sum_{1 \le e < t} \mathrm{len}_l^-(e) \le 3 d_E t^2 \log(tN) + C \sum_{\substack{l \mid o \\ l \text{ odd}}} l\, d_E^2 t^2 \log(tlN)^2.$$

Idea of proof. We describe the standard route to a bound of this kind. The minus part of $Ш_l(E, e)$ is the sum of the eigenspaces for the critical characters η of $\mathrm{Gal}(E/\mathbb{Q})$. For an odd l with $l \nmid [E : \mathbb{Q}]$ and η critical for e, the length of the η-eigenspace is bounded by the l-adic valuation of a generalised Bernoulli number attached to $e+1$ and η. This bound comes from Poitou–Tate duality and an orthogonality statement for unramified classes: a system of elements of cyclotomic fields with norm relations (Gauss sums), the valuations given by Stickelberger's theorem, and derivative classes of Kolyvagin type produce enough classes to bound the eigenspace from one side. Summing over l uses that for a nonzero integer B,

$$\sum_l v_l(B) \log l = \log |B|,$$

so one bound on the size of a product of Bernoulli numbers controls every l at once. The second term of the bound accounts for the odd primes dividing the degree o, where $l \mid [E : \mathbb{Q}]$ is possible. □

Remark 10.3.2 (the minus part is a one-sided bound). The bound on the minus part does not use the main conjecture of Iwasawa theory for abelian fields [**MW84**]. It needs only an inequality in one direction, which comes from duality, the orthogonality of unramified classes and the Stickelberger valuations of Gauss sums. No equality and no characteristic ideal enter. The situation is classical: on the critical eigenspaces an L-value is available and bounds the group.

10.4. The plus part

Definition 10.4.1 (coefficient rings, orbits and the hypotheses GD). Let l be a prime, $N \ge 1$ an integer and $\overline{\mathbb{Q}}$ an algebraic closure of $\mathbb{Q}$.

(1) A *coefficient ring* for l is a $\mathbb{Z}_l$-subalgebra $\mathcal{O}$ of an algebraic closure of $\mathbb{Q}_l$ that is a discrete valuation ring and a finitely generated $\mathbb{Z}_l$-module; $e_{\mathcal{O}}$ is its ramification index over $\mathbb{Z}_l$ and $F_{\mathcal{O}}$ its field of fractions.

(2) For a subgroup $H' \le (\mathbb{Z}/N\mathbb{Z})^\times$, the *orbit* $\mathcal{X}_{\mathcal{O}}(H')$ is the finite set of Dirichlet characters modulo N with values in $\mathcal{O}$ whose kernel is exactly H'. For a host datum $\mathfrak{h}$ and $e \ge 0$, $\mathcal{X}_{\mathcal{O}}^+(\mathfrak{h}, e)$ is the set of characters η modulo $N_{\mathfrak{h}}$ with values in $\mathcal{O}$ that are trivial on $H_{\mathfrak{h}}$ and plus for e.

(3) Let η be a character modulo N with values in $\mathcal{O}$, viewed as a character of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, let ω_l be the l-adic cyclotomic character

and let $e \geq 0$. The *eigen-length* $h^1_{\mathcal{O}}(e,\eta) \in \mathbb{Z}_{\geq 0} \cup \{\infty\}$ is the $\mathcal{O}$-length of the group of classes in $H^1(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), (F_{\mathcal{O}}/\mathcal{O})(\eta^{-1}\omega_l^{-e}))$ that are unramified at the primes not dividing lN and trivial on the decomposition groups at the primes dividing lN. Its *normalised length* is $\bar{h}^1_{\mathcal{O}}(e,\eta) = h^1_{\mathcal{O}}(e,\eta)/e_{\mathcal{O}}$ when $h^1_{\mathcal{O}}(e,\eta)$ is finite. Neither length depends on the choice of $\overline{\mathbb{Q}}$.

(4) Let $\mathbb{K} \subset \overline{\mathbb{Q}}$ be a number field, e an integer and η a character modulo N. The *hypotheses* $\mathrm{GD}(\mathbb{K}, l, e, \eta)$ hold if: l is odd; $e \geq 1$; N is even; l does not divide the order of η; $(\eta|_l)^2 = 1$; η is plus for e; $\mathbb{K}$ is an imaginary quadratic field; $l \nmid |\operatorname{disc}\mathbb{K}|\, h_{\mathbb{K}}$; $\gcd(|\operatorname{disc}\mathbb{K}|, N) = 1$; $(l-1) \nmid 2e$; and the intersection of the kernel of the mod-l cyclotomic character of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ with the kernel of $\chi'_l(\eta)$ is not contained in $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{K})$.

The hypotheses $\mathrm{GD}(\mathbb{K}, l, e, \eta)$ are the conditions under which the deformation argument in the proof of Theorem 10.4.2 applies: l is odd and prime to the auxiliary data, the twist e is not divisible by $(l-1)/2$, and $\mathbb{K}$ is not contained in the field cut out by the mod-l cyclotomic character and $\chi'_l(\eta)$, so that these characters stay distinct on $G_{\mathbb{K}}$. For an abelian CM field E the length of the plus part of $Ш_l(E, e)$ is a sum of eigen-lengths over the plus characters of $\mathrm{Gal}(E/\mathbb{Q})$, after the values of the characters are adjoined to the coefficients; this is how the counts below bound the plus part.

THEOREM 10.4.2 (Alpöge). *Let $\mathbb{K}$ be a number field. There are natural numbers C_1 and C_2 with the following property. Let $e \geq 1$ and $N \geq 1$ be integers, $H' \leq (\mathbb{Z}/N\mathbb{Z})^\times$ a subgroup of index i, and p a prime with $p \nmid 30N|\operatorname{disc}\mathbb{K}|$. There is a nonzero integer $\mathcal{N}$ with*

$$\log|\mathcal{N}| \leq C_1\,\big|(\mathbb{Z}/i\mathbb{Z})^\times\big|\,(N(e+1))^{C_2}\log p$$

such that, for every prime $l \neq p$ and every coefficient ring $\mathcal{O}$ for l: if $\mathrm{GD}(\mathbb{K}, l, e, \eta)$ holds for some $\eta \in \mathcal{X}_{\mathcal{O}}(H')$, then $h^1_{\mathcal{O}}(e,\eta) < \infty$ for every $\eta \in \mathcal{X}_{\mathcal{O}}(H')$ and

$$\sum_{\eta \in \mathcal{X}_{\mathcal{O}}(H')} \bar{h}^1_{\mathcal{O}}(e,\eta) \leq 2v_l(\mathcal{N}).$$

The integer $\mathcal{N}$ is chosen before l and $\mathcal{O}$, so one integer controls every prime $l \neq p$ at once. The statement has content only for imaginary quadratic $\mathbb{K}$, since GD requires it. The automorphy input of the proof is Theorem 11.1.4(b) of Chapter 11.

PROOF (OUTLINE). We give Alpöge's route at the level of its main steps. Fix $\mathbb{K}$, e, N, H' and p, and let σ be a Frobenius element at p. A nonzero class in the eigenspace of η, of the wrong parity, gives a nontrivial tangent vector, of weight -1 for the torus action, at the Eisenstein point $\Psi \oplus 1 \oplus \Psi^{-1}$ of an ordinary deformation ring of a residually reducible three-dimensional representation of $G_{\mathbb{K}}$, where Ψ is built from η and the cyclotomic character. The length of the class is bounded by the length of a

graded piece of the cotangent space of this ring at the Eisenstein point. Every characteristic-zero point of the ring is automorphic. This automorphy-lifting theorem of $R = T$ type for residually reducible representations is part of Alpöge's argument, and is not taken from the literature: it is proved from a patching statement for residually reducible determinant rings, together with Hida theory and pro-modularity, and its automorphic input is Theorem 11.1.4(b). The automorphic points are of two kinds: irreducible three-dimensional representations whose cyclotomic twist is attached to a regular algebraic conjugate self-dual cuspidal representation of $\mathrm{GL}_3(\mathbb{A}_{\mathbb{K}})$, and representations $W|_{G_{\mathbb{K}}} \oplus 1$ with W attached to a newform of weight $2e+1$. The cotangent space is then bounded by the congruences between the Eisenstein trace at σ and the traces at the automorphic points: $\mathcal{N}$ is a norm of the product of the differences $\operatorname{tr}\rho(\sigma) - t_{\mathrm{Eis}}$ over these points, where t_{Eis} is the trace of the Eisenstein point at σ. It is nonzero because the automorphic traces are smaller than the Eisenstein trace. In the unitary normalisation the Jacquet–Shalika estimate [**JS81**] gives $p^{-1/2} < |\alpha| < p^{1/2}$ for each Satake parameter α at p of a unitary cuspidal representation of GL_n, and so, for a three-dimensional point, only $|\operatorname{tr}\rho(\sigma)| < 3p^{1/2}$. The bound $|\operatorname{tr}\rho(\sigma)| \le \sqrt{p}+2$ used here follows when the eigenvalues of $\rho(\sigma)$ have the form $\{\zeta, \alpha, \alpha^{-1}\}$ with $|\zeta| = 1$, as for the points $W|_{G_{\mathbb{K}}} \oplus 1$ when $\det W(\sigma) = 1$: then $|\operatorname{tr}\rho(\sigma)| \le 1 + |\alpha| + |\alpha|^{-1} < 1 + p^{1/2} + p^{-1/2} < \sqrt{p}+2$. That the eigenvalues have this form at all the automorphic points is assumed here, not derived; it is where conjugate self-duality enters. With the bound $\sqrt{p}+2$ the Eisenstein trace is larger (Figure 10.5b shows the case $e = 1$, with q in the role of p). □

THEOREM 10.4.3 (Alpöge). *Let ψ be a primitive Dirichlet character. There are a natural number C and primes r_1, r_2 with the following property. Let $\mathfrak{h}$ be a host datum with character ψ whose primes q, d exceed r_1 and r_2, let t be an integer for which $\mathfrak{h}$ is large, and let $0 < e < t$. Let $\overline{\mathbb{Q}}$ be an algebraic closure of $\mathbb{Q}$, choose a coefficient ring $\mathcal{O}_l$ for every prime l, and let L be a finite set of odd primes l with $l \notin B_\psi$ and $(l-1) \nmid 2e$. Then $h^1_{\mathcal{O}_l}(e,\eta) < \infty$ for every $l \in L$ and $\eta \in \mathcal{X}^+_{\mathcal{O}_l}(\mathfrak{h}, e)$, and*

$$\sum_{l \in L} \log l \sum_{\eta \in \mathcal{X}^+_{\mathcal{O}_l}(\mathfrak{h},e)} \bar{h}^1_{\mathcal{O}_l}(e,\eta) \le (N_{\mathfrak{h}} t)^C.$$

The condition on l in Theorem 10.4.3 says that every triple (l, e, η) is in Case I; it depends only on (l, e), which is why the sum runs over all of $\mathcal{X}^+_{\mathcal{O}_l}(\mathfrak{h}, e)$. The automorphy input of Theorem 10.4.3, through Theorem 10.4.2, is Theorem 11.1.4(b).

PROOF (OUTLINE). We give Alpöge's route at the level of its main steps. The proof applies a form of Theorem 10.4.2 summed over orbits and over primes l, with $\mathbb{K} = \mathbb{Q}(\sqrt{-p_K})$ for the auxiliary prime p_K of Definition 10.2.2, and with two fixed primes $r_1, r_2 > 30 N_\psi p_K$ in the role of p. The discriminant

of $\mathbb{K}$ is $-p_K$ and its class number is h^*, and both p_K and the prime divisors of h^* lie in B_ψ; this is why $l \notin B_\psi$ is imposed. The sum over l costs nothing, since $\sum_l v_l(\mathcal{N}) \log l = \log|\mathcal{N}|$, and the bound on $\log|\mathcal{N}|$ is polynomial in $N_\mathfrak{h}$ and t. □

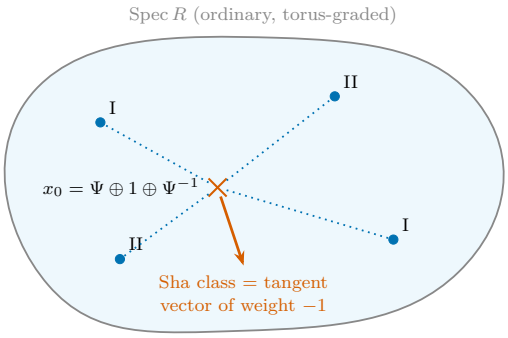
Remark 10.4.4 (the plus part is bounded by automorphy lifting). Theorem 10.4.2 is not proved as a converse to Ribet's theorem [**Rib76**] by a lattice argument on a unitary group in three variables. The route is a cotangent bound on a deformation ring together with the automorphy of all its characteristic-zero points, and no unitary group occurs in the statements. The group in the background is $\mathrm{GL}_3 \rtimes \{\pm 1\}$, which encodes conjugate self-duality. A lattice construction in the style of Ribet appears only in the construction of one rank-two representation used along the way, and no Ramanujan bound is used: the traces are bounded through the Jacquet–Shalika estimate. Figure 10.5 shows the route.

Definition 10.4.5 (Kato configurations). Let l be a prime, $e \geq 1$ and $N \geq 1$ integers, $\mathbb{K}$ a number field with a character θ of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, η a character of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ with values in $\overline{\mathbb{Q}}_l^\times$, and $W\colon \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\overline{\mathbb{Q}}_l)$ a continuous representation. Let ω_l be the l-adic cyclotomic character and put $\Psi = \omega_l^e \eta$. The tuple $(l, e, N, \mathbb{K}, \eta, W)$ is a *Kato configuration* if:

(1) l is odd, $(l-1) \nmid 2e$, N is even, $\mathbb{K}$ is imaginary quadratic, $l \nmid |\operatorname{disc}\mathbb{K}| h_\mathbb{K}$, $\gcd(|\operatorname{disc}\mathbb{K}|, N) = 1$, and θ is the quadratic character of $\mathbb{K}$;
(2) η has finite order prime to l, η^2 is trivial on the inertia groups at l, and $\Psi(c) = 1$ for some complex conjugation c;
(3) W has level dividing $N|\operatorname{disc}\mathbb{K}|$ and Hodge–Tate weights $(e, -e)$, $\det W = \theta$, $\operatorname{tr} W = \Psi + \Psi^{-1}$ on the inertia groups at l, and $\operatorname{tr} W \equiv \theta\Psi + \Psi^{-1}$ modulo the maximal ideal;
(4) W is ordinary at l, with a stable flag at a place above l on whose graded pieces inertia acts by $\omega_l^{\pm e}$ times a common character;
(5) η and W satisfy further local conditions: a condition on η relative to the roots of unity of order N, and, over every imaginary quadratic field over which W is irreducible, a condition at the places above l.

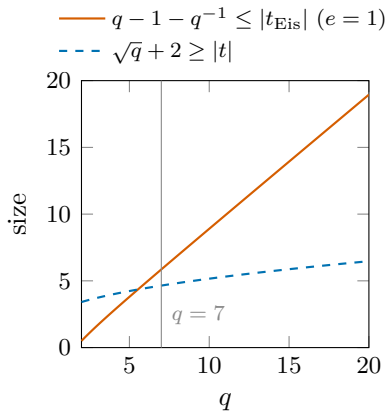
The conditions in (5) are part of the definition and are not written out here. The data mirror those of the hypotheses GD of Definition 10.4.1, with W a two-dimensional representation of the kind that occurs among the automorphic points in the proof of Theorem 10.4.2.

Theorem 10.4.6 (Alpöge). *Let l be an odd prime, $e \geq 1$, $N \geq 1$, $\mathbb{K}$ an imaginary quadratic field with quadratic character θ, η a character of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ and $W\colon \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\overline{\mathbb{Q}}_l)$ a continuous representation, such that $(l, e, N, \mathbb{K}, \eta, W)$ is a Kato configuration. Suppose that $W(-e)$ is attached to a newform of weight $2e+1$. Then the Bloch–Kato Selmer groups of W and of $W \otimes \theta$ vanish.*



Dots: characteristic-zero points, all automorphic (I: irreducible, attached to GL_3 over $\mathbb{K}$; II: $W|_{G_{\mathbb{K}}} \oplus 1$, W from a weight $2e+1$ newform). Dotted "distance" to x_0: $v_l(\mathrm{tr}\,\rho(\sigma) - t)$, $t = \Psi(\sigma) + 1 + \Psi(\sigma)^{-1}$. Bound: length of the Sha class $\leq 2\sum_\rho v_l(\mathrm{tr}\,\rho(\sigma) - t)$. *schematic, not to scale*

(A)



(B)

FIGURE 10.5. (a) The route to the plus-part bound (Theorems 10.4.2 and 10.4.3), schematically: a Sha class of the wrong parity is a tangent vector of weight -1 at the Eisenstein point x_0 of an ordinary deformation ring, and its length is bounded by the congruences between the trace at one Frobenius and the traces at the automorphic points. (b) For $e = 1$, $|t_{\mathrm{Eis}}| \geq q - 1 - q^{-1} > \sqrt{q} + 2$ once $q \geq 6$, so a trace with $|t| \leq \sqrt{q}+2$ differs from the Eisenstein trace. This is the bound of the proof of Theorem 10.4.2, which follows from the Jacquet–Shalika estimate when the eigenvalues at the point pair off as α, α^{-1}; the general Jacquet–Shalika bound $3\sqrt{q}$ separates the traces only for $q \geq 13$. The argument takes $q \geq 7$.

The statement is taken, in a weaker form, from [**Alp26a**, `KatoTwoW.lean`, lines 10–14], where it is the proposition named `KatoTwoW`. No proof of Theorem 10.4.6 is given in this book, and no published source for it is known to us. It is a vanishing theorem of Kato type [**Kat04**]; Kato treats elliptic modular newforms over $\mathbb{Q}$ and their Dirichlet twists, and we do not settle here whether the twist by θ lies in that range. The vanishing is of the kind that Kato's Euler system gives for modular forms.

THEOREM 10.4.7 (Alpöge). *There is a constant C_K with the following property. Let l be a prime, L a totally complex number field containing the l-th roots of unity (the fourth roots of unity if $l = 2$), and $e \geq 1$. Then $H^2(G_{L,S_l}, \mathbb{Z}_l(e+1))$ is finite and*

$$\log|H^2(G_{L,S_l}, \mathbb{Z}_l(e+1))| \leq \exp((2e[L:\mathbb{Q}])^{C_K} \log(2|\operatorname{disc} L|)).$$

The statement is that of [**Alp26a**, `LemmaK.lean`, lines 9–17], where it is the proposition named `LemmaK`. No proof of Theorem 10.4.7 is given in this book. It is a finiteness statement of Soulé type: Soulé [**Sou79**] proves the finiteness of these groups, and the doubly exponential bound is an addition, for which no published source is known to us. The bound is weak, and it is used only for a bounded number of exceptional twists and in the sub-case V_s of Case V (Theorem 10.4.9).

THEOREM 10.4.8 (Alpöge). *There are functions R_0 and R_l (for primes l) from abelian data to the natural numbers and a constant $C_K > 0$ with the following property. Let $C_t \geq 1$. Let M be an abelian datum, B' a finite set of odd primes, and for each $l \in B'$ let M_l be an abelian datum whose fields contain the l-th roots of unity. Let $\varepsilon \in (0, 1/52)$ and $A_1 > 0$, and put $R = R_0(M) + \sum_{l \in B'} R_l(M_l)$ and $J = 2^{R+2}(2C_K C_t/\varepsilon)^R$. There is X_0 such that for all functions ν, t on $[2, \infty)$ with $2 \leq \nu(x)$, $2^{\nu(x)} \leq x^{A_1}$ and $2 \leq t(x) \leq x^{C_t}$, every $X \geq X_0$ and every prime pair (q, d) for M with $\log(qd) \leq X^{\varepsilon/(4J)}$, there is Y with $X^{1/J} \leq Y$ and $Y^2 \leq X$ such that for every $x \in [Y, Y^2]$:*

- *(a) every field E for $(M, \nu(x), q, d)$ satisfies $\mathrm{len}_2(e) < \infty$ for $0 < e < t(x)$ and $\sum_{1 \leq e < t(x)} \mathrm{len}_2(e) \leq \exp(x^{4\varepsilon/5})$;*
- *(b) for every $l \in B'$, every field L_0 cut out by the datum M_l and every field $L_1 = L_0(\mu_{l^k})$, $k \geq 0$, with $[L_1 : L_0] = l^{j_l}$, where $j_l = \lfloor \log_l R_l(M_l) \rfloor + 3$, the field L_1 satisfies $\mathrm{len}_l(e) < \infty$ for $0 < e < t(x)$ and $\sum_{1 \leq e < t(x)} \mathrm{len}_l(e) \leq \exp(x^{4\varepsilon/5})$.*

Theorem 10.4.8 is unconditional, but its proof uses other results: Theorem 10.4.7, Poitou–Tate duality, the finiteness of local Galois cohomology and local Tate duality, and the theorem of Ferrero and Washington [**FW79**]; the exceptional twists are treated by Theorem 10.4.7 alone.

PROOF (OUTLINE). We give Alpöge's route at the level of its main steps. Ferrero–Washington [**FW79**], in the form used, bounds the l-ranks of the class groups uniformly up a cyclotomic tower. A tower bound then bounds the Sha lengths at a layer of the tower linearly in the layer, for every twist e outside a fixed l-adic neighbourhood of each of at most R centres. This leaves at most R exceptional twists in a given range, and for those only the doubly exponential bound of Theorem 10.4.7 is available. An exceptional twist e_i does harm at the scale x only when it lies between a small power x^{κ} and x^{C_t}, that is, for $\log x$ in an interval of fixed ratio. Among at most R such intervals a pigeonhole argument finds, in a range of length depending only on R and the constants, a window $[Y, Y^2]$ free of all of them. This is the origin of the exponent J. □

Case V is divided into two sub-cases. Attached to an odd prime l and an eigen-character η are two natural numbers $R_1(l, \eta)$ and $c_2(\eta)$, explicit in terms of l, η and the host datum, whose definitions we omit. A triple $(l, 1, \eta)$

in Case V is in the *main sub-case* if $R_1(l,\eta) \le 2^{c_2(\eta)}$, and in the *sub-case* V_s otherwise.

THEOREM 10.4.9 (Alpöge). *Let ψ be a primitive Dirichlet character. There is a natural number C with the following property. Let $\mathfrak{h}$ be a host datum with character ψ, t an integer for which $\mathfrak{h}$ is large, $\overline{\mathbb{Q}}$ an algebraic closure of $\mathbb{Q}$, l an odd prime, $\mathcal{O}$ a coefficient ring for l, $0 < e < t$, and $\eta \in \mathcal{X}^+_{\mathcal{O}}(\mathfrak{h}, e)$. Suppose that (l, e, η) is in Case II, III or Z, or in the main sub-case of Case V. Then $h^1_{\mathcal{O}}(e,\eta) < \infty$ and*

$$\bar{h}^1_{\mathcal{O}}(e,\eta) \log l \le (l\, N_{\mathfrak{h}}\, e)^C.$$

Cases IV and V_s are not covered by the theorem: they are bounded without automorphy, Case IV by restriction to a layer of a cyclotomic tower and Theorem 10.4.8, and Case V_s by Theorem 10.4.7 at $e = 1$. The automorphy statements used for Cases II, III and the main sub-case of V are, like that of Theorem 10.4.3, consequences of Theorem 11.1.4(b).

ON THE PROOF. We describe the route by case; this account is ours. Cases II and III follow the pattern of the proof of Theorem 10.4.2 in rank four: ordinary deformations, with values in GSp_4, of a sum of four characters; an automorphy statement in which the automorphic points are attached, over an imaginary quadratic field, to cuspidal representations of GL_4 (no Siegel modular forms occur); a rigidity statement; a congruence number; and an input on the residual nonvanishing of Bernoulli numbers. The main sub-case of Case V has the same shape over $\mathbb{Q}(i)$. Case Z concerns the trivial character, and its bound uses the arithmetic of Bernoulli numbers only, with no automorphy; for $l \ge 5$, $e \ge 2$ even and $(l-1) \mid 2e$ the relevant eigen-length of the trivial character vanishes. □

REMARK 10.4.10 (comparison with Vandiver and Greenberg). Theorems 10.4.3 and 10.4.9 bound the eigenspaces of the wrong parity of $Ш^1(\mathbb{Q}_l/\mathbb{Z}_l(-e))$ over abelian fields, polynomially in the conductor and uniformly in l. For $e = 0$ statements of this kind concern the plus parts of class groups of abelian fields, and are of the type of Vandiver's conjecture and of Greenberg's conjecture on the Iwasawa invariants of totally real fields [**Was97, Gre76**]. We know of no published result of this strength.

10.5. The Sha bound on windows

THEOREM 10.5.1 (Alpöge). *Every primitive Dirichlet character has Sha bounds on windows with constant* 401 *(Definition 10.1.2).*

So the Sha bound on windows holds. The automorphy used in the case bounds, Theorems 10.4.3 and 10.4.9, is Theorem 11.1.4(b); this dependence is recorded in Remark 11.3.1.

ON THE PROOF. We describe how the bounds combine. Fix a primitive ψ, $\varepsilon \in (0, 1/52)$ and an adapted ν, and let E be a ψ-field for $(\psi, \nu(x), q, d)$

part	bound as stated	source	log of the bound on a window
2-part	$\sum_e \mathrm{len}_2(e) \le \exp(x^{4\varepsilon/5})$	Thm. 10.4.8(a)	$x^{4\varepsilon/5}$
minus part, odd l	$3d_E t^2 \log(tN) + C\sum_{l\mid o} l\, d_E^2 t^2 \log(tlN)^2$	Thm. 10.3.1	$O(\log x)$
Case I	$\sum_l \log l \sum_\eta \bar{h}^1_{\mathcal{O}_l}(e,\eta) \le (N_\mathfrak{h} t)^C$ for each e	Thm. 10.4.3	$O(x^{\varepsilon/5}\log^2 x)$
Cases II, III, Z, main V	$\bar{h}^1_{\mathcal{O}}(e,\eta)\log l \le (l\, N_\mathfrak{h}\, e)^C$ for each (l,e,η)	Thm. 10.4.9	$O(x^{\varepsilon/5}\log^2 x)$
Case IV	$\sum_e \mathrm{len}_l(e) \le \exp(x^{4\varepsilon/5})$ for the tower layer L_1, then restriction	Thm. 10.4.8(b)	$x^{4\varepsilon/5}$ for L_1; restriction not quantified
Case V_s	at $e=1$: $\log\lvert H^2\rvert \le \exp\big((2[L:\mathbb{Q}])^{C_K}\log(2\lvert\operatorname{disc} L\rvert)\big)$	Thm. 10.4.7	not quantified
allowed total	$\exp(x^\varepsilon)^{1/2}$	Def. 10.1.2	$x^\varepsilon/2$

FIGURE 10.6. The pieces of the Sha budget of a ψ-field E in the proof of Theorem 10.5.1, for x in a window. There $t = \lfloor x^{401}\rfloor$, $[E:\mathbb{Q}] = d_E$ is about $x^{2+8\varepsilon}$ and $\log N_\mathfrak{h} = O(x^{\varepsilon/5}\log^2 x)$, so the polynomial bounds have logarithm $O(x^{\varepsilon/5}\log^2 x)$ or less, against $x^\varepsilon/2$ for the budget. The sums over $e < t$, over η and, in Cases II, III, Z and main V, over $l < 2t$ or $l \in B_\psi$ (proof of Proposition 10.2.3) multiply by factors polynomial in t and $N_\mathfrak{h}$. The text does not state the size of the Case V_s contribution or of the restriction step in Case IV on a window.

with x in a window. The Sha budget of E splits into three parts: the 2-part, the minus part at odd primes and the plus part at odd primes, so a bound B_2 for the 2-lengths and bounds B_-, B_+ for the other two give the budget $(t, B_2\log 2 + B_- + B_+)$. The window and the 2-part come from Theorem 10.4.8(a), with $C_t = 401$, applied to the datum of ψ with $t(x) = \lfloor x^{401}\rfloor$; the same window serves the bad primes $l \in B_\psi$, with M_l the datum of $\mathbb{Q}(\zeta_{lN_\psi})$, through clause (b). The minus part is bounded by Theorem 10.3.1. For the plus part, the length at an odd prime is the sum of the eigen-lengths over the plus characters, and Proposition 10.2.3 sorts every eigenspace into one of the six cases: Case I is bounded by Theorem 10.4.3; Cases II, III, Z and the main sub-case of V by Theorem 10.4.9; Case IV by restriction to a layer of a cyclotomic tower and the window bound of clause (b); and Case V_s by Theorem 10.4.7. Poitou–Tate duality, the finiteness of local Galois cohomology and local Tate duality [**Tat63, Poi67**] are used to pass between the groups of the field E and the eigen-lengths over $\mathbb{Q}$, and between $Ш^1$ and H^2.

On a window, $\log q$ and $\log d$ are at most $x^{\varepsilon/5}\log^2 x$ and $2^{\nu(x)} \leq 2x^{2+8\varepsilon}$, so $\log N_{\mathfrak{h}} = O(x^{\varepsilon/5}\log^2 x)$. This is why bounds polynomial in $N_{\mathfrak{h}}$, t and l are small against the budget $\exp(x^{\varepsilon})^{1/2}$ (Figure 10.6). □

REMARK 10.5.2 (where the windows and constants come from). The windows do not come from an Iwasawa growth formula $\lambda_{\mathrm{Iw}}\nu + \nu_{\mathrm{Iw}}$ for the Sha lengths in a cyclotomic tower, with invariants controlled only doubly exponentially. No growth formula is used. The windows come from the pigeonhole argument in the proof of Theorem 10.4.8, applied to a bounded number of exceptional twists, and the constants are only doubly exponential because Theorem 10.4.7 is. The arithmetic input exists only on windows, so the analytic criterion of Chapter 4 is windowed (Theorem 4.3.3). The window exponent of the Sha bound is 2, and the supply of hosts in Chapter 12 is stated with the exponent 3/2. How the one becomes the other is Question 12.1.3, which follows the proof of Theorem 12.1.2. One further constant has the same origin as the windows. The constant 401 makes the budget cover $\lfloor x^{401} \rfloor$ Tate twists, while the number of twists needed is the rank parameter t of Definition 7.1.1, which satisfies $t \leq x^{400}$. So 401 leaves a margin of one in the exponent; we do not identify a step that uses the margin.

REMARK 10.5.3 (named inputs that the argument does not use). The argument for Theorem 10.5.1 outlined here does not use Clozel's purity lemma [**Clo90**], the main theorem of complex multiplication or the theorem of Deligne and Ribet [**DR80**]. The bound on traces that makes the congruence number nonzero in Theorem 10.4.2 is the Jacquet–Shalika bound. No bound on Satake parameters from purity is used, and the purity lemma in question concerns infinity types.

CHAPTER 11

Galois representations and local–global compatibility

This chapter states Theorem 11.1.4: weak local Langlands maps exist, and truncated local–global compatibility holds for regular algebraic polarized cuspidal representations over CM fields. Part (a) is stated without an outline of proof, and part (b) with an outline of the named steps of the construction. The chapter also states the elementary Proposition 11.2.3 on compact groups, which is the step of the construction that recovers an a-dimensional representation from an annihilating polynomial. In Chapter 3, Theorem 11.1.4 is the first step of the proof of the main theorem. It is used as a black box twice: as the automorphy input of the Sha bound on windows of Chapter 10 (Remark 11.3.1), and, through its weak local Langlands maps, in the existence of admissible-family data (Theorem 12.3.4(c), Chapter 12).

The automorphy statements of Chapter 10 need Galois representations attached to regular algebraic polarized cuspidal representations over CM fields, with a truncated form of local–global compatibility. This chapter states what is needed, compares it with the literature, and outlines how the representations are constructed. What is needed is weaker than the published theorems in two ways: the local Langlands maps are required to satisfy only a list of laws, and the compatibility is asserted only at the places where the monodromy vanishes.

11.1. Weak local Langlands maps and truncated compatibility

Definition 11.1.1 (weak local Langlands maps). Let E be a number field and w a finite place of E, with residue characteristic p.

(1) An irreducible admissible representation of some $\mathrm{GL}_n(E_w)$ is *good* if its cuspidal support consists of characters, or, when $E_w \cong \mathbb{Q}_p$, of twists of depth-zero cuspidal representations.

(2) A map rc from the irreducible admissible representations of the groups $\mathrm{GL}_n(E_w)$, $n \geq 1$, to Weil–Deligne representations of the Weil group of E_w *satisfies the laws* if:
   (a) it preserves dimension and is local class field theory on characters;
   (b) it is compatible with the Zelevinsky classification by segments, sends cuspidal representations to irreducible ones, and respects

unramified twists and the absolute value of the central character;
   (c) on good representations it respects twists by characters, sends the central character to the determinant and the contragredient to the dual, is injective, and sends depth zero to tame;
   (d) it satisfies a compatibility of types over finite extensions.
(3) A *weak local Langlands map* for (E, w) is a map satisfying the laws.

Clauses (2)(b) and (2)(d) are summarised; their precise form is a list of compatibilities that we do not need. The uniqueness that goes with this definition holds only on good representations: two weak local Langlands maps for (E, w) agree on good representations.

Remark 11.1.2 (the local Langlands input is weaker than the correspondence). The argument does not use the local Langlands correspondence for GL_n [**HT01, Hen00, Sch13**]. It uses only the existence of weak local Langlands maps. These are not required to be bijective, and no condition on L- or ϵ-factors is imposed. On representations that are not good, a weak local Langlands map is constrained only by laws (a), (b) and (d), and the existence statement makes an arbitrary choice there. The correspondence is the stronger statement that for every number field E and finite place w there is exactly one map from the irreducible admissible representations of the groups $\mathrm{GL}_n(E_w)$ to Frobenius-semisimple Weil–Deligne representations that is a bijection, satisfies the laws of Definition 11.1.1 on all representations and preserves the L- and ϵ-factors of pairs, with these maps compatible for different fields.

Definition 11.1.3 (truncated local–global compatibility). Let E be a CM field with maximal totally real subfield E^+ and complex conjugation c, let l be a prime, $\iota\colon \overline{\mathbb{Q}}_l \cong \mathbb{C}$ an isomorphism and $a \geq 1$. Let π be a cuspidal automorphic representation of $\mathrm{GL}_a(\mathbb{A}_E)$.

(1) π is *polarized* if $\pi^c \cong \pi^\vee \otimes (\chi \circ N_{E/E^+})$ for a character χ of $\mathbb{A}_{E^+}^\times/(E^+)^\times$ with the sign condition at the real places of E^+.
(2) A *compatible system for* π consists of a continuous semisimple representation $r\colon G_E \to \mathrm{GL}_a(\overline{\mathbb{Q}}_l)$, an integer wt and, for each embedding $\tau\colon E \to \overline{\mathbb{Q}}_l$, a multiset HT_τ of a integers, such that:
   (a) (weights) for every τ, HT_τ, viewed in $\mathbb{C}$ by ι, is the multiset of the numbers $(a-1)/2 - z$ for z in the infinitesimal character of π at $\iota \circ \tau$, and $\mathrm{HT}_{\tau c} = \mathrm{wt} - \mathrm{HT}_\tau$;
   (b) (places not above l) at every finite place $w \nmid l$ and for every weak local Langlands map rc for (E, w): the dimensions of $\mathrm{rc}(\pi_w)$ and of $r|_{G_{E_w}}$ agree; if $\mathrm{rc}(\pi_w)$ is a sum of characters with monodromy $N = 0$, then the Weil–Deligne representation of $r|_{G_{E_w}}$ is $\mathrm{rc}(\pi_w)$; and if $E_w = \mathbb{Q}_p$ with $a = 1$ or $p > 16a^2$, and $\mathrm{rc}(\pi_w)$ is tame, multiplicity free and has $N = 0$, then

the Weil–Deligne representation of $r|_{G_{E_w}}$ has $N = 0$ and the characteristic polynomials of inertia on the two agree;

(c) (places above l) at every place $w \mid l$: if π_w is spherical, then $r|_{G_{E_w}}$ is crystalline with Hodge–Tate weights $(\mathrm{HT}_\tau)_\tau$; and for every weak local Langlands map rc such that $\mathrm{rc}(\pi_w)$ has $N = 0$ and is either a sum of characters or satisfies the three conditions of (b), the restriction $r|_{G_{E_w}}$ is potentially semistable with these Hodge–Tate weights, and its Weil–Deligne representation is $\mathrm{rc}(\pi_w)$ in the first case and, in the second, has $N = 0$ and the same characteristic polynomials of inertia as $\mathrm{rc}(\pi_w)$.

(3) *Truncated local–global compatibility* is the statement that for every CM field E, every prime l, every ι, every $a \geq 1$ and every regular algebraic polarized cuspidal automorphic representation π of $\mathrm{GL}_a(\mathbb{A}_E)$ there is a compatible system for π.

Figure 11.1 lists the local clauses case by case. Compared with local–global compatibility in the literature, nothing is asserted about the monodromy operator except where it vanishes, nothing about purity, and at places above l only the cases listed in (2)(c).

THEOREM 11.1.4 (Alpöge). *(a) For every number field E and every finite place w of E, a weak local Langlands map for (E, w) exists.*

(b) Truncated local–global compatibility (Definition 11.1.3) holds.

In the literature the corresponding results are due to Kottwitz, Clozel, Harris–Taylor, Shin, Chenevier–Harris, Caraiani and others [**Kot92, HT01, TY07, Shi11, CH13, Car12, Car14, BLGGT14**]. The construction below agrees with that literature in its frame and, as far as we know, differs from it at three inner steps.

PROOF (OUTLINE). *Proof of (a).* We do not outline this part. On representations that are not good, the map it provides is an arbitrary choice compatible with laws (a), (b) and (d).

Proof of (b). We describe only the named steps of the construction.

(1) The Shimura varieties are those of the unitary similitude group of a hermitian space of odd dimension a over E, of signature $(1, a-1)$ at one infinite place and definite at the others. The order of the PEL datum has rank one, so no division algebra occurs. The cohomology used is that of the Kuga–Sato powers of the universal abelian scheme, in all degrees.

(2) The Hecke eigensystem of a form on the definite unitary group is put into this cohomology by a holomorphic Lefschetz formula for correspondences, compared with a Hecke trace on the definite group. This gives occurrence of the eigensystem, not its multiplicity.

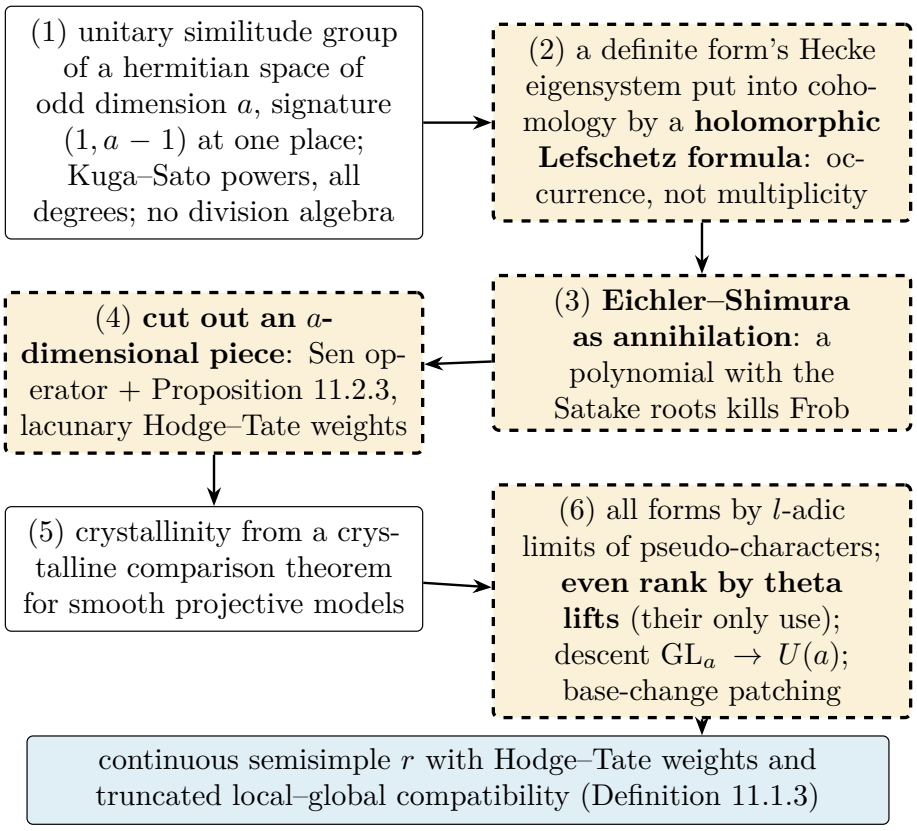
condition	asserted about $r\vert_{G_{E_w}}$
places $w \nmid l$; every row is for every weak local Langlands map rc *for* (E, w)	
always	$\dim r\vert_{G_{E_w}} = \dim \mathrm{rc}(\pi_w)$
$\mathrm{rc}(\pi_w)$ a sum of characters, $N = 0$	its Weil–Deligne representation is $\mathrm{rc}(\pi_w)$
tame row: $E_w = \mathbb{Q}_p$ with $a = 1$ or $p > 16a^2$; $\mathrm{rc}(\pi_w)$ tame, multiplicity free, $N = 0$	its Weil–Deligne representation has $N = 0$, and its characteristic polynomials of inertia are those of $\mathrm{rc}(\pi_w)$
otherwise	nothing beyond the dimension
places $w \mid l$; the second and third rows are for every weak local Langlands map rc	
π_w spherical	crystalline, Hodge–Tate weights $(\mathrm{HT}_\tau)_\tau$
$\mathrm{rc}(\pi_w)$ a sum of characters, $N = 0$	potentially semistable, Hodge–Tate weights $(\mathrm{HT}_\tau)_\tau$; its Weil–Deligne representation is $\mathrm{rc}(\pi_w)$
$\mathrm{rc}(\pi_w)$ satisfies the conditions of the tame row	potentially semistable, Hodge–Tate weights $(\mathrm{HT}_\tau)_\tau$; its Weil–Deligne representation has $N = 0$, and its characteristic polynomials of inertia are those of $\mathrm{rc}(\pi_w)$
otherwise	nothing

FIGURE 11.1. The local clauses (2)(b) and (2)(c) of Definition 11.1.3: in each row, what is asserted when the condition on the left holds. Conditions can hold together, and then all their assertions apply. Purity, and the monodromy operator outside the case $N = 0$, are not asserted at any place. For comparison, the Galois representations of the literature satisfy, under the hypotheses of the papers cited, local–global compatibility including the monodromy operator up to Frobenius semisimplification, at every place (Remark 11.1.5).

(3) The Eichler–Shimura relation is used only as an annihilation statement: a polynomial whose roots are the Satake parameters kills the Frobenius.
(4) An a-dimensional piece is cut out of the cohomology by a statement on the Sen operator together with Proposition 11.2.3, using a set of Hodge–Tate weights that is lacunary (Definition 11.2.1).
(5) Crystallinity comes from a crystalline comparison theorem for all cohomology degrees of smooth projective models.
(6) General forms on the definite group are reached by l-adic approximation and limits of pseudo-characters. Even rank is reached from odd rank by theta lifts, which occur in the construction only for this step. The passage from GL_a to the definite unitary group is a descent by a twisted trace formula, and general CM fields are reached by base-change patching.

□

The frame (descent by the trace formula, limits of pseudo-characters, base-change patching) is that of the literature. The departures are steps (2), (3)–(4) and the use of theta lifts in step (6). As far as we know, the published constructions compute the trace of Frobenius on the cohomology by counting points, in the manner of Langlands and Kottwitz [**Kot92, HT01, Shi11**], which gives the whole characteristic polynomial of Frobenius. Here only an annihilating polynomial is available, and Step (4) supplies what is missing: the Eichler–Shimura relation gives the eigenvalues without their multiplicities, and the Sen operator with Proposition 11.2.3 recovers an a-dimensional representation from them. Figure 11.2 shows the steps.



Plain boxes: the published frame. Dashed boxes: the three departures, namely step (2), steps (3)–(4) together, and the theta lifts in step (6).

FIGURE 11.2. The construction behind Theorem 11.1.4. Plain boxes follow the published constructions, as far as we know; dashed boxes mark the three departures from them (step (2), steps (3)–(4), and the theta lifts in step (6); Remark 11.1.5). Step (4) uses Proposition 11.2.3.

REMARK 11.1.5 (the Galois inputs are proved in the argument). The argument does not take two separate Galois-theoretic inputs from the literature as assumptions, namely Caraiani's local–global compatibility and the existence of Galois representations attached to polarized cuspidal representations. Both places need only truncated local–global compatibility, the same statement, and Theorem 11.1.4 supplies it, so no further hypothesis enters (Alpöge's claim; see Remark 1.7.1). The third input of this kind, the existence of weak local Langlands maps, is also part of Theorem 11.1.4.

Truncated compatibility is weaker than the known theorems. For a regular algebraic polarized cuspidal π over a CM field, the Galois representations constructed in [**HT01, Shi11, CH13**] satisfy, under the hypotheses of the papers cited, local–global compatibility including the monodromy operator up to Frobenius semisimplification, at the places not above l [**TY07, Car12**] and at the places above l [**Car14**], [**BLGGT14**, Thm. 2.1.1]. Truncated compatibility asks for less: no monodromy beyond the case $N = 0$, no purity, and at each place only the clauses of Definition 11.1.3. Since clauses (2)(b) and (2)(c) are asserted for every weak local Langlands map, they say something only where the laws determine the map. For example, if $\mathrm{rc}(\pi_w)$ is a sum of characters, then by laws (a) and (b) the cuspidal support of π_w consists of characters, so π_w is good, and on good representations two weak local Langlands maps agree (laws (a) and (b) are used here in the summarised form of Definition 11.1.1). Provided the local Langlands correspondence itself satisfies the laws, the clauses of Definition 11.1.3 are therefore consequences of the theorems just cited. Under that proviso, truncated compatibility, as a hypothesis, is implied by the literature. It is also what the application needs, for the abelian CM host fields of Chapter 5.

11.2. A proposition on compact groups

DEFINITION 11.2.1 (lacunary sets and product-generic polynomials). Let K be a field of characteristic zero.

(1) A finite family $(e_1, \dots, e_n)$ of elements of K is *lacunary* if no two disjoint nonempty sets $A, B \subset \{1, \dots, n\}$ satisfy $|B| \sum_{i \in A} e_i = |A| \sum_{i \in B} e_i$.
(2) A monic polynomial $P \in K[X]$ of degree n is *product-generic* if it is separable and the product of its roots is not the product of any other multiset of n of its roots, taken with repetition.

Equivalently, a family is lacunary if no two disjoint nonempty subfamilies have equal averages.

DEFINITION 11.2.2 (tangent to the Zariski closure). Let K be a field of characteristic zero, Γ a group, $\rho\colon \Gamma \to \mathrm{GL}_N(K)$ a representation and K' an extension of K. A matrix $X \in M_N(K')$ is *tangent to the Zariski closure* of ρ if for every polynomial function f on M_N over K that vanishes on $\rho(\Gamma)$, the differential of f at the identity vanishes on X.

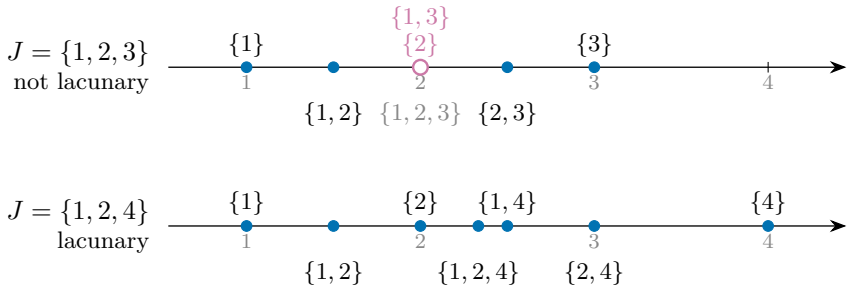


FIGURE 11.3. The averages of the nonempty subsets of J, for the two sets of Example 11.2.4. For $J = \{1,2,3\}$ the disjoint subsets $\{2\}$ and $\{1,3\}$ have the same average 2, so J is not lacunary (Definition 11.2.1); the subset $\{1,2,3\}$ also has average 2, but it meets both. For $J = \{1,2,4\}$ the seven averages are distinct, so J is lacunary.

PROPOSITION 11.2.3 (Alpöge). *Let Γ be a compact topological group, K an algebraically closed field of characteristic zero with a Hausdorff ring topology, K' an algebraically closed extension of K, and $N, n \geq 1$ integers. Let $\rho\colon \Gamma \to \mathrm{GL}_N(K)$ be a continuous semisimple representation. Let $S \subset \Gamma$ be dense, and for each $s \in S$ let $P_s \in K[X]$ be monic of degree n with $P_s(\rho(s)) = 0$. Let $X \in M_N(K')$ be tangent to the Zariski closure of ρ, and suppose that X has n distinct eigenvalues that form a lacunary family. Let $\psi\colon \Gamma \to K^\times$ be a continuous character with $\psi(s) = (-1)^n P_s(0)$ for $s \in S$. Then there is a continuous semisimple representation $r\colon \Gamma \to \mathrm{GL}_n(K)$ such that:*

- *(a) r is a subrepresentation of ρ, and ρ is isomorphic to a subrepresentation of $r^{\oplus k}$ for some $k \geq 1$;*
- *(b) for every $g \in \Gamma$ the characteristic polynomial of $r(g)$ kills $\rho(g)$, and $P_s(r(s)) = 0$ for every $s \in S$;*
- *(c) $\det r = \psi$;*
- *(d) the set of $g \in \Gamma$ at which the characteristic polynomial of $r(g)$ is separable is open and dense;*
- *(e) for $s \in S$, the characteristic polynomial of $r(s)$ equals P_s whenever the characteristic polynomial of $r(s)$ is separable, and whenever P_s is product-generic.*

The statement is elementary: it concerns compact groups, finite-dimensional representations and Zariski closures only, and it can be checked independently of the rest of the argument. In the application Γ is the image of a Galois group, ρ is a piece of the cohomology of step (1), the polynomials P_s come from the Eichler–Shimura relation of step (3), and X is the Sen operator, whose eigenvalues are the Hodge–Tate weights.

On the proof. We do not outline the proof. Part (a) is the main content of the statement: every $\rho(s)$ has at most n distinct eigenvalues, and the conclusion is that ρ is built, up to multiplicity, from a single representation of dimension n. The lacunary eigenvalues of X are the hypothesis that excludes the other ways in which the eigenvalues could be constrained. Parts (b) to (e) record the properties of r that step (4) of the proof of Theorem 11.1.4 uses. □

Example 11.2.4 (lacunary sets). The set $\{1, 2, 3\}$ is not lacunary: $\{1, 3\}$ and $\{2\}$ both have average 2. More generally an arithmetic progression $a, a + \delta, a + 2\delta$ is never lacunary, since $\{a, a + 2\delta\}$ and $\{a + \delta\}$ have the same average. The sets $\{1, 2, 4\}$, $\{1, 10, 100\}$ and $\{1, 3, 9, 27\}$ are lacunary. Table 11.1 records an exhaustive check for these and a few more small sets, and Figure 11.3 shows the subset averages of the first two.

set J	lacunary?	witness (two disjoint subsets with equal averages)
$\{1, 10, 100\}$	yes	—
$\{1, 2, 4\}$	yes	—
$\{1, 2, 3\}$	no	$\{1, 3\}$ and $\{2\}$, average 2
$\{5, 6, 7\}$	no	$\{5, 7\}$ and $\{6\}$, average 6
$\{0, 1, 3, 9\}$	yes	—
$\{1, 2, 4, 8\}$	yes	—
$\{1, 3, 9, 27\}$	yes	—
$\{1, 2, 5, 13\}$	yes	—

Table 11.1. A finite set J is lacunary if no two disjoint nonempty subsets of J have equal averages (Definition 11.2.1), the condition on the Hodge–Tate weights in Proposition 11.2.3. Exhaustive check for small sets. Arithmetic progressions of length at least 3 are never lacunary.

11.3. What the Sha bound uses

Remark 11.3.1 (the Sha bound rests on Theorem 11.1.4). The only input of Chapter 10 beyond classical Galois cohomology and the theorems stated there is Theorem 11.1.4(b). Truncated local–global compatibility (Definition 11.1.3) supplies the Galois representations attached to regular algebraic polarized cuspidal representations of GL_3 and GL_4 over imaginary quadratic fields, together with the local properties that make them points of the ordinary deformation rings in the proofs of Theorems 10.4.2, 10.4.3 and 10.4.9. Read the other way, the argument of Chapter 10 shows that truncated local–global compatibility implies the Sha bound on windows; a base change statement used along the way is also part of Alpöge's argument.

CHAPTER 12

Existence of families

This chapter states the existence results that the Chebyshev inequality on windows of Chapter 9 needs, and deduces from them the zero-freeness of $L(s,\psi)$ for primitive ψ in $\operatorname{Re} s > 1/2$ (Theorem 12.4.1). The existence results are the supply of hosts (Theorem 12.1.2), Green's parametrization (Theorem 12.2.1), the existence of packages and of admissible-family data (Theorem 12.3.4) and the existence of tied families (Theorem 12.3.5); their proofs are given in outline or by a list of inputs, and the deduction of Theorem 12.4.1 is proved in full. All five theorems are black boxes in the proof of the main theorem in Chapter 3; Section 12.5 says which step uses which.

Chapters 6–9 reduce the zero-freeness of the $L(s,\psi)$ to two kinds of existence. First, for each primitive ψ there must be host fields at every scale of a system of windows, meeting the explicit-formula condition and carrying a small Sha budget. Second, there must be families of packages (Definition 8.1.1) tied at every large scale of their windows (Definition 9.1.4), so that the Chebyshev inequality on windows (Theorem 9.4.1) applies. The chain of implications that proves Theorem 3.1.1 is assembled in Section 3.12 of Chapter 3. The mathematical content is in the results quoted; the assembly itself is short (Remark 3.12.2).

12.1. The supply of hosts

DEFINITION 12.1.1 (host supply; schematic). Let $\lambda > 1$.

(1) A *host requirement* $\mathscr{R}$ is a property of a scale (x,ε,J) (Definition 5.3.1), a frame (Definition 5.1.1), a host $\mathcal{H}$ for that frame (Definition 5.1.2), a finite set $\mathcal{T} \subset A$, where A is the group of the frame, and a positive integer t.

(2) Let (x,ε,J) be a scale, with $b = \exp(x^{\varepsilon})$, let $\mathbf{k} = (k,m,t)$ be ranks (Definition 7.1.1), let C_1, C_2, C_3, C_4 be real numbers and n_0 a natural number. A host $\mathcal{H}$ for $\mathbf{k}$, with auxiliary primes q, d and field $E \supset K = \mathbb{Q}(\sqrt{-d})$, is *good at* (x,ε,J) for these constants if it satisfies size conditions at this scale; it has norm primes ϖ_0, ϖ_1 (Definition 5.3.2); it has primes that split completely in E; the characters $\chi_q = (\cdot/q)$, $\chi_d = (\cdot/d)$ and $\chi_q\chi_d$ satisfy pointwise bounds at scale x and have zero-free boxes; $|\operatorname{Re} L'(1,\chi_{-d})/L(1,\chi_{-d})| \le x^{\varepsilon}$; the thin-scale bound of Definition 6.2.1 holds; and the number of bad Frobenius classes is bounded.

(3) The host requirement $\mathscr{R}$ is *supplied on windows of exponent* λ if for every $\varepsilon \in (0, 1/52)$ there are real numbers $C_1, \dots, C_5, x_0$ with $C_1 \geq 1$ and $C_2, C_3, C_4 > 0$, and a system of scales of exponent λ (Definition 5.3.1), such that for every scale (x, ε, J) of the system with $x \geq x_0$ there are ranks $\mathbf{k} = (k, m, t)$ valid at this scale (Definition 7.1.1), a host $\mathcal{H}$ for $\mathbf{k}$ with constant C_5, with top places $\mathcal{T}_{\mathcal{H}}$, and a natural number n_0, such that $\mathcal{H}$ is good at (x, ε, J) for C_1, C_2, C_3, C_4, n_0 and $\mathscr{R}$ holds at the scale, the frame of $\mathbf{k}$, $\mathcal{H}$, $\mathcal{T}_{\mathcal{H}}$ and t.

(4) For a primitive Dirichlet character ψ, the *requirement of* ψ is the host requirement $\mathscr{R}_\psi$ that holds at a scale (x, ε, J), a frame, a host $\mathcal{H}$ with field E, a set $\mathcal{T}$ and an integer t if $\mathcal{H}$ satisfies the explicit-formula condition $\mathrm{EF}(\mathscr{L}_\psi, \mathcal{T}, x, \varepsilon, 4)$ of Definition 5.3.3, for the cyclic level $\mathscr{L}_\psi$ (Definition 4.4.1), and E has Sha budget $(t, b^{1/2})$ (Definition 10.1.2), with $b = \exp(x^\varepsilon)$.

Clause (2) is descriptive. The size conditions of a good host, the finer size conditions on ϖ_0, ϖ_1, the bounds and boxes for χ_q, χ_d and $\chi_q\chi_d$, the bound on bad Frobenius classes, the validity of ranks at a scale and the constant C_5 of a host for given ranks depend on the constants and on n_0; clause (2) is schematic in the sense of Section 4.1, and these conditions are not written out here. In $\mathscr{R}_\psi$ the Sha budget is part of the requirement, and this is how the Sha bound of Chapter 10 reaches the hosts.

THEOREM 12.1.2 (Alpöge). *Let ψ be a primitive Dirichlet character. If ψ has Sha bounds on windows with constant* 401 *(Definition 10.1.2), then $\mathscr{R}_\psi$ is supplied on windows of exponent* $3/2$.

PROOF (OUTLINE). We give Alpöge's route at the level of its main steps. Fix $\varepsilon \in (0, 1/52)$. The host field is the field of Theorem 5.2.4, with a frame of large exponent and twist form for ψ. In each window the auxiliary primes (q, d) are a good pair, which Theorem 5.4.8 supplies. A good pair bounds the bounded twists (Proposition 5.4.9), and this gives the explicit-formula condition $\mathrm{EF}(\mathscr{L}_\psi, \mathcal{T}, x, \varepsilon, 4)$ in $\mathscr{R}_\psi$. The Sha budget in $\mathscr{R}_\psi$ is the hypothesis on ψ.

The two windowings are compatible, because one pair (q, d) serves both. Theorem 5.4.8 with $M = N_\psi$, $\Delta = J$ and $Y = X^{1/J}$, for large X, gives a window pair (q, d) for $(\varepsilon, X^{1/J})$ that is good for $(N_\psi, \varepsilon, X^{1/J}, X)$. For this pair the Sha bound on windows gives a window $[Y', Y'^2] \subset [X^{1/J}, X]$. At every $x \in [Y', Y'^2]$ the pair is good for $(N_\psi, \varepsilon, X^{1/J}, x)$, since $X^{1/J} \leq x \leq X$, so Proposition 5.4.9 applies at x; the frame exponent of Theorem 5.2.4 is taken to be $\nu(x)$, adapted to (ε, x) in the sense of Definition 10.1.2. The Sha bound gives windows $[Y, Y^2]$, and the theorem is stated with the exponent $3/2$; the step between the two is not reconstructed here (Question 12.1.3, Figure 4.4). □

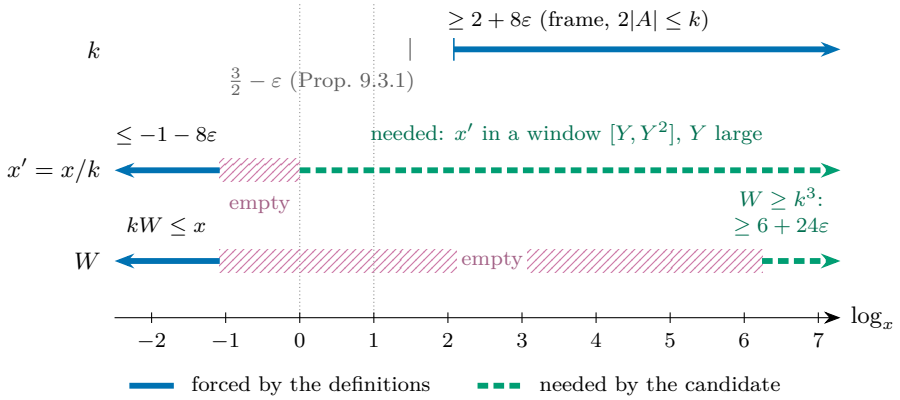


FIGURE 12.1. The rescaling candidate of Remark 12.1.4, on the axis of $\log_x$ of each quantity. The frame and the ranks force $k \geq 2x^{2+8\varepsilon}$, which is stronger than the bound $k \geq x^{3/2-\varepsilon}$ from Proposition 9.3.1. Then $x = kx'$ forces $x' < 1$, and $kW \leq x$ is incompatible with $W \geq k^3$. The constants 2 and 8 are dropped; they only strengthen the bounds drawn. Drawn to scale for $\varepsilon = 0.01$. The figure rules out this candidate only; Question 12.1.3 remains open (Remark 3.14.8).

QUESTION 12.1.3 (the window exponent 3/2). By what step do the windows $[Y, Y^2]$ of the Sha bound (Theorem 10.5.1) become the windows $[W, W^{3/2}]$ of the supply of hosts (Theorem 12.1.2)?

One candidate is a rescaling by the rank: the analytic scale x of a host would be k times the scale x' at which the Sha bound is applied, where k is the first entry of the ranks datum, and a window $[W, W^2]$ for x' would give the window $[kW, (kW)^{3/2}]$ for x, since $(kW)^{3/2} \leq W^2$ once $W \geq k^3$. This candidate is inconsistent with the definitions.

REMARK 12.1.4 (rescaling by the rank does not give 3/2). Let (k, m, t) be the ranks datum of a host at the scale x. The group A of the frame surjects onto a group of order 2^ν (Definition 5.1.1), with ν adapted to the scale, so its order $|A|$ is at least $2^\nu \geq x^{2+8\varepsilon}$; and Definition 7.1.1 requires $2|A| \leq k$. Hence $k \geq 2x^{2+8\varepsilon}$ (Figure 12.1). Independently, the hypotheses $xm \leq k$ and $m \geq x^{1/2-\varepsilon}$ of Proposition 9.3.1 give $k \geq x^{3/2-\varepsilon}$. So $x = kx'$ would force $x' \leq x^{-1-8\varepsilon}/2 < 1$, whereas the Sha bound is applied at scales in a window $[Y, Y^2]$ with Y large; and $W \geq k^3$ would force $W \geq 8x^{6+24\varepsilon}$, which is incompatible with $kW \leq x$. The rank k is not bounded independently of x, so it is not a fixed factor either. One observation is elementary: a set with windows of exponent 2 has windows of exponent 3/2 (Definition 4.4.2), because $[Y, Y^{3/2}] \subset [Y, Y^2]$. Whether this is all that

happens, or whether the scale changes between the two statements, is not determined here. Remark 3.14.8 lists it among the open points.

12.2. Green's parametrization

THEOREM 12.2.1 (Green [**Gre55**]). *Let p be a prime and $n \geq 1$. There is a map $P \mapsto \chi_P$ from the monic irreducible polynomials $P \in \mathbb{F}_p[X]$ of degree n with $P(0) \neq 0$ to the cuspidal irreducible characters of $\mathrm{GL}_n(\mathbb{F}_p)$ such that*

(a) $\chi_P(1) = \prod_{i=1}^{n-1}(p^i - 1)$ for every such P;
(b) $P \mapsto \chi_P$ is injective;
(c) every cuspidal irreducible character of $\mathrm{GL}_n(\mathbb{F}_p)$ is χ_P for some such P.

The map depends on a choice of labels: an identification of the characters of $\mathbb{F}_{p^n}^\times$ with $\mathbb{F}_{p^n}^\times$, under which the Frobenius orbits of regular characters correspond to the roots of the polynomials P.

EXAMPLE 12.2.2. For $p = 3$ and $n = 2$ the monic irreducible quadratics over $\mathbb{F}_3$ are X^2+1, X^2+X+2 and X^2+2X+2. Hence $\mathrm{GL}_2(\mathbb{F}_3)$ has three cuspidal irreducible characters, each of degree $p-1 = 2$. Table 12.1 gives the counts for small p and n.

p	n	$\#\{\text{monic irreducible } P, \deg P = n, P(0) \neq 0\}$	$\chi_P(1) = \prod_{i=1}^{n-1}(p^i-1)$
2	1	1	1
2	2	1	1
2	3	2	3
2	4	3	21
3	1	2	1
3	2	3	2
3	3	8	16
3	4	18	416
5	1	4	1
5	2	10	4
5	3	40	96
5	4	150	11904

TABLE 12.1. Cuspidal irreducible characters of $\mathrm{GL}_n(\mathbb{F}_p)$ correspond to monic irreducible polynomials P of degree n over $\mathbb{F}_p$ with $P(0) \neq 0$ (Theorem 12.2.1). The table gives their number, by the Möbius formula, and the degree $\chi_P(1) = \prod_{i=1}^{n-1}(p^i-1)$.

REMARK 12.2.3 (where Green's theorem and the exponent 3/2 enter). Theorem 12.2.1 and the exponent 3/2 appear in adjacent steps of the proof of Theorem 3.1.1. The exponent does not come from Green's theorem, and

Green's theorem is not an analytic input. Green's parametrization enters only through local types at the norm primes ϖ_0, ϖ_1: it is used to show that the rank defect of each member is small (the irreducibility link, clause (2) of Definition 12.3.1) and to label the local types in the ties. The exponent $3/2$ is the window exponent of Theorem 12.1.2; its relation to the exponent 2 of the Sha bound on windows is Question 12.1.3.

12.3. Admissible families

DEFINITION 12.3.1 (the links; schematic). Let $\mathfrak{F}$ be a family at ε. For a member $\mathcal{P}$ at x with size datum Z, coefficient field L, ranks (k, m, t) with derived quantity f (Definition 7.1.1) and lift index j, put $b = \exp(x^{\varepsilon})$, let B_0 and N_c be the size quantities of Z (Definition 7.2.4), and let

$$\Sigma_j = \sum_{\mathfrak{p}} e_{\mathfrak{p}} \, \frac{\log N\mathfrak{p}}{[L : \mathbb{Q}]} \, \eta_j(\mathfrak{p}),$$

the sum over the places $\mathfrak{p}$ of the support of Z at which the local type of $\mathcal{P}$ is twisted, where $\eta_j(\mathfrak{p}) \geq 0$ is the congruence invariant of $\mathcal{P}$ at $\mathfrak{p}$ and lift index j (Definition 7.2.3). For a prime l put $v(l) = \log l \sum_{1 \leq e < t} \mathrm{len}(l, e)$, with $\mathrm{len}(l, e)$ the lengths of Definition 7.2.3. Let $D_0 \geq 1$ be the rank count of $\mathcal{P}$. The *links* of $\mathfrak{F}$ are the following statements, each for all large x and every member at x.

(1) (Near-orthogonality.) The theta kernels of distinct words are nearly orthogonal for the Hilbert–Schmidt pairing, and the kernel of each word is nearly an isometry, with relative defects at most $\exp(-x^{\varepsilon/8})$.

(2) (Irreducibility.) The rank defect of each member is nonnegative and at most $\exp(-x^{e_z/7})$, where $e_z > 0$ is an exponent attached to the family.

(3) (Congruence.) For every lift index j,

$$\Sigma_j \leq D_0 \Big(C(f m b \, x^{2\varepsilon} + N_c \log^2 x) + \sum_{l} v(l) \Big),$$

for a constant C of the family.

(4) (Divisibility.) For every lift index j, $\Sigma_j \leq D_0 \, C x^{\varepsilon} (B_0 + N_c)$.

(5) (Global.) The chosen data and the inputs of clauses (1)–(4), the upper link (Definition 7.5.1) and the lower link (Definition 7.6.1) imply that every large $x \in \mathrm{Win}$ is tied (Definition 9.1.4).

The rank count D_0 is an integer attached to the package. The precise form of the near-orthogonality in (1), and the further hypotheses of (5), are schematic (Section 4.1). The divisibility bound (4) is a hypothesis of the upper link.

DEFINITION 12.3.2 (admissible-family data and the top existence statement; schematic). Let $\lambda > 1$, let $\mathscr{R}$ be a host requirement (Definition 12.1.1) and let $\mathcal{C}$ and $\mathcal{A}$ be classes of families (Definition 8.1.1).

(1) A family at ε *satisfies the Sha hypothesis* if there is x_0 such that for every $x \geq x_0$ the host field of every member at x has Sha budget $(t, b^{1/2})$ (Definition 10.1.2), with t the rank parameter and $b = \exp(x^{\varepsilon})$.
(2) The *class of* $\mathscr{R}$ *at* λ, $\mathcal{C}_{\mathscr{R}}(\lambda)$, consists of the families with windows of exponent λ all of whose members beyond the threshold meet $\mathscr{R}$.
(3) The class $\mathcal{C}$ *supplies tied families* if there are real numbers e_D, e_L and $\varepsilon_1 > 0$ such that for every $\varepsilon \in (0, \varepsilon_1)$ some family at ε in $\mathcal{C}$ is tied (Definition 9.1.4) at every large $x \in \mathrm{Win}$, with these e_D, e_L and some constants.
(4) The *top existence statement* is: for every $\lambda > 1$ and every host requirement $\mathscr{R}$, if $\mathscr{R}$ is supplied on windows of exponent λ and every family of $\mathcal{C}_{\mathscr{R}}(\lambda)$ satisfies the Sha hypothesis, then $\mathcal{C}_{\mathscr{R}}(\lambda)$ supplies tied families.
(5) *Admissible-family data* for $\mathcal{A}$ consist of realisation data, constants and a list of propositions about the families of $\mathcal{A}$, in four groups:
    (a) automorphic forms, theta lifts and mass formulas: class number equals mass times index; volume ratios by the Siegel–Weil formula [**Wei65, KR88, Ich04**]; decay of matrix coefficients; integrality of row maps; Hecke compatibility; isotypic decomposition; integrality of periods up to the order of the stabiliser;
    (b) Galois representations and local–global compatibility: twist places, and the congruence link, clause (3) of Definition 12.3.1;
    (c) Arakelov geometry and covolume: propositions obtained from the covolume condition (Definition 8.2.1);
    (d) counting and design: the existence of a large family in the class, using Minkowski's bound on finite subgroups of $\mathrm{GL}_n(\mathbb{Z})$ [**Min87**], applied to $\mathrm{GL}_n(\mathbb{Q})$; sizes of level primes; thin-set statements; near-orthogonality; good local types; level projections.

    The propositions are required to hold as closed properties (Definition 8.1.2) on $\mathcal{A}$ or on subclasses of $\mathcal{A}$ attached to the data; among them, the congruence link is closed on such a subclass.

Item (5) is descriptive: the realisation data, the constants and the statements of the individual propositions are schematic (Section 4.1); the list in (5) names the topics they cover. Minkowski's bound applies to $\mathrm{GL}_n(\mathbb{Q})$ because a finite subgroup of $\mathrm{GL}_n(\mathbb{Q})$ fixes a lattice. The top existence statement does not mention $\mathcal{A}$.

The class $\mathcal{C}_{\mathrm{cov}}$ of Definition 8.2.1 is not one of the classes $\mathcal{C}$, $\mathcal{A}$ of this definition. It is a fixed class, cut out by conditions on the packages of a family, and the covolume condition is asserted on it. The admissible-family data of a class $\mathcal{A}$ are built with the covolume condition as one input (Theorem 12.3.4(c)), and their propositions in group (c) are then required on $\mathcal{A}$ or on its subclasses.

Remark 12.3.3 (where the Sha bound is spent). The Sha bound enters on the archimedean side. It does not enter as a separate correction on the finite side at the twisted places. The divisibility bound, clause (4) of Definition 12.3.1, follows from the Sha hypothesis (Definition 12.3.2(1)), the congruence link and a rank statement, and it is a hypothesis of the upper link (Definition 7.5.1). The Sha budget $b^{1/2}$ is absorbed into the bound $D_0 C x^{\varepsilon}(B_0 + N_c)$ because $b \leq B_0$.

Theorem 12.3.4 (Alpöge). *Let $\varepsilon \in (0, 1/52)$.*

(a) For every scale (x, ε, J) with x sufficiently large there are packages (Definition 7.2.3) whose automorphic forms have the weights, level and local types prescribed at that scale, on the tower of unitary groups of Definition 7.1.2 over a CM field of degree about $x^{2+8\varepsilon}$.

(b) The members of the families built from these packages have points, and for each lift index j the lattice M_j of the covolume condition (Definition 8.2.1) is the lattice of integral cusp forms on $U(W_j)$.

(c) There is a class of families of these packages that carries admissible-family data (Definition 12.3.2).

Informal statement. At every large scale there are automorphic forms on the tower of unitary groups over a CM field of degree about $x^{2+8\varepsilon}$ with the prescribed weights, level and local types. Their integral cusp forms give the lattices M_j of the covolume bound, and families of these packages satisfy the propositions listed in Definition 12.3.2(5).

Idea of proof of (a) and (b). We indicate the standard method. On the definite unitary groups of the tower the forms are algebraic automorphic forms [**Gro99**], so forms with prescribed local components can be sought among functions on finite class sets. The chain of theta kernels (Definition 7.1.4) links the levels of the tower, and its analysis rests on the Siegel–Weil formula [**Wei65, KR88, Ich04**] and Howe duality [**How89, Wal90, GT16**]. □

On the proof of (c). The data are assembled from the following inputs: the separation of functionals by unramified Howe duality [**How89, Wal90, GT16**] and the Siegel–Weil formula; the rationality of theta lifts; a fit of Fourier coefficients; theta integrals, by Mínguez's results [**Mín08**]; row functions, through discrete series and the Casselman–Wallach theory [**Cas89, Wal92**]; the covolume bound (Theorem 8.2.3); the weak local Langlands maps of Theorem 11.1.4; and a bound at the places 2 and 3. □

Informal statement of Theorem 12.3.5. Suppose some class of families carries admissible-family data. Let $\mathscr{R}$ be a host requirement that is supplied on windows of exponent λ, and suppose the host fields of the families built on these hosts satisfy the Sha bound. Then for every small ε there is a family of packages, built on the supplied hosts, that is tied (Definition 9.1.4) at every large scale of its windows, with exponents $e_D = 2$ and $e_L = 0$.

THEOREM 12.3.5 (Alpöge). *Suppose that some class $\mathcal{A}$ of families carries admissible-family data (Definition 12.3.2). Let $\lambda > 1$, and let $\mathscr{R}$ be a host requirement (Definition 12.1.1) that is supplied on windows of exponent λ and such that every family of $\mathcal{C}_{\mathscr{R}}(\lambda)$ satisfies the Sha hypothesis. Then $\mathcal{C}_{\mathscr{R}}(\lambda)$ supplies tied families.*

In other words, the top existence statement of Definition 12.3.2 holds.

PROOF (OUTLINE). We give Alpöge's route at the level of its main steps. Let $\lambda > 1$ and let $\mathscr{R}$ be a host requirement such that $\mathscr{R}$ is supplied on windows of exponent λ and every family of $\mathcal{C}_{\mathscr{R}}(\lambda)$ satisfies the Sha hypothesis. Consider the families of $\mathcal{A}$ that lie in $\mathcal{C}_{\mathscr{R}}(\lambda)$.

The links of Definition 12.3.1, the upper and lower links and the remaining propositions of the data are closed on the subclass attached to the data, so for every small ε one threshold serves for all of them. Green's parametrization (Theorem 12.2.1) enters here, in the closed form of the irreducibility link. The design statement of the data (item (5)(d) of Definition 12.3.2), fed by the hosts with which $\mathscr{R}$ is supplied, gives for every small ε and every threshold a family at ε of the class that is large for it. For that family the global link shows that every large $x \in \mathrm{Win}$ is tied, with $e_D = 2$ and $e_L = 0$. Hence the class supplies tied families.

In terms of the objects of Chapters 7–9, the data encode nine steps:

(1) build the packages and their lattices (Theorem 12.3.4);
(2) form the period map (Remark 7.2.6);
(3) apply the slope inequality (Theorem 7.4.2);
(4) bound the archimedean side (the upper link);
(5) bound the covolume (Theorem 8.2.3);
(6) bound the finite side (the lower link);
(7) balance the two sides (Theorem 9.2.1);
(8) bound the defects at the twisted places by Sha lengths and the Sha hypothesis (Remark 12.3.3);
(9) supply near-orthogonality, good local types, masses and class numbers.

□

REMARK 12.3.6 (quantifiers). Supplying tied families is an existence statement that cannot hold vacuously. The closed properties choose the threshold before the family, and the family produced is large for that threshold. An empty class cannot supply tied families, since the definition asks for a tied family at every small ε. Since the top existence statement does not mention $\mathcal{A}$, Theorem 12.3.5 says that one class carrying admissible-family data suffices.

12.4. Zero-freeness for primitive characters

THEOREM 12.4.1 (Alpöge). *Let $\lambda > 1$. Assume that the top existence statement (Definition 12.3.2) holds and that $\mathscr{R}_\psi$ is supplied on windows of*

exponent λ for every primitive Dirichlet character ψ. Then $L(s,\psi) \neq 0$ for every primitive ψ and every s with $\operatorname{Re} s > 1/2$, except $s = 1$ when ψ is trivial.

PROOF. Fix a primitive ψ, and let $\mathcal{C} = \mathcal{C}_{\mathscr{R}_\psi}(\lambda)$; Figure 12.3 shows the order in which the proof fixes its objects. Every member of a family in $\mathcal{C}$ beyond the threshold meets $\mathscr{R}_\psi$. Hence every family in $\mathcal{C}$ satisfies the Sha hypothesis, because the Sha budget is part of $\mathscr{R}_\psi$, and the uniform explicit-formula condition for $\mathscr{L}_\psi$ (Definition 9.1.3), because the explicit-formula condition is. By assumption $\mathscr{R}_\psi$ is supplied on windows of exponent λ. The top existence statement therefore applies to λ and $\mathscr{R}_\psi$, and $\mathcal{C}$ supplies tied families: there are e_D, e_L and $\varepsilon_1 > 0$ such that for every $\varepsilon \in (0, \varepsilon_1)$ some family at ε in $\mathcal{C}$ is tied at every large $x \in \mathrm{Win}$, with these e_D, e_L and some constants.

Let s satisfy $\operatorname{Re} s > 1/2$, with $s \neq 1$ when ψ is trivial, put $\mu = \min(\operatorname{Re} s - 1/2, 1/2) > 0$, and choose $\varepsilon \in (0, \varepsilon_1)$ with $(e_D + e_L + 7)\varepsilon \leq \mu/2$. By the previous paragraph there is a family $\mathfrak{F}$ at this ε in $\mathcal{C}$ and there are real numbers c_W, c_S, e_K, C_L and x_0 such that $\mathfrak{F}$ is tied with constants $e_D, e_L, c_W, c_S, e_K, C_L$ at every $x \in \mathrm{Win}$ with $x \geq x_0$. By the Chebyshev inequality on windows (Theorem 9.4.1), applied to $\mathscr{L}_\psi$ and $\mathfrak{F}$, there is x_1 such that

$$G_{\mathscr{L}_\psi}(x) \leq x^{1/2+(e_D+e_L+7)\varepsilon} \leq x^{\sigma_0}, \qquad \sigma_0 = \tfrac{1}{2} + \tfrac{\mu}{2},$$

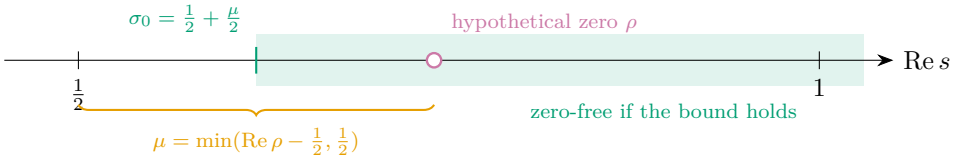
for every $x \in \mathrm{Win}$ with $x \geq x_1$. Since $\mathfrak{F}$ lies in $\mathcal{C}$, the set Win has windows of exponent λ. Hence the set $\{x : G_{\mathscr{L}_\psi}(x) \leq x^{\sigma_0}\}$, which contains every large element of Win, has windows of exponent λ, and $\sigma_0 \in (0,1)$. By the windowed Landau criterion for levels (Theorem 4.4.3), $\mathscr{L}_\psi$ satisfies the window criterion at λ, so $L(s', \chi) \neq 0$ for every $\chi \in \mathrm{Kept}(\mathscr{L}_\psi)$ and every s' with $\operatorname{Re} s' > \sigma_0$, where $s' \neq 1$ when χ is the trivial character. The character ψ is kept by $\mathscr{L}_\psi$, and $\operatorname{Re} s > \sigma_0$ because $\mu \leq \operatorname{Re} s - 1/2$. Hence $L(s,\psi) \neq 0$. □

For the families of Theorem 12.3.5 one has $e_D = 2$ and $e_L = 0$, so $e_D + e_L + 7 = 9$; this is the constant in Figure 12.2.

12.5. What this chapter supplies to the main proof

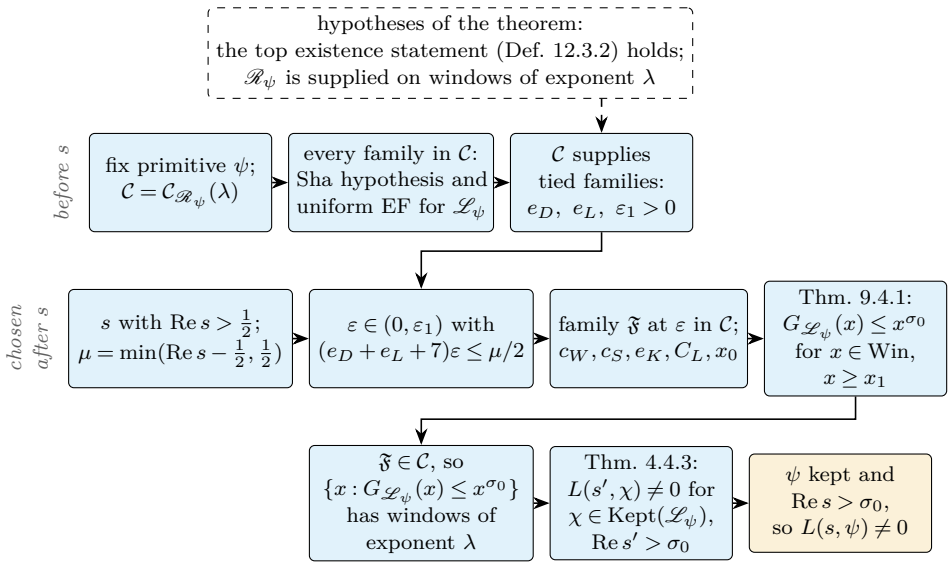
The results of this chapter, together with those of Chapters 8–11, complete the inputs to the proof of Theorem 3.1.1 given in Section 3.12 of Chapter 3. That proof is a chain of implications, and it uses the following statements as black boxes, in this order.

(1) Theorem 11.1.4 (Chapter 11): weak local Langlands maps exist, and truncated local–global compatibility (Definition 11.1.3) holds.
(2) Theorem 10.5.1 (Chapter 10): every primitive Dirichlet character has Sha bounds on windows with constant 401; its automorphy input is the previous item (Remark 11.3.1).



Chebyshev inequality on windows: $G_{\mathscr{L}}(x) \le x^{1/2+(e_D+e_L+7)\varepsilon}$, with $e_D = 2$ and $e_L = 0$ for the families of Theorem 12.3.5, so $x^{1/2+9\varepsilon}$. Choosing $9\varepsilon \le \mu/2$ makes the exponent at most σ_0; the windowed Landau theorem then rules out zeros in the shaded half-plane $\operatorname{Re} s > \sigma_0$, which contains ρ. The closer ρ is to the line, the smaller ε must be. *schematic, not to scale*

FIGURE 12.2. The final step (Theorem 12.4.1). A hypothetical zero ρ with $\operatorname{Re}\rho > 1/2$ is excluded by choosing ε with $9\varepsilon \le \mu/2$ in the Chebyshev inequality on windows (Theorem 9.4.1) and applying the windowed Landau criterion for levels (Theorem 4.4.3). Schematic.



schematic; $\sigma_0 = \frac{1}{2} + \frac{\mu}{2}$

FIGURE 12.3. The order of choices in the proof of Theorem 12.4.1. The exponents e_D, e_L and ε_1 are fixed with ψ, before s is given. The point s determines μ, then ε, then the family $\mathfrak{F}$, its constants and x_1. Dashed boxes are the hypotheses of the theorem. Schematic.

(3) Theorem 12.1.2: given the previous item, $\mathscr{R}_\psi$ is supplied on windows of exponent $3/2$ for every primitive ψ.

(4) Theorem 12.3.4(c): some class of families carries admissible-family data. Its proof uses the covolume bound (Theorem 8.2.3, Chapter 8) and the weak local Langlands maps of the first item. With Green's parametrization (Theorem 12.2.1), Theorem 12.3.5 then gives the top existence statement of Definition 12.3.2.
(5) Theorem 12.4.1 with $\lambda = 3/2$, whose proof applies the Chebyshev inequality on windows (Theorem 9.4.1, Chapter 9) and the windowed Landau criterion for levels (Theorem 4.4.3): $L(s, \psi) \neq 0$ for every primitive ψ and every s with $\operatorname{Re} s > 1/2$, except $s = 1$ when ψ is trivial.

The one remaining step of the proof in Chapter 3, the passage to all Dirichlet characters and to the critical line, uses Proposition 4.5.2 and Theorem 4.5.1 of Chapter 4, not the results of this chapter. The analytic part of the chain is the gap bound (Theorem 9.2.1) with the core bound (Theorem 6.3.4). The constant $e_D + e_L + 7 = 9$ of Figure 12.2 comes from Theorem 12.3.5, which gives $e_D = 2$ and $e_L = 0$.

APPENDIX A

Software and AI methods

At Taylor Dupuy's direction (2026-11-21), all material on the use of software and AI in this book is collected in this appendix, apart from the short discussion that closes the Introduction (page 7). The heading "Software and AI methods" is part of the same instruction. The book has no abstract.

Summary. AI agents assembled and rewrote every chapter of this book on 2026-11-21, in Claude Code [**Ant26a**, product page] (desktop app, version 2.1.328, with subagents), with the models Claude Opus 5.5 (`claude-opus-5-5`) [**Ant26c**, model page] and Claude Fable 5.2 (`claude-fable-5-2`) [**Ant26b**, model page]; the subagents ran at effort setting *medium*, and the coordinator's effort was not recorded. They worked from four sources, all drafted by AI agents on 2026-11-20 and 2026-11-21 and none reviewed by a human: a short account of the idea of the argument (now Chapter 1); an account of the Stepanov–Bombieri proof and an account of the sign-only argument (now Chapter 2); and a survey of Alpöge's Lean archive (now Chapters 3–12). The statement of Theorem 3.1.1 was drafted for that survey by an agent in the Codex desktop app [**Ope26b**, product page] with GPT-6 [**Ope26a**, model guide] on 2026-11-21; its exact variant and reasoning setting were not exposed. All mathematical prose of the book is therefore AI-drafted. The results that originated with the agents are listed in Section A.7, each with the place where the book states it. Chapters 1–3 had one review by an AI agent; Chapters 4–12 had no review as chapters of this book (Figure A.1). No human has reviewed the book, and no agent review is a verification. This appendix was drafted by an AI agent (Section A.6).

Starting material. The following were available before the agent work on the book.

(a) Alpöge's note [**Alp26b**], reproduced after the preface (human-authored, public), and his Lean archive [**Alp26a**, home page], imported unmodified on 2026-11-20 and described by its author as AI-assisted; that description was not checked.

(b) The requests of Taylor Dupuy (human): the title, the remark in the frontmatter, the order of the chapters (the idea first, then the Stepanov–Bombieri proof as an analogy, then the whole proof with black boxes, then the tools), the sign-only argument woven into Chapter 2, the style of lecture notes for professional mathematicians

and graduate students, the exclusion of the project's material on the Deninger program, and one appendix for all AI material.

(c) The four sources (AI, not human-reviewed). The account of the idea was drafted on 2026-11-20 by `claude-opus-5-5` in Claude Code from the agent's earlier notes on the argument (AI), the remark, and Lean statements read as text. The Stepanov–Bombieri account was written on 2026-11-20 by `claude-opus-5-5` from a research report of `claude-fable-5-2` (see Section A.3). The sign-only account was drafted on 2026-11-21 by `claude-opus-5-5` from the text of the remark and the project's notes. The survey was rewritten on 2026-11-21 by `claude-opus-5-5` agents from an earlier, longer report of 2026-11-20, which rested on readings of the archive by `claude-fable-5-2` and on two earlier research notes of the project (AI, `claude-fable-5-2`).

(d) A plan for the book, written on 2026-11-21 by a `claude-fable-5-2` agent and reviewed for wording by a `claude-opus-5-5` agent (AI), and two transcriptions of the remark into LaTeX (AI, one by `claude-fable-5-2` in the plan and one by `claude-opus-5-5`).

(e) Vakil's book [**Vak25**] (human, published), assumed throughout, and the bibliographic data of the sources' bibliographies.

(f) The project's written conventions for style, LaTeX, layout and AI disclosure (drafts; see Section A.2).

A.1. Typesetting

The book is typeset with the class `amsbook` 2.20.6 of the AMS document classes [**Ame20**, CTAN page] and built with pdfTeX 1.40.27 [**T$^+$25**, CTAN page] (TeX Live 2025) through `latexmk` 4.86a [**Col24**, CTAN page]. The figures are those of the sources, unchanged, drawn with TikZ/PGF 3.1.10 [**Tan23**, CTAN page], pgfplots 1.18.1 [**Feu21**, CTAN page] and tikz-cd 1.0 [**Sto21**, CTAN page]. Some builds of the survey used XeTeX [**KH25**, home page].

Skills are named in `typewriter` type. Where a skill is in a public snapshot of the skill repository [**Dup26a**, README], its name links to its copy there. The copies executed were loaded from a local checkout of the workflow repository [**Dup26b**, repository]; no pinned public version of it is recorded, and the executed copies are not shown to equal the snapshot. The skills `finish-writing`, `using-writingpowers`, `check-ste100`, `take-notes`, `create-figure`, `check-result-statement`, `check-definition`, `check-notation` and `check-claims` are not in the public snapshot.

The book was made in three steps on 2026-11-21. First, a plan, with `frontier-dump`: the planning agent read the four sources, compared their preambles, labels, figure paths and bibliographies by script, and built two

throwaway prototypes of the assembly. One prototype dropped spaces between words and was rejected. The plan listed eighteen problems of assembly, among them colliding macros and labels, two different figure files of the same name, and a shift by two chapters in every number of the survey, so that a citation of the survey by number does not match this book. Second, a skeleton: one preamble for the four sources, a merged bibliography, label renames for the collisions, and the remark, merged from the two transcriptions by a comparison of their words with page images made by the Poppler tools [**The26c**, home page]. Third, four writers ported the sources in parallel under `using-latexpowers`, which passed the work to `math-workflow`: Chapters 1 and 2 (`claude-fable-5-2`), Chapter 3 (`claude-fable-5-2`), Chapters 4–7 (`claude-opus-5-5`) and Chapters 8–12 (`claude-opus-5-5`). The writers invoked no further skill. Chapter 3 restates twenty results of the later chapters as black boxes; a script found each restatement identical to its target, before and after revision.

The mechanical checks were `check-labels-and-refs`, `check-citations`, `check-style`, `check-latex`, `check-latex-hygiene` and `check-layout`, run by a `claude-opus-5-5` agent. The citation check failed: most citations of the book lack a pinpoint locator, and many bibliography entries are not cited. The hygiene check asked for revision, the layout check failed in an advisory mode, and the style check passed in an advisory mode; these findings are open. Two prose passes (`claude-opus-5-5`) with `using-writingpowers`, `finish-writing` and `purge-claudisms` changed wording only, in the openings of the chapters, in Chapters 1–3 and in the preface.

The sources had been made as follows. The account of the idea was written in one pass, revised after a review, and given figures by `claude-opus-5-5` agents with `generate-graphics` and `add-figure`. The Stepanov–Bombieri account was harvested from its research report with `harvest-dumps`, revised so that each corrected expectation stands where the claim is met, given figures by `claude-opus-5-5` agents with `create-figure` and `add-figure`, and shortened from 29 to 10 pages at Taylor Dupuy's request; the book also uses the commented-out longer version. The sign-only account was written in three iterations with `take-notes`. The survey was rewritten from the earlier report by a coordinating agent and its subagents: an outline (`rapid-prototype`, `create-exposition`) reviewed with `fp-finder-latex`, five writers, an assembling agent that checked every statement, and a hygiene pass.

A.2. Agent management

The project's written conventions are six contracts, each amended only through its own writer skill: the style contract (`new-repo-style-policy`), the LaTeX contract (`new-repo-latex-policy`), the layout contract (`new-repo-layout-policy`), the AI-disclosure contract

(`new-repo-ai-policy`), the agents file (`new-repo-agents-policy`) and the decision record (`new-repo-adr-policy`). The AI-disclosure, style, layout and LaTeX contracts are drafts that have not been adopted.

The plan was one task: a coordinating session (`claude-opus-5-5`, Claude Code 2.1.330, desktop app) briefed a planning agent (`claude-fable-5-2`, effort *medium*) and a wording reviewer (`claude-opus-5-5`, effort *medium*). The book was a second task: a coordinating session (`claude-opus-5-5`, Claude Code 2.1.328, desktop app) and twelve subagents, all at effort *medium* except the record-keeping agent, whose effort was not recorded. They were the skeleton agent, the four writers, a subagent of the first writer that checked its proofs adversarially, a referee, the agent of the mechanical checks, a revision agent, two prose agents and a record-keeping agent. The writers of Chapters 1–3, the referee and the revision agent were `claude-fable-5-2`; the others were `claude-opus-5-5`. The models are those recorded in every billed message of each agent's transcript. The coordinator set the rulings for the revision and edited two sentences of Chapters 10 and 11 on the status of their results.

The sources were made by other agent teams, described in Section A.1. For the survey: the reading of the archive was one task, in which a coordinating session (`claude-opus-5-5`) briefed a worker (`claude-fable-5-2`) that split the archive into four branches read by four subagents; the earlier report was a second task with 44 agents (`claude-opus-5-5`); the five further readings were five tasks with five `claude-fable-5-2` workers and fourteen subagents; the rewrite was done by a `claude-opus-5-5` coordinator and its subagents.

A.3. Natural language proofs

The skill `frontier-dump` sends a prompt to one model and saves its report; here the model was always `claude-fable-5-2`. It was used four times for the material of this book.

(1) On 2026-11-20, in Claude Code 2.1.326 at effort *xhigh*, for the proof of the Riemann hypothesis for curves in the style of Stepanov and Bombieri, with an analogy to Alpöge's argument. Its result was a research report with a full proof, two variants of the argument, a list of corrected expectations and a script of numerical checks; the Stepanov–Bombieri sections of Chapter 2 are built from it.

(2) On 2026-11-20, for a reading of the Lean archive. Its result was a reconstruction of the proof architecture (a refined Chebyshev method giving a one-sided bound on windows of scales, fed into a windowed Landau criterion), an assumption ledger, a check of the final statements and an outline; the earlier report was built from it with `harvest-dumps`, `create-exposition`, `rapid-prototype` and `fill-in-prototype`.

(3) On 2026-11-20, five further readings, each following one part of the argument below the top level. They gave the deduction from the character of modulus 1 (Remark 4.5.3), the frame theorem behind the slope clause (Theorem 7.4.2), the routes to the plus and minus parts of the Sha bound (Remarks 10.4.4 and 10.3.2), the construction of the Galois representations (Theorem 11.1.4) and the role of arithmetic geometry (Remarks 8.3.1, 8.3.2 and 8.3.3).
(4) On 2026-11-21, at effort *medium*, for the plan of this book (Section A.1).

The proofs of the book were written as follows. In Chapter 1, the proofs of Propositions 1.1.1, 1.2.1, 1.3.2, 1.3.4 and 1.4.2 were written by `claude-opus-5-5` on 2026-11-20 following the standard arguments of the cited sources, which were not reopened; the writer of the chapter (`claude-fable-5-2`) replaced the proof of Proposition 1.1.1 by a sketch citing Chapter 2. In Chapter 2, the proofs of the Stepanov–Bombieri sections were written by `claude-opus-5-5` on 2026-11-20 from the research report of item (1), and the writer of the chapter restored eleven statements from the longer version of that account. The proofs of Propositions 2.9.1 and 2.10.1 and of Theorem 2.10.2 are reconstructions by `claude-opus-5-5` (2026-11-21) of the sketch in Alpöge's note; hypothesis (H2) of Theorem 2.10.2 is not proved (Question 2.10.10). The proofs and outlines of Chapters 3–12 are those of the survey, written by `claude-opus-5-5` with `fill-in-prototype` and revised in the rewrite; the proof of Theorem 3.1.1 in Section 3.12 was expanded by the writer of Chapter 3 (`claude-fable-5-2`), each step with the black box it calls.

Outside the scope of this book, and not withdrawn: six earlier findings of the project on the conjecture of Birch and Swinnerton-Dyer, and a block of 250 refutations concerning the Deninger program. None is a premise of the argument of this book.

A.4. Adversarial reviews

For the book, one referee review with `referee-report`, by a `claude-fable-5-2` agent that had not drafted the text, covered Chapters 1–3 and asked for revision. A `claude-fable-5-2` subagent of the writer of Chapters 1 and 2 checked the sign-only sections of Chapter 2 adversarially, following the prompt of the skill `doubt` without invoking it. Chapters 4–12 had no referee review as chapters of this book.

The sources had these reviews. The account of the idea had one review with `critical-review` by a `claude-opus-5-5` subagent. The Stepanov–Bombieri account had one review with `fp-finder-latex` by a `claude-opus-5-5` agent, the same model as the writer, which asked for minor revision and made the revisions itself, with no return to the referee. The sign-only account had no review. The earlier report behind the survey had seven simulated referee reviews with `fp-finder-latex`, two reviews of its

Part of the book	Source (drafter, date)	Review of the source	Writer	Review in the book's run
Chapter 1	account of the idea (`opus`, 2026-11-20)	`critical-review` (`opus`)	`fable`	referee (`fable`)
Chapter 2, Stepanov–Bombieri	Stepanov–Bombieri account (`opus`, 2026-11-20, from a `fable` research report)	`fp-finder-latex` (`opus`); revised by the reviewer	`fable`	referee (`fable`)
Chapter 2, sign-only	sign-only account (`opus`, 2026-11-21)	*none*	`fable`	referee (`fable`); adversarial check by a subagent (`fable`)
Chapter 3	survey*	survey reviews†	`fable`	referee (`fable`)
Chapters 4–12	survey*	survey reviews†	`opus`	*none*
Appendix A	the project's records	–	`opus`	audit by its own drafter

*Rewritten by `opus` agents on 2026-11-21 from an earlier report of 2026-11-20, which rested on readings of the archive by `fable` agents; the statement of Theorem 3.1.1 was drafted for it by an agent with GPT-6. †Of the earlier report: seven simulated referee reviews with `fp-finder-latex`, and further reviews and checks. Of the rewrite: a review of its outline and an integrated review of its body.

FIGURE A.1. Who wrote and who reviewed each part of the book. Model names omit the prefix `claude-` and the version. All reviews were by agents of the same model family as the drafters, and no human has reviewed any part (Sections A.1–A.4 and A.6).

figures, a review of its introduction with `critical-review` and checks of its proved propositions with `doubt`; the rewrite had a review of its outline and an integrated review of its body, which asked for a major revision and reported no mathematical error in a proved statement. All these reviews were by agents of the same model family as the drafters.

A.5. AI assisted computations

No computer algebra system was run for the book, and no Lean code was run. The book's computations are Python scripts written and run by the agents in 2026-11-21 (Python version not recorded): the comparison of the black-box restatements of Chapter 3 with their targets; the values of hypothesis (H2) in characteristic 2 and at $k = 4, 6, 7$ for odd $q \leq 29$ (Remark 2.10.7); the factor formula of Remark 2.9.2; the orbit count of Remark 2.10.9; and the ratios of Remark 2.7.6. The adversarial subagent of Section A.4 checked the identity $qN = M^k + R$ and the bound of Proposition 2.9.1 numerically for $q = 5, 7, 9$ and $k = 2, \ldots, 5$. These are outputs of the scripts; they test small cases and are not proofs.

For the Stepanov–Bombieri account, a Python 3.13.7 script [**Pyt25**, release page] using only the standard library checked the parameter inequalities, the constructions on small explicit curves, point counts and L-polynomials by brute force, and some identities on random inputs; the reviewer checked one identity numerically over $\mathbb{F}_3$ and $\mathbb{F}_9$. For the survey, the counts of the static scans (Remark 3.14.4) are outputs of a ripgrep 15.1.0 script [**The25a**, release 15.1.0] and of Python 3.13.7 scripts, written and run by the reading agents (`claude-fable-5-2`) and not recomputed. The data of the figures and example tables of Chapters 3–12 come from Python scripts written and run by the figure agents (`claude-opus-5-5`), some with exact rational arithmetic and two with mpmath 1.3.0 [**The23**, home page]. The zeros of ζ in Figures 4.1 and 4.3 come from a script run in SageMath 10.8 [**The25b**, home page] (NumPy 2.3.2 [**HMvdW$^+$20**, DOI] as Sage reported it), with byte-identical output on a rerun. On 2026-11-21 an agent of the rewrite reran the scripts behind Examples 6.1.6, 10.2.5, 11.2.4 and 12.2.2 and Figure 6.3, with byte-identical outputs, and the value $L'(1,\chi_{-7})/L(1,\chi_{-7}) = 0.015635689993720378957$ was reproduced independently twice with mpmath. The inequalities of Proposition 9.3.1 were derived by hand by an agent and re-derived by a reviewer. For the account of the idea and the sign-only account no machine computation was made; their parameter balances and the bound at $k = 3$ were checked by hand by the agents. These computations illustrate the text; no result of the book rests on them.

A.6. AI statement generation

This appendix was drafted on 2026-11-21 by a `claude-opus-5-5` subagent in Claude Code (harness version and effort not recorded) with `write-materials-and-methods`, from the project's provenance and token records, the logs of the book's run and the disclosures of the four sources. The record-writing skills `update-ai-usage` and `update-tokens` were only consulted. The same agent then audited it with the read-only skill `latex-ai-statement`; that audit is by the drafting model and is not independent. The AI notice of the preface was drafted by the skeleton agent (`claude-opus-5-5`) and revised by a prose agent; the AI and Software Discussion of the Introduction was drafted by the introduction agent (`claude-fable-5-2`). The disclosures of the four sources were not ported into the chapters; their content is in this appendix. No human has reviewed or accepted this appendix. The author's statement of responsibility for the book has not yet been written.

A.7. Proofreading and adversarial review

The findings of the referee of the book were applied in a revision pass with `revise` (`claude-fable-5-2`), under the coordinator's rulings, with `track-down-reference` for the references. The following results of the book

originated with the agents in the book's run. None has been checked by a human; each entry names its origin and how far it was checked.

(i) Hypothesis (H2) of Theorem 2.10.2 fails for the path quadric at $k = 3$ (Remark 2.10.7). Found by the adversarial subagent and written by the writer of Chapter 2. The criterion through the Künneth formula was added in the revision. Proved in the text. The computed failures at $k = 4, 6, 7$ and in characteristic 2 in the same remark are numerical evidence only.

(ii) The hyperbolic reading of the form, with its bound of order $q^{3k/4}$ (Remark 2.9.2), and the torus-orbit remark (Remark 2.10.9); from the referee, written in the revision, proved in the text and checked by script.

(iii) The exclusion of a real zero by the sign of its residue in the proof of Proposition 1.3.4; from the referee, written in the revision.

(iv) Rescaling the window exponent by the rank is inconsistent with Definitions 5.1.1 and 7.1.1 and with Proposition 9.3.1 (Remark 12.1.4); from the referee, proved in the revision from the definitions. The step from the exponent 2 to $3/2$ is left open (Question 12.1.3); the archive's formal statement of host supply, which would decide it, was not consulted.

(v) The remarks 2.10.6, 2.10.5, 2.7.6 and 2.3.14; from the writer of Chapter 2 and the revision, proved or computed in the text. Remark 2.3.14 rests on Clifford's theorem, used without proof or locator.

(vi) The $y(c)$ of the remark are theta coefficients only for $n = 3$, not for every n (Section 2.12); from the referee, written in the revision after the cited sources were traced.

(vii) Corrections to the sources, each made in place: the theta exponent $a/p+a/p^3$; Proposition 2.6.5 at $n = 1$; the Parseval saving $q^{(k-1)/2}$; the hypothesis $g \geq 1$ of Theorem 2.6.2; and the restriction of an attribution to Kazhdan and Patterson to $n = 3$; and the observation that the hypothesis $0 \notin U$ of Theorem 2.10.2 is not used. From the writers and the revision, and confirmed by the referee.

The following results came from the sources, with their origin.

(viii) Stepanov–Bombieri account: the corrected range of (2.6.2), from its writer; and Remark 2.3.14 in its first form, from its reviewer. The negative results of Chapter 2 come from the research report of Section A.3, item (1), and its harvest.

(ix) Account of the idea: the agent's first claim that Landau's theorem does the job of Bombieri's lower bound was incomplete; the one-sided criterion works for ζ because ζ has no real zero in $(0, 1)$, and fails as stated over a finite field (Remark 1.1.3).

(x) Sign-only account: the hypothesis $0 \notin U$ of Theorem 2.10.2, added in its second iteration (and found unused in the book, item (vii));

the link between the path quadric and Gabber's hypersurface in Katz's argument [**Kat15**] is the agent's inference.

(xi) Survey (`claude-opus-5-5` unless stated, from the reviews and readings of Sections A.3 and A.4): the proof of the balance identity (Proposition 6.1.4), by orthogonality of characters, checked line by line by two reviewers; the bookkeeping of Proposition 9.3.1; the slope clause is not Bost's inequality but has the form of the conclusion of the frame theorem (Remark 7.4.3, Theorem 7.4.2), weakened after a review; the plus-part route (Theorems 10.4.2 and 10.4.3, Remark 10.4.4); the spread of a slope record is set by no tie (Definition 9.1.4); the covolume condition and its one consumer of arithmetic geometry (Definition 8.2.1, Remark 8.2.2, Theorem 8.2.3, Remark 8.3.1); the cancellations of the upper link (Remarks 9.2.2 and 12.3.3); the minus part and its windows (Remark 10.3.2, Theorem 10.3.1, Theorem 10.4.8, Remark 10.5.2); the route of the Galois construction and its three departures from the published constructions (Theorem 11.1.4, Proposition 11.2.3, Remarks 11.1.2 and 11.1.5); the divisor data (Remarks 8.3.3 and 8.3.2); the standing hypotheses and parity character of the case split (Definition 10.2.2, Proposition 10.2.3, Remark 10.2.4), from a review; the deduction from the character of modulus 1 (Remark 4.5.3); the strength of the arithmetic-geometry inputs and three inputs not used (Remarks 8.3.4 and 10.5.3), from the agents' memory and a textual search, not a term-level dependency check; Remark 12.2.3 and the exponents of Theorem 12.3.5; corrections to Definitions 7.1.1 and 7.1.2 and Remark 7.1.3, made while drawing the figures; the interpretation of the exponent 1/2 (Remark 9.4.3); the hypothesis on the trivial character in Proposition 6.3.2; the size of the modulus in the idea of proof of Proposition 5.4.3; the dependence on Δ of the constant in the route sketched for Proposition 5.4.7; and part (c) of Theorem 12.3.4. The readings that produced these points were by `claude-fable-5-2` agents; the derivations were checked only as stated.

A.8. Literature review

For the book, the revision agent located references with `track-down-reference`: the locators in Vakil's book [**Vak25**] were checked in the public PDF of its version of October 2025; the pages of [**Che10**] against the publisher's page; the publisher of [**Wal90**] from citing papers, not from the volume; and [**Pat77, KP84, CFH12**] only through a secondary source, without opening them. No search for prior occurrences was run for the results (i)–(vi) of Section A.7; their novelty is not claimed, and whether they are known is unresolved.

For the sources: the Stepanov–Bombieri report read Bombieri's Bourbaki exposé [**Bom74**] (pp. 234–240) and Tao's blog post [**Tao14**] directly, through the web tools of Claude Code [**Ant26d**, tools reference], `curl` [**The24**, home page] and a Semantic Scholar search connector, which serves the Semantic Scholar API [**All26**, API page] and refused queries for rate limiting; other references were taken from publisher or index pages or cited from memory. No source was found for Theorems 2.6.3 and 2.6.9 or Propositions 2.6.8 and 2.6.10 in this form, and whether Proposition 2.3.16 is known is unresolved. The sign-only account cited [**Kat15, Poo04, Del74**] from memory and earlier records, without rereading them. For the survey, two `claude-opus-5-5` agents checked the references with `track-down-reference` against zbMATH Open [**Eur26**, home page], Crossref [**Cro26**, REST API], Numdam [**Mat26**, home page] and publishers, and ran bounded searches for its distinctive steps through the arXiv API [**arX26**, API page] in the manner of `search-arxiv`, web search, and the Semantic Scholar API in the manner of `search-scholar`. A search without hits is not a claim of novelty. For the windowed Landau criterion no prior occurrence was found; the closest work [**BR26**, Thm. 3.1] starts from a pointwise bound (Remark 4.3.8, Question 4.3.9). For a Chebyshev or slope inequality through theta lifts, none was found. For a uniform Vandiver-type bound on Sha groups of Tate twists none was found; the related [**Kur92**, §5, p. 232] works over $\mathbb{Q}$, one prime at a time (Remark 10.4.10). For the balance identity, the orthogonality it uses is standard [**Con26**, Cor. 4.2, Ex. 4.3], and no prior occurrence of the identity was found. The comparisons of the plus part, the minus part and the Galois construction with published work were made from the reading agents' memory, except the check of Remark 11.1.5 against [**BLGGT14**, Thm. 2.1.1] in full text and [**Car12, Car14**] in abstract; one point is open, the twist by a power of $|\det|$ in the normalisation of the local Langlands correspondence. Six entries added in the rewrite [**Kaw82, Vie82, Nad90, Mes72, ALB23, Wal36**] were checked as bibliographic records, not for what they say. No human has checked the references of the book.

A.9. Formalization

This book formalizes nothing, and no Lean skill was run for it. Its object is Alpöge's Lean archive [**Alp26a**], on Lean 4 v4.33.1 [**The26a**, release v4.33.1] [**dMU21**, DOI] and Mathlib revision `0df444a3` [**The26b**, commit]. It was read as text and never built. The formal names of the statements are kept only in source comments next to each statement. The readings covered definitions and statements, and proofs only at the top two to four levels of each branch, eight to twelve levels for the construction of the Galois representations; about 175,000 of its 175,187 files were not opened. A build of the main file, its axiom report and an independent replay of the kernel check [**lea23**, repository] would decide whether the final statements

are proved relative to the kernel; static reading cannot. Elaboration runs the archive's meta code, so such a check belongs in a disposable sandbox.

The readings established facts about the archive. Each fact that bears on a statement of the book is stated next to that statement. The final statements of the archive are compared with Theorem 3.1.1 in Remark 3.1.3. The scan for assumptions follows Remark 3.14.3, as Remark 3.14.4. The facts on imports and on transcribed theorems follow Remark 3.14.5, as Remarks 3.14.7 and 3.14.6. One fact concerns the method of reading.

Remark A.9.1 (names). At every level, the names of the files and declarations of the archive do not describe their content.

Further facts found by the readings: the archive contains meta code that adds declarations and defers type checking to the kernel, so that "it elaborates" is no evidence, and it shows signs of mechanical assembly (duplicated declarations, 1727 one-line wrappers); none of this is evidence of an error. It claims in-archive proofs of resolution of singularities in characteristic zero, the arithmetic Hilbert–Samuel theorem, the Siegel–Weil formula, Howe duality and about 190 further named theorems. A Weil–Deligne parameter has default values when none exists, whose branches were not checked to be unreached. No counterexample to a statement of the archive was found, and no ill-posed definition at the level of the final statements.

A.10. Record keeping and token accounting

The project keeps one provenance record and one token record. They are working records and are not distributed with this book. The provenance record, written only with `update-ai-usage`, gives for each task the agents with their models and harness, the skill chain, the inputs and their origin, the results with their derivation, the searches, the reviews and the limits. The token record, written only with `update-tokens`, gives per agent the token counts by class where available, the rates with their date, and the cost, labelled as measured or estimated. "Not recorded" means not measured, not zero.

Table A.1 gives the token use of the tasks that wrote the four sources and the book. Measured rows sum the per-message counts of the agents' transcripts by token class, cache reads included, and price each class at its rate; most of them are lower bounds, because a coordinating session or the last turns of a record-keeping agent were not counted. The estimated rows are carried from the survey's own table: per-agent totals without a split by class, as the harness reported them; they probably exclude cache reads. Their costs price `claude-opus-5-5` totals at the input rate and `claude-fable-5-2` totals with an assumed split of 25% output and 75% cache write, with bounds at a low and a high rate. Costs are token charges in US dollars at the rates per million tokens recorded on 2026-11-20 and checked again on 2026-11-21 on the provider's pricing page, which shows no effective date: for `claude-opus-5-5` input 4.00, output 20.00, cache write

5.00 (5 minutes) and 8.00 (1 hour), cache read 0.20; for `claude-fable-5-2` input 8.00, output 40.00, cache write 10.00 and 16.00, cache read 0.20. Not recorded: the import of the archive, the introduction and Theorem 3.1.1 drafted in the Codex app, the rewrite's coordinator after a point and the last turns of its body revision, the coordinator's last turns in the book's run, and the drafting and audit of this appendix.

Task	Model	Tokens	USD	Source
Chapter 1: account of the idea (2026-11-20)				
Draft	`opus`	3,961,002	2.04	measured
Review and revision	`opus`	9,656,142	3.96	measured
Figures	`opus`	14,149,338	4.63	measured
Figure insertion	`opus`	11,032,067	2.87	measured
Chapter 2: Stepanov–Bombieri (2026-11-20) and sign-only (2026-11-21)				
Research report	both	7,283,238	18.50	measured
Harvest and review	`opus`	18,264,997	9.35	measured
Expectations placed	`opus`	16,291,335	5.70	measured
Figures	`opus`	50,892,392	24.92	measured
Shortening	`opus`	26,162,340	8.31	measured
Figure insertion	`opus`	15,492,962	4.54	measured
Sign-only, iteration 1	`opus`	6,464,552	4.32	measured
Sign-only, iterations 2–3	`opus`	3,945,668	2.21	measured
Chapters 3–12: the survey (2026-11-20 unless stated)				
Import of the archive	`opus`	not recorded	–	–
Note on Birch–Swinnerton-Dyer	both	24,829,392	28.15	measured
Note on the Deninger program	both	66,519,676	55.97	measured
Reading of the archive	`fable`	338,616	5.93	estimated
Further readings 1–5	`fable`	4,851,181	84.90	estimated
Earlier report	`opus`	7,407,733	29.63	estimated
Figures of the earlier report	`opus`	4.36–11.88 M	17.42–47.53	estimated
Rewrite into the survey (2026-11-21)	`opus`	222,670,060	84.01	measured
The book (2026-11-21)				
Transcription of the remark	`opus`	2,320,296	1.25	measured
Plan	both	12,720,908	10.58	measured
Skeleton, chapters, review, revision, prose	both	104,511,334	70.73	measured

TABLE A.1. Token use of the tasks behind this book. Model names omit the prefix `claude-` and the version; "both" means agents of the two models, priced per model by token class. The measured rows sum to 617,167,699 tokens and 342.04 US dollars; most are lower bounds. The estimated rows of the survey sum to 12,597,530 tokens and 120.45 US dollars (bounds 43.00 to 355.75), and its figure pass is a range from 4,355,591 to 11,881,382 tokens, because its forks re-read the coordinator's context. In the book's run, `claude-fable-5-2` agents used 50,581,916 tokens (47.77 US dollars) and `claude-opus-5-5` agents 53,929,418 (22.97 US dollars).

Bibliography

[AB95] Ahmed Abbes and Thierry Bouche, *Théorème de Hilbert–Samuel "arithmétique"*, Annales de l'Institut Fourier **45** (1995), no. 2, 375–401, Zbl 0818.14011; MR 96e:14024.

[ALB23] Johannes Anschütz and Arthur-César Le Bras, *Prismatic Dieudonné theory*, Forum of Mathematics, Pi **11** (2023), Paper No. e2, Zbl 1515.14032; arXiv:1907.10525.

[All26] Allen Institute for AI, *Semantic Scholar academic graph API*, `https://www.semanticscholar.org/product/api`, 2026, Operated by Ai2. Page opened 2026-11-20. Every query made for this report was refused for rate limiting.

[Alp26a] Levent Alpöge, *Lean archive*, Published at `https://alpo.ge/` as `lean.tar.gz.part01`–`part14`, 2026, Version used: the archive as published on alpo.ge and imported locally on 2026-11-20 unmodified (commit 155fd7918658af466e105c6ad58f63b3b3aac119; sha256 of the joined tarball 9631b4f020199e580ba86178cb56c6e9b09f690c2369a80dc4334aa93f49c7fc). Toolchain Lean v4.33.1; Mathlib 0df444a360eaa60ab8c11dca51a86af692955474. Home page title "Levent Alpöge", opened 2026-11-20. Not built or verified.

[Alp26b] ———, *Remark*, `https://alpo.ge/remark.pdf`, 2026, Unsigned four-page note; attribution as published on alpo.ge. PDF creation date 2026-11-20; retrieved 2026-11-21, SHA-256 3fc97784b71ddd44ff8aa0a177e99622a1c61b13cfa843a0989ee4a7a0cfeb8d.

[Ame20] American Mathematical Society, *amscls: AMS document classes for LaTeX; class `amsbook`, version 2.20.6*, `https://ctan.org/pkg/amscls`, 2020, Local `amsbook.cls` reports "2020/05/29 v2.20.6" (TeX Live 2025). CTAN page opened 2026-11-21.

[Ant26a] Anthropic, *Claude Code*, `https://claude.com/claude-code`, 2026, Agentic coding harness. Version 2.1.329 recorded for one subagent of this manuscript task and reported by the local command-line installation on 2026-11-20; desktop app and other sessions of 2026-11-20 (version not recorded); product page opened 2026-11-20.

[Ant26b] ———, *Claude Fable 5.2 (model `claude-fable-5-2`)*, `https://platform.claude.com/docs/en/models/fable-5-2/overview`, 2026, Large language model; model ID confirmed on `https://platform.claude.com/docs/en/models/overview`, opened 2026-11-20.

[Ant26c] ———, *Claude Opus 5.5 (model `claude-opus-5-5`)*, `https://platform.claude.com/docs/en/models/opus-5-5/overview`, 2026, Large language model; model ID confirmed on `https://platform.claude.com/docs/en/models/overview`, opened 2026-11-20.

[Ant26d] ———, *Tools reference, Claude Code documentation*, `https://code.claude.com/docs/en/tools-reference`, 2026, Documents the built-in

tools `WebFetch` (fetches a URL, converts it to Markdown and runs an extraction prompt on it, "lossy by design") and `WebSearch`. Opened 2026-11-20.

[arX26] arXiv, *arXiv API access*, `https://info.arxiv.org/help/api/index.html`, 2026, Public API of the arXiv e-print repository. Page opened 2026-11-20.

[BGKK05] José I. Burgos Gil, Jürg Kramer, and Ulf Kühn, *Arithmetic characteristic classes of automorphic vector bundles*, Documenta Mathematica **10** (2005), 619–716, Zbl 1080.14028; arXiv:math/0502085.

[BH13] Reinier Bröker and Jeff Hoffstein, *Fourier coefficients of sextic theta series*, Preprint, arXiv:1312.0568v1, 2013, Version 1 of 2 December 2013, 30 pages; Section 1 read on 2026-11-21.

[BLGGT14] Thomas Barnet-Lamb, Toby Gee, David Geraghty, and Richard Taylor, *Potential automorphy and change of weight*, Annals of Mathematics (2) **179** (2014), no. 2, 501–609, Zbl 1310.11060; arXiv:1010.2561.

[Bom74] Enrico Bombieri, *Counting points on curves over finite fields (d'après S. A. Stepanov)*, Séminaire Bourbaki, Vol. 1972/73, Exposés 418–435, Lecture Notes in Mathematics, vol. 383, Springer, Berlin, 1974, Exp. No. 430; Numdam SB_1972-1973___15___234_0. Pages 234–240 read from the Numdam scan on 2026-11-20; page 241 not read, pp. 234–241.

[Bos96] Jean-Benoît Bost, *Périodes et isogénies des variétés abéliennes sur les corps de nombres (d'après D. Masser et G. Wüstholz)*, Séminaire Bourbaki, Vol. 1994/95, Exposés 790–804, Astérisque, vol. 237, Société Mathématique de France, Paris, 1996, Exp. No. 795; Zbl 0936.11042; MR 1423622; Numdam SB_1994-1995___37___115_0, full text opened; subtitle as printed on p. 115, pp. 115–161.

[Bos01] ———, *Algebraic leaves of algebraic foliations over number fields*, Publications Mathématiques de l'IHÉS **93** (2001), 161–221, Zbl 1034.14010; MR 1863738; full text opened (Numdam).

[BR26] Fredrik Broucke and Szilárd Gy. Révész, *On Landau's "local method" for obtaining zero-free regions*, `https://arxiv.org/abs/2611.03356`, 2026, arXiv:2611.03356v1, submitted 2026-11-02 (arXiv API record opened 2026-11-20); Thm. 3.1 read in the arXiv HTML version.

[Car12] Ana Caraiani, *Local-global compatibility and the action of monodromy on nearby cycles*, Duke Mathematical Journal **161** (2012), no. 12, 2311–2413, Zbl 1405.22028; arXiv:1010.2188.

[Car14] ———, *Monodromy and local-global compatibility for $l = p$*, Algebra & Number Theory **8** (2014), no. 7, 1597–1646, Zbl 1310.11061; arXiv:1202.4683.

[Cas89] W. Casselman, *Canonical extensions of Harish-Chandra modules to representations of G*, Canadian Journal of Mathematics **41** (1989), no. 3, 385–438, Zbl 0702.22016.

[CFH12] G. Chinta, S. Friedberg, and J. Hoffstein, *Double Dirichlet series and theta functions*, Contributions in Analytic and Algebraic Number Theory: Festschrift for S. J. Patterson, Springer, 2012, Not opened; data from the reference list of arXiv:1312.0568.

[CH13] Gaëtan Chenevier and Michael Harris, *Construction of automorphic Galois representations, II*, Cambridge Journal of Mathematics **1** (2013), no. 1, 53–73, Zbl 1310.11062.

[Che10] Huayi Chen, *Convergence des polygones de Harder–Narasimhan*, Mémoires de la Société Mathématique de France, Nouvelle Série **120** (2010), 1–120, Zbl 1208.14001 gives 112 pp.; Crossref gives pp. 1–120 and the publisher's

page (read 2026-11-21) 120 pages; arXiv:0706.2649; French title as in the Crossref record.

[Clo90] Laurent Clozel, *Motifs et formes automorphes: applications du principe de fonctorialité*, Automorphic Forms, Shimura Varieties, and L-functions, Vol. I (Ann Arbor, MI, 1988), Perspectives in Mathematics, vol. 10, Academic Press, Boston, MA, 1990, Zbl 0705.11029 (English title there); French title and publisher as commonly cited, not checked on a publisher page, pp. 77–159.

[Col93] Pierre Colmez, *Périodes des variétés abéliennes à multiplication complexe*, Annals of Mathematics (2) **138** (1993), no. 3, 625–683, Zbl 0826.14028.

[Col24] John Collins, *Latexmk, version 4.86a*, `https://ctan.org/pkg/latexmk`, 2024, Local `latexmk -v` reports "Latexmk, John Collins, 27 Dec. 2024. Version 4.86a". CTAN page opened 2026-11-20 (current version there 4.89, 2026-11-08; GPL 2; home page `https://www.cantab.net/users/johncollins/latexmk/`).

[Con26] Keith Conrad, *Characters of finite abelian groups*, `https://kconrad.math.uconn.edu/blurbs/grouptheory/charthy.pdf`, 2026, Expository note. Copy opened 2026-11-20 (PDF metadata date 2026-10-20); orthogonality relations in Cor. 4.2, Dirichlet's detection of a residue class in Ex. 4.3.

[Cro26] Crossref, *Crossref REST API*, `https://www.crossref.org/documentation/retrieve-metadata/rest-api/`, 2026, Base URL `https://api.crossref.org/`. Documentation page and `https://www.crossref.org/` opened 2026-11-20.

[Dav00] Harold Davenport, *Multiplicative number theory*, third ed., Graduate Texts in Mathematics, vol. 74, Springer, New York, 2000, Revised and with a preface by Hugh L. Montgomery; Zbl 1002.11001.

[Del74] Pierre Deligne, *La conjecture de Weil. I*, Institut des Hautes Études Scientifiques. Publications Mathématiques **43** (1974), 273–307.

[dMU21] Leonardo de Moura and Sebastian Ullrich, *The Lean 4 theorem prover and programming language*, Automated Deduction – CADE 28, Lecture Notes in Computer Science, vol. 12699, Springer, 2021, Metadata reused from latex/notes.bib key Moura2021lean (verified there 2026-10-14), pp. 625–635.

[DR80] Pierre Deligne and Kenneth A. Ribet, *Values of abelian L-functions at negative integers over totally real fields*, Inventiones Mathematicae **59** (1980), no. 3, 227–286, Zbl 0434.12009.

[Dup26a] Taylor Dupuy, *agent-skills-public: agent skills for research mathematicians*, `https://github.com/tdupu/agent-skills-public/tree/8d13efb5e66306aeae6e3b948bc463e7e2db1b65`, 2026, Public reference snapshot, commit 8d13efb5e66306aeae6e3b948bc463e7e2db1b65 (2026-11-12), containing skills/frontier-dump, skills/harvest-dumps, skills/using-latexpowers, skills/create-exposition and skills/rapid-prototype (opened 2026-11-20). The copies executed on 2026-11-20 were loaded from local checkouts and are not shown to equal this snapshot.

[Dup26b] Taylor Dupuy, *mathcity: mathematical research workflows and policies*, 2026, Source of the frontier-dump, harvest-dumps, using-latexpowers and math-workflow skills. The version executed on 2026-11-20 is the local installation; no pinned public snapshot recorded.

[Eur26] European Mathematical Society, FIZ Karlsruhe and Heidelberg Academy of Sciences, *zbMATH Open*, `https://zbmath.org/`, 2026, Bibliographic database of mathematics, formerly Zentralblatt MATH, open since January

2021. Editors as stated in M. Schubotz, D. Trautwein, O. Teschke, "zbMATH is open: a practical guide to open an information service", Zenodo, 2021, `https://doi.org/10.5281/zenodo.4646668` (opened 2026-11-20); the zbmath.org pages themselves answered the agent's fetches with a browser challenge (HTTP 403).

[FC90] Gerd Faltings and Ching-Li Chai, *Degeneration of abelian varieties*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3), vol. 22, Springer-Verlag, Berlin, 1990, Zbl 0744.14031.

[Feu21] Christian Feuersänger, *Manual for package pgfplots, version 1.18.1*, `https://ctan.org/pkg/pgfplots`, 2021, Version date 2021-05-15 (local `pgfplots.revision.tex`); CTAN page opened 2026-11-21.

[FW79] Bruce Ferrero and Lawrence C. Washington, *The Iwasawa invariant μ_p vanishes for abelian number fields*, Annals of Mathematics (2) **109** (1979), no. 2, 377–395, Zbl 0443.12001.

[Gal70] P. X. Gallagher, *A large sieve density estimate near $\sigma = 1$*, Inventiones Mathematicae **11** (1970), 329–339, Zbl 0219.10048.

[Gra84] Daniel R. Grayson, *Reduction theory using semistability*, Commentarii Mathematici Helvetici **59** (1984), 600–634, Zbl 0564.20027.

[Gre55] J. A. Green, *The characters of the finite general linear groups*, Transactions of the American Mathematical Society **80** (1955), no. 2, 402–447, Zbl 0068.25605.

[Gre76] Ralph Greenberg, *On the Iwasawa invariants of totally real number fields*, American Journal of Mathematics **98** (1976), no. 1, 263–284, Zbl 0334.12013.

[Gro99] Benedict H. Gross, *Algebraic modular forms*, Israel Journal of Mathematics **113** (1999), 61–93, Zbl 0965.11020.

[GS92] Henri Gillet and Christophe Soulé, *An arithmetic Riemann–Roch theorem*, Inventiones Mathematicae **110** (1992), no. 3, 473–543, Zbl 0777.14008.

[GT16] Wee Teck Gan and Shuichiro Takeda, *A proof of the Howe duality conjecture*, Journal of the American Mathematical Society **29** (2016), no. 2, 473–493, Zbl 1342.11051; arXiv:1407.1995; Crossref online date 2015-07-13.

[HB95] D. R. Heath-Brown, *A mean value estimate for real character sums*, Acta Arithmetica **72** (1995), no. 3, 235–275, Zbl 0828.11040.

[Hen00] Guy Henniart, *Une preuve simple des conjectures de Langlands pour* $\mathrm{GL}(n)$ *sur un corps p-adique*, Inventiones Mathematicae **139** (2000), no. 2, 439–455, Zbl 1048.11092.

[Hir64] Heisuke Hironaka, *Resolution of singularities of an algebraic variety over a field of characteristic zero. I*, Annals of Mathematics (2) **79** (1964), 109–203, Zbl 0122.38603.

[HMvdW+20] Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, et al., *Array programming with NumPy*, Nature **585** (2020), no. 7825, 357–362, Version used: NumPy 2.3.2 (PyPI release record, uploaded 2025-07-24).

[How89] Roger Howe, *Transcending classical invariant theory*, Journal of the American Mathematical Society **2** (1989), no. 3, 535–552, Zbl 0716.22006.

[HT01] Michael Harris and Richard Taylor, *The geometry and cohomology of some simple Shimura varieties*, Annals of Mathematics Studies, vol. 151, Princeton University Press, Princeton, NJ, 2001, With an appendix by Vladimir G. Berkovich; Zbl 1036.11027; Crossref gives the ebook DOI date 2002-12-31.

[Ich04] Atsushi Ichino, *A regularized Siegel–Weil formula for unitary groups*, Mathematische Zeitschrift **247** (2004), no. 2, 241–277, Zbl 1071.11022.

[Ing32] A. E. Ingham, *The distribution of prime numbers*, Cambridge Tracts in Mathematics and Mathematical Physics, vol. 30, Cambridge University

Press, London, 1932, Zbl 0006.39701; reprinted Cambridge Mathematical Library, 1990 (Zbl 0715.11045).

[JS81] H. Jacquet and J. A. Shalika, *On Euler products and the classification of automorphic representations. I*, American Journal of Mathematics **103** (1981), no. 3, 499–558, Zbl 0473.12008.

[Kat04] Kazuya Kato, *p-adic Hodge theory and values of zeta functions of modular forms*, Cohomologies p-adiques et applications arithmétiques. III (Pierre Berthelot, Jean-Marc Fontaine, Luc Illusie, Kazuya Kato, and Michael Rapoport, eds.), Astérisque, vol. 295, Société Mathématique de France, Paris, 2004, Zbl 1142.11336; MR 2104361; Numdam AST_2004___295___117_0, pp. 117–290.

[Kat15] Nicholas M. Katz, *A note on Riemann Hypothesis for curves and hypersurfaces over finite fields*, 2015, Preprint, 9 pages, available from the author's web page. Read 2026-11-20 for ai/2026-11-20-stepanov-bombieri-varieties (pp. 1–9); bibliographic details not re-verified here.

[Kaw82] Yujiro Kawamata, *A generalization of Kodaira–Ramanujam's vanishing theorem*, Mathematische Annalen **261** (1982), no. 1, 43–46, Zbl 0476.14007.

[KH25] Jonathan Kew and Khaled Hosny, *XeTEX, version 3.141592653-2.6-0.999997*, `https://tug.org/xetex/`, 2025, Distributed with TEX Live 2025 (package `xetex`, revision 73850); `xetex --version` reports copyright 2025 SIL International, Jonathan Kew and Khaled Hosny. TUG page opened 2026-11-21.

[Kot92] Robert E. Kottwitz, *Points on some Shimura varieties over finite fields*, Journal of the American Mathematical Society **5** (1992), no. 2, 373–444, Zbl 0796.14014.

[KP84] D. A. Kazhdan and S. J. Patterson, *Metaplectic forms*, Publications Mathématiques de l'IHÉS **59** (1984), 35–142, Not opened; data from the reference list of arXiv:1312.0568.

[KR88] Stephen S. Kudla and Stephen Rallis, *On the Weil–Siegel formula*, Journal für die reine und angewandte Mathematik **387** (1988), 1–68, Zbl 0644.10021.

[Kur92] Masato Kurihara, *Some remarks on conjectures about cyclotomic fields and K-groups of* $\mathbf{Z}$, Compositio Mathematica **81** (1992), no. 2, 223–236, Zbl 0747.11055, MR 1145807. Numdam record `https://www.numdam.org/item/CM_1992__81_2_223_0/` and full text opened 2026-11-20.

[Lan05] Edmund Landau, *Über einen Satz von Tschebyschef*, Mathematische Annalen **61** (1905), no. 4, 527–550, JFM 37.0224.04 and EuDML 158244 give 1905; Crossref gives issue date December 1906.

[Lan13] Kai-Wen Lan, *Arithmetic compactifications of PEL-type Shimura varieties*, London Mathematical Society Monographs Series, vol. 36, Princeton University Press, Princeton, NJ, 2013, Zbl 1284.14004.

[lea23] leanprover (GitHub organization), *lean4checker: replay the Environment for a given Lean module, ensuring that all declarations are accepted by the kernel*, `https://github.com/leanprover/lean4checker`, 2023, Apache-2.0. Repository created 2023-09-25; last commit 91a7f0e8e9dffe927089f5a6edcfeeb8a0e07709 (2026-03-25), a deprecation notice: the tool was re-integrated into Lean 4 as `leanchecker` (leanprover/lean4 PR 11887, merged 2026-01-08) and ships with toolchains from v4.28.0. Opened 2026-11-20 via the GitHub API.

[Leg08] Adrien-Marie Legendre, *Essai sur la théorie des nombres*, seconde ed., Courcier, Paris, 1808, xxiii+480 pp.; zbMATH 1395.11004. Metadata only; the page of Legendre's formula for $v_p(n!)$ was not opened.

[Lin44] Yu. V. Linnik, *On the least prime in an arithmetic progression. I. The basic theorem*, Matematicheskii Sbornik, Novaya Seriya **15(57)** (1944), 139–178, Zbl 0063.03584.

[Lit28] J. E. Littlewood, *On the class-number of the corpus* $P(\sqrt{-k})$, Proceedings of the London Mathematical Society (2) **27** (1928), 358–372, JFM 54.0206.02.

[Mat26] Mathdoc, *Numdam, la bibliothèque numérique française de mathématiques*, `https://www.numdam.org/`, 2026, Developed by Mathdoc, supported by CNRS (INSMI) and Université Grenoble Alpes. Home page opened 2026-11-20.

[Mes72] William Messing, *The crystals associated to Barsotti–Tate groups: with applications to abelian schemes*, Lecture Notes in Mathematics, vol. 264, Springer-Verlag, Berlin, 1972, Zbl 0243.14013.

[Min87] Hermann Minkowski, *Zur Theorie der positiven quadratischen Formen*, Journal für die reine und angewandte Mathematik **101** (1887), 196–202, JFM 19.0189.01.

[Mín08] Alberto Mínguez, *Correspondance de Howe explicite: paires duales de type II*, Annales Scientifiques de l'École Normale Supérieure (4) **41** (2008), no. 5, 717–741, Zbl 1220.22014.

[Mon71] Hugh L. Montgomery, *Topics in multiplicative number theory*, Lecture Notes in Mathematics, vol. 227, Springer-Verlag, Berlin, 1971, Zbl 0216.03501.

[Mor09] Atsushi Moriwaki, *Continuity of volumes on arithmetic varieties*, Journal of Algebraic Geometry **18** (2009), no. 3, 407–457, Zbl 1167.14014; arXiv:math/0612269.

[Mum77] D. Mumford, *Hirzebruch's proportionality theorem in the non-compact case*, Inventiones Mathematicae **42** (1977), 239–272, Zbl 0365.14012.

[MV07] Hugh L. Montgomery and Robert C. Vaughan, *Multiplicative number theory. I. Classical theory*, Cambridge Studies in Advanced Mathematics, vol. 97, Cambridge University Press, Cambridge, 2007, Zbl 1142.11001 (2007 hbk); Crossref gives online date 2006-11-16 for the DOI.

[MW84] B. Mazur and A. Wiles, *Class fields of abelian extensions of* $\mathbb{Q}$, Inventiones Mathematicae **76** (1984), no. 2, 179–330, Zbl 0545.12005.

[Nad90] Alan Michael Nadel, *Multiplier ideal sheaves and Kähler–Einstein metrics of positive scalar curvature*, Annals of Mathematics (2) **132** (1990), no. 3, 549–596, Zbl 0731.53063.

[NIS26] NIST Digital Library of Mathematical Functions, *§5.11 asymptotic expansions (chapter 5, gamma function)*, `https://dlmf.nist.gov/5.11`, 2026, Release 1.2.8 of 2026-09-15; eqs. 5.11.1 (Stirling's series) and 5.11.3 (Stirling's formula); opened 2026-11-20.

[Obu13] Andrew Obus, *On Colmez's product formula for periods of CM-abelian varieties*, Mathematische Annalen **356** (2013), 401–418, Bibliographic data checked on Crossref 2026-11-20; text not opened.

[Ope26a] OpenAI, *GPT-6 and other models in ChatGPT*, `https://help.openai.com/en/articles/20001354-gpt-6-and-other-models-in-chatgpt`, 2026, Model guide, accessed 2026-11-21.

[Ope26b] ———, *Introducing the Codex app*, `https://openai.com/index/introducing-the-codex-app/`, 2026, Product page, 2026-02-02.

[Pag35] A. Page, *On the number of primes in an arithmetic progression*, Proceedings of the London Mathematical Society (2) **39** (1935), no. 1, 116–141, Zbl 0011.14905.

[Pat77] S. J. Patterson, *A cubic analogue of the theta series*, Journal für die reine und angewandte Mathematik **296** (1977), 125–161, Not opened; data from

the reference list of arXiv:2109.07463, where the coefficients of the cubic theta function are cited as Theorem 8.1 of this paper.

[Pat14] ———, *The Fourier coefficients of metaplectic theta series on GL(2) over rational function fields*, Preprint, arXiv:1411.7594v1, 2014, Version 1 of 27 November 2014; abstract and introduction read on 2026-11-21.

[Poi67] G. Poitou, *Cohomologie galoisienne des modules finis*, Travaux et Recherches Mathématiques, vol. 13, Dunod, Paris, 1967, Séminaire de l'Institut de Mathématiques de Lille; Zbl 0161.04203.

[Poo04] Bjorn Poonen, *Bertini theorems over finite fields*, Annals of Mathematics (2) **160** (2004), no. 3, 1099–1127.

[Pyt25] Python Software Foundation, *Python, version 3.13.7*, `https://www.python.org/downloads/release/python-3137/`, 2025, Released 2025-08-14 (release page opened 2026-11-20); local `python3 -version` reports 3.13.7.

[Ral84] S. Rallis, *On the Howe duality conjecture*, Compositio Mathematica **51** (1984), 333–399, Zbl 0624.22011.

[Rib76] Kenneth A. Ribet, *A modular construction of unramified p-extensions of* $\mathbb{Q}(\mu_p)$, Inventiones Mathematicae **34** (1976), no. 3, 151–162, Zbl 0338.12003.

[Rie59] Bernhard Riemann, *Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse*, Monatsberichte der Königlich Preußischen Akademie der Wissenschaften zu Berlin, November 1859, 1859, Venue and date as given by D. R. Wilkins's edition, `https://www.maths.tcd.ie/pub/HistMath/People/Riemann/Zeta/` (opened 2026-11-20); page numbers not verified. Reprinted in Riemann's Gesammelte mathematische Werke (Teubner, 1876; JFM 08.0231.03).

[SC67] Atle Selberg and S. Chowla, *On Epstein's zeta-function*, Journal für die reine und angewandte Mathematik **227** (1967), 86–110, Zbl 0166.05204 (author order Selberg, Chowla as in zbMATH).

[Sch31] Friedrich Karl Schmidt, *Analytische Zahlentheorie in Körpern der Charakteristik p*, Math. Z. **33** (1931), 1–32, Bibliographic data verified on 2026-11-20; content not read.

[Sch73] Wolfgang M. Schmidt, *Zur Methode von Stepanov*, Acta Arith. **24** (1973), 347–367, Bibliographic data verified on 2026-11-20; content not read.

[Sch76] ———, *Equations over finite fields: An elementary approach*, Lecture Notes in Math., vol. 536, Springer, Berlin, 1976, Cited from memory, unverified.

[Sch13] Peter Scholze, *The local Langlands correspondence for* GL_n *over p-adic fields*, Inventiones Mathematicae **192** (2013), no. 3, 663–715, Zbl 1305.22025; arXiv:1010.1540.

[Shi11] Sug Woo Shin, *Galois representations arising from some compact Shimura varieties*, Annals of Mathematics (2) **173** (2011), no. 3, 1645–1741, Zbl 1269.11053.

[Sie35] Carl Ludwig Siegel, *Über die Classenzahl quadratischer Zahlkörper*, Acta Arithmetica **1** (1935), no. 1, 83–86, Zbl 0011.00903.

[Sou79] C. Soulé, *K-théorie des anneaux d'entiers de corps de nombres et cohomologie étale*, Inventiones Mathematicae **55** (1979), no. 3, 251–295, Zbl 0437.12008.

[Ste69] S. A. Stepanov, *On the number of points of a hyperelliptic curve over a finite prime field*, Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya **33** (1969), 1171–1181, Zbl 0192.58002 (English title as in zbMATH; original Russian). English transl.: Math. USSR-Izv. 3 (1969), 1103–1114. Bibliographic data verified by index search on 2026-11-20; content not read.

[Sti09] Henning Stichtenoth, *Algebraic function fields and codes*, 2 ed., Grad. Texts in Math., vol. 254, Springer, Berlin, 2009, Section numbers reliable; §5.2 (p. 197), Theorems 5.2.1 and 5.2.3 confirmed from a search snippet on 2026-11-20. All other theorem numbers cited here are from memory, unverified.

[Sto21] Augusto Stoffel, *The tikz-cd package: commutative diagrams with TikZ, version 1.0*, `https://ctan.org/pkg/tikz-cd`, 2021, Version date 2021-05-04 (local `tikz-cd.sty`, TeX Live 2025, revision 59133); CTAN page opened 2026-11-21, listing version 1.0 of 2021-05-04.

[Stu76] Ulrich Stuhler, *Eine Bemerkung zur Reduktionstheorie quadratischer Formen*, Archiv der Mathematik **27** (1976), 604–610, Zbl 0338.10024.

[T+25] Hàn Thế Thành et al., *pdfTeX, version 3.141592653-2.6-1.40.27*, `https://ctan.org/pkg/pdftex`, 2025, Local `pdftex -version` reports "pdfTeX 3.141592653-2.6-1.40.27 (TeX Live 2025/Homebrew)". CTAN page opened 2026-11-21.

[Tan23] Till Tantau, *The TikZ and PGF packages, version 3.1.10*, `https://ctan.org/pkg/pgf`, 2023, Version date 2023-01-15 (local `pgf.revision.tex`); CTAN page opened 2026-11-21.

[Tao14] Terence Tao, *The Bombieri–Stepanov proof of the Hasse–Weil bound*, Blog post, 2 May 2014, 2014, Page text read on 2026-11-20; comments not read.

[Tat63] John Tate, *Duality theorems in Galois cohomology over number fields*, Proceedings of the International Congress of Mathematicians (Stockholm, 1962), 1963, Zbl 0126.07002, pp. 288–295.

[Tch52] P. Tchebichef, *Mémoire sur les nombres premiers*, Journal de Mathématiques Pures et Appliquées (1re série) **17** (1852), 366–390, Numdam JMPA_1852_1_17___366_0; full text opened.

[The23] The mpmath development team, *mpmath: a Python library for arbitrary-precision floating-point arithmetic, version 1.3.0*, `https://mpmath.org/`, 2023, PyPI release record for 1.3.0 (author Fredrik Johansson), uploaded 2023-03-07, opened 2026-11-21.

[The24] The curl project, *curl, version 8.7.1*, `https://curl.se/`, 2024, Local `curl -version` reports 8.7.1 (installed version on 2026-11-21; the version at the run time of 2026-11-20 was not recorded). Home page opened 2026-11-21.

[The25a] The ripgrep developers, *ripgrep, version 15.1.0*, `https://github.com/BurntSushi/ripgrep/releases/tag/15.1.0`, 2025, Released 2025-10-22 (GitHub release record, opened 2026-11-20); local `rg -version` reports 15.1.0.

[The25b] The Sage Developers, *SageMath, the Sage Mathematics Software System, version 10.8*, `https://www.sagemath.org/`, 2025, Released 2025-12-18 (changelog `https://www.sagemath.org/changelogs/sage-10.8.txt`, opened 2026-11-21).

[The26a] The Lean developers (leanprover/lean4), *Lean 4, release v4.33.1*, `https://github.com/leanprover/lean4/releases/tag/v4.33.1`, 2026, Published 2026-08-21 (GitHub release record, opened 2026-11-20); the toolchain pinned by the Alpöge archive. For the system cite also [**dMU21**].

[The26b] The mathlib Community, *mathlib4: the Lean mathematical library, commit 0df444a360eaa60ab8c11dca51a86af692955474*, `https://github.com/leanprover-community/mathlib4/commit/0df444a360eaa60ab8c11dca51a86af692955474`, 2026, Commit "chore: bump toolchain to v4.33.1", 2026-08-21 (GitHub API record, opened 2026-11-20); revision pinned in the Alpöge archive's lake-manifest.json. Differs from the repository's own Mathlib (key Mathlib2026library).

[The26c] The Poppler developers, *Poppler: a PDF rendering library, with the utility `pdftotext`*, `https://poppler.freedesktop.org/`, 2026, Home page

opened 2026-11-20 (“a PDF rendering library based on the xpdf-3.0 code base”). Local `pdftotext -v` reports version 26.04.0.

[Tit39] E. C. Titchmarsh, *The theory of functions*, 2 ed., Oxford University Press, 1939, Cited from memory, unverified.

[TY07] Richard Taylor and Teruyoshi Yoshida, *Compatibility of local and global Langlands correspondences*, Journal of the American Mathematical Society **20** (2007), no. 2, 467–493, Zbl 1210.11118; arXiv:math/0412357; Crossref online date 2006-07-10.

[Vak25] Ravi Vakil, *The rising sea: Foundations of algebraic geometry*, Princeton University Press, Princeton, NJ, 2025, Publication date October 2025 (publisher page); the author's public version of October 21, 2025 is at `https://math.stanford.edu/~vakil/216blog/FOAGoct2125public.pdf`. Data checked on the author's and the publisher's web pages on 2026-11-21; chapter and theorem numbers differ between versions. All locators in this book are in the numbering of the public version of October 21, 2025, in which they were checked on 2026-11-21.

[Vie82] Eckart Viehweg, *Vanishing theorems*, Journal für die reine und angewandte Mathematik **335** (1982), 1–8, Zbl 0485.32019.

[Voe11] Vladimir Voevodsky, *On motivic cohomology with $\mathbb{Z}/l$-coefficients*, Annals of Mathematics (2) **174** (2011), no. 1, 401–438, Zbl 1236.14026; arXiv:0805.4430.

[Wal36] Arnold Walfisz, *Zur additiven Zahlentheorie. II*, Mathematische Zeitschrift **40** (1936), no. 1, 592–607, JFM 61.1070.02.

[Wal90] J.-L. Waldspurger, *Démonstration d'une conjecture de dualité de Howe dans le cas p-adique, $p \neq 2$*, Festschrift in Honor of I. I. Piatetski-Shapiro on the Occasion of his Sixtieth Birthday, Part I, Israel Mathematical Conference Proceedings, vol. 2, Weizmann, Jerusalem, 1990, Zbl 0722.22009 (English title there: “Proof of a duality conjecture of Howe in the p-adic case, $p \neq 2$”); French title as commonly cited, not checked on a publisher page, pp. 267–324.

[Wal92] Nolan R. Wallach, *Real reductive groups II*, Pure and Applied Mathematics, vol. 132-II, Academic Press, Boston, MA, 1992, Zbl 0785.22001.

[Was97] Lawrence C. Washington, *Introduction to cyclotomic fields*, second ed., Graduate Texts in Mathematics, vol. 83, Springer, New York, 1997, Zbl 0966.11047.

[Wei48] André Weil, *Sur les courbes algébriques et les variétés qui s'en déduisent*, Hermann, Paris, 1948, Cited from memory, unverified.

[Wei65] André Weil, *Sur la formule de Siegel dans la théorie des groupes classiques*, Acta Mathematica **113** (1965), 1–87, Zbl 0161.02304.

[Wid41] David Vernon Widder, *The Laplace transform*, Princeton Mathematical Series, vol. 6, Princeton University Press, Princeton, N.J., 1941, Zbl 0063.08245; JFM 67.0384.01.