Families and dichotomies in the circle method (unofficial version)

Victor Y. Wang

(See `http://arks.princeton.edu/ark:/88435/dsp01rf55zb86g` for the official version.)

Abstract. In the eighties, Hooley applied the Grand Riemann Hypothesis, and what practically amounts to the general Langlands reciprocity (automorphy) conjecture, in a fresh new way, over certain families of cubic threefolds. This eventually led to conditional near-optimal bounds for the number N of integral solutions to $x_1^3+\cdots+x_6^3=0$ in expanding boxes.

Elsewhere, building on Hooley's work, we have given new applications of large-sieve hypotheses, the Square-free Sieve Conjecture, and predictions of Random Matrix Theory type, over the same geometric families—for instance, conditional optimal asymptotics for N in a large class of regions, with applications to sums of three cubes. The underlying harmonic analysis—which in rough form goes back to Kloosterman—picks up equally significant contributions from the classical major and minor arcs in the circle method.

Here, we mainly provide extended summaries, commentary, and other complementary material, leaving complete traditional accounts to papers available elsewhere. Two central themes of this thesis are families (of arithmetic or analytic objects) and dichotomies (between structure and randomness). We especially consider (mainly in relation to the aforementioned cubic questions) families of regions, weights, point counts, oscillatory integrals, exponential sums, Hasse–Weil L-functions, and quadratic equations; and dichotomies for point counts over finite and infinite fields.

Contents

Acknowledgements

Many thanks to Professors Amit Ghosh and Peter Sarnak for suggesting the main problem considered here. In general, I am most grateful to Professor Sarnak for pushing me in the direction of this thesis; for sharing much knowledge and insight; for positively influencing my taste in math; and for being a wonderful adviser, patient and encouraging to an endlessly ignorant student. I would also like to thank him—and the audience members of his informal Fall 2020 seminar—for the chance to speak about, clarify, and simplify preliminary versions of my work. In addition, I thank the late Professor Christopher Hooley, who has, through his papers, amply influenced almost all aspects of this thesis; in my mind, he has been much like a second adviser, inspiring and full of wit.

Furthermore, I thank Professors Manjul Bhargava, Tim Browning, Bill Duke, and Bob Vaughan very much for their support. I also thank Professor Bhargava for serving as a Reader, and for exposing me to many beautiful ideas in arithmetic statistics and related areas. Similarly, I thank Professors Nick Katz and Shou-Wu Zhang for examining my defense, for sharing many funny and inspiring stories over the years, and for helping me understand algebraic geometry a little bit better; I have especially learned much about "life over finite fields" from Professor Katz.

Also, I must thank Jill LeClair and Dr. Jennifer Johnson, as well as Ankit Tak, Will Crow, and others, for their general kindness, and for their excellent help with logistics and other matters throughout my time at Princeton.

Aside from those listed above, many other people have contributed professionally to my academic development in the last few years, in ways large and small. At the risk of listing both too many and too few names, I would like to thank Jonathan Bober, Andy Booker, Brian Conrey, Izzet Coskun, Samit Dasgupta, Simona Diaconu, Alex Gamburd, Jayce Getz, Roger Heath-Brown, Yotam Hendel, Bob Hough, Edna Jones, Ilya Khayutin, Valeriya Kovaleva, Chung-Hang Kwan, Bao Le Hung, Mark Levi, Zhiyuan Li, Yuchen Liu, Jasmin Matz, Danny Neftin, Fabien Pazuki, Lillian Pierce, Morten Risager, Anurag Sahay, Will Sawin, Chad Schoen, Freydoon Shahidi, Efthymios Sofos, Christof Sparber, Ramin Takloo-Bighash, Yuri Tschinkel, Akshay Venkatesh, Hong Wang, Junho Whang, Trevor Wooley, Liyang Yang, Shing-Tung Yau, Ruixiang Zhang, Wei Zhang, my academic siblings and neighbors, and others I have failed to mention, for their general comments, questions, answers, suggestions, interest, support, and help.

On a more personal note, I thank my academic siblings, and many other fellow graduate students and members of the math department, for making math and life at Princeton much more fun and interesting than it would be otherwise. Thanks also to Calvin Deng and Wayne Zhao for being fantastic roommates.

I am also grateful to many other people, including many old friends, teachers, and mentors, who never cease to inspire me. And most of all, I thank my family for their patience, love, support, and spirit.

To friends, family, and adventure, I dedicate this thesis.

Conventions

We let $\mathbb{Z}_{\geq 0} := \{n \in \mathbb{Z} : n \geq 0\}$, and define $\mathbb{Z}_{\geq 2}, \mathbb{R}_{<0}, \ldots$ similarly.

We let $\mathbf{1}_E$ denote the *indicator value* of an event E; i.e. we let $\mathbf{1}_E := 1$ if E holds, and $\mathbf{1}_E := 0$ otherwise. When it would be too cumbersome to "restrict" a sum explicitly via indicator notation, we use the shorthand $\sum'$ to denote a *restricted sum*, whose variables are restricted according to context.

In number-theoretic contexts, p will denote a prime, and d a positive divisor. We let

$v_p(-)$ denote the usual p-adic valuation. For $n \in \mathbb{Z}_{\ge 1}$, we let $\tau(n) := \sum_{d \mid n} 1$; $\omega(n) := \sum_{p \mid n} 1$; $\phi(n) := n \prod_{p \mid n}(1 - p^{-1})$; $\operatorname{rad}(n) := \prod_{p \mid n} p$; and $\mu(n) := \mathbf{1}_{n = \operatorname{rad}(n)} \cdot (-1)^{\omega(n)}$. Given arithmetic functions a, b, we let $a * b$ denote their *Dirichlet convolution*.

All L-functions will be analytically normalized (to have critical line $\Re(s) = 1/2$, or an analogous property in the case of general Hasse–Weil L-functions).

By default, $\|\boldsymbol{x}\|$ will refer to the ℓ^∞-norm $\|\boldsymbol{x}\|_\infty := \max_i(|x_i|)$ when $\boldsymbol{x}$ is a vector. And in the context of indices, $[n]$ will denote the set $\{1, 2, \dots, n\}$ when $n \in \mathbb{Z}_{\ge 0}$.

We let $e(t) := e^{2\pi i t}$ (if t "makes sense" in $\mathbb{R}/\mathbb{Z}$), and $e_r(t) := e(t/r)$ (for $r \in \mathbb{R}^\times$).

We will use algebro-geometric notation freely, both for convenience and for rigor; but most of our varieties and schemes will be explicitly embedded in a projective (or affine) space, and *applied* to concrete questions. In general, we let $V_U(f)_{/R}$, or (by a minor abuse of notation) $V_U(f)_R$, denote the closed subscheme of U_R cut out by $f = 0$; we let $V_U(f_1, \dots, f_r)_R$ denote the scheme-theoretic intersection $\bigcap_{i \in [r]} V_U(f_i)_R$. If the base ring R is clear from context, we may omit it.

Given a polynomial $f(x_1, \dots, x_m)$, we let $\operatorname{Hess}(f)$ denote the usual $m \times m$ Hessian matrix, $\det(\operatorname{Hess} f)$ the Hessian determinant, $\operatorname{hess}(f)_R := V_{\mathbb{A}^m}(\det(\operatorname{Hess} f))_R$ the affine Hessian vanishing locus over R, and (when f is homogeneous) $\operatorname{hess}(V_{\mathbb{P}^{m-1}}(f)_R) := V_{\mathbb{P}^{m-1}}(\det(\operatorname{Hess} f))_R$ the projective Hessian hypersurface over R.

We use the subscript notation $\partial_x := \partial/\partial x$ for derivatives. We adopt multi-index notation for multivariable calculus, especially in the context of derivatives; e.g. we let $\partial_{\boldsymbol{c}}^{\boldsymbol{r}} := \partial_{c_1}^{r_1} \cdots \partial_{c_m}^{r_m}$ (for $\boldsymbol{r} \in \mathbb{Z}_{\ge 0}^m$) and $|\boldsymbol{b}| := \sum_{i \in S} b_i$ (for $\boldsymbol{b} \in \mathbb{Z}^S$).

We write $f \ll_S g$, or $g \gg_S f$, to mean $|f| \lesssim_S g$, i.e. $|f| \le Cg$ for some $C = C(S) > 0$. We write $f \asymp_S g$ if $f \ll_S g \ll_S f$. We let $O_S(g)$ denote a quantity that is $\ll_S g$; and similarly, $\Omega_S(f)$ a quantity $\gg_S f$, and $\Theta_S(g)$ a quantity $\asymp_S g$. As usual, the implied constants C throughout an argument will depend on one another in a logical fashion. A few clarifications on our use of inexplicit inequalities may be helpful:

(1) We will often attach "size adjectives" to inexplicit inequalities when we really mean to describe their implied constants; e.g. "if $f \ll_S g$ is small" would mean "if C is small in terms of S, and if $|f| \le Cg$".

(2) Typically "$|f| \le Cg$" will either appear in a statement as (i) a *hypothesis*, in which case we allow C to be arbitrary, unless there is some restriction given explicitly (e.g. "sufficiently small") or by context (e.g. when applying a previous bound (X), the constant C would simply be "copied" from (X)); or (ii) a *conclusion*, in which case C would "follow" or "result" from the proof.

(3) In the context of the previous point, a phrase of the form "P unless Q" should be read "if $\neg P$, then Q" (with "hypothesis" $\neg P$ and "conclusion" Q); but we will try to minimize our use of such "negative" phrases, since "positive" phrases like "if P, then Q" or "Q, provided P" (with "hypothesis" P and "conclusion" Q) are generally expressive enough for us.

Chapter 1

Introduction

1.1 General motivation

Diophantine equations in the tradition of Hardy–Littlewood, and L-functions in the tradition of Riemann, are both central objects in number theory. Some natural problems and questions about them are the following:

(1) Count, produce, or bound solutions to algebraic equations over the integers ($\mathbb{Z}$) or related rings (e.g. $\mathbb{F}_p[t]$ or $\mathbb{F}_p$, for various primes p).

(2) Prove approximations to the Grand Riemann Hypothesis (GRH) for individual L-functions, or analyze statistics (especially those of Random Matrix Theory type) over families.

(3) To what extent are (1)–(2) related?

Example 1.1.1 (BSD)**.** Let C be a soluble cubic curve in $\mathbb{P}^2_{\mathbb{Q}}$, cut out by $F = 0$ for some ternary cubic form $F \in \mathbb{Z}[x, y, z]$ with nonzero discriminant. (For example, one could take $F = x^3+y^3+60z^3$, but not $F = 3x^3+4y^3+5z^3$.) Then Birch and Swinnerton-Dyer conjectured an equality between two integers: (1) rank $J(C)(\mathbb{Q})$, which measures how many primitive integral solutions $(x, y, z) \in [-X, X]^3$ to $F = 0$ there are as $X \to \infty$; and (2) $\operatorname{ord}_{s=1/2} L(s, C)$, where $L(s, C)$—the *Hasse–Weil L-function* associated to C—roughly encodes the number of solutions to $F \equiv 0 \bmod p$ as p varies. In general, the "$\geq$" direction, i.e. "producing" points, remains especially mysterious (even over function fields), though both directions are difficult and interesting. But the modularity theorem for the elliptic curve $J(C)$ often allows one to produce points, via Heegner points and the Gross–Zagier theorem; contrast with the use of modularity in Wiles' proof of Fermat's last theorem (showing that certain points do *not* exist).

Example 1.1.2 (Quadratic equations)**.** The most difficult part of the solution of Hilbert's eleventh problem (up to questions of effectiveness), namely the part regarding integral representations of integers by ternary quadratic forms with integral coefficients (due to Iwaniec, Duke, and Schulze-Pillot over $\mathbb{Q}$), also makes essential use of automorphic forms, through subconvex L-function bounds obtained through the study of L-function families. See e.g. [Sar00] for details.

Rational representations are much simpler, with a very clean existence theory (a local-to-global principle with no exceptions) given by Hasse–Minkowski, quantifiable by the sharpest forms of the circle method (see e.g. [DFI93, HB96] over $\mathbb{Q}$).

Since much of what we discuss below will be conditional on standard number-theoretic conjectures, let us first make two remarks, to give some reasons to view "standard conjectures" as valuable working hypotheses.

Remark 1.1.3. Given an (appropriately normalized) automorphic L-function $L(s,\pi)$, GRH is *equivalent* to the "naive probabilistic conjecture" $\sum_{n\le N}\mu_\pi(n)\ll_{\pi,\epsilon} N^{1/2+\epsilon}$ [IK04, Proposition 5.14]. But what makes GRH so fascinating is that in addition to strong direct heuristic and numerical evidence (at least in low dimensions) for its truth, there is striking indirect evidence—from various comparisons with function field analogs and Random Matrix Theory (RMT) models—going beyond what the "naive probabilistic conjecture" above would seem to merit.

Meanwhile, given a (positive definite) ternary integral quadratic form such as $F_0 = ax^2+by^2+cz^2$, the "naive conjecture" for $r_{F_0}(n) := \#\{(x,y,z)\in\mathbb{Z}^3 : F_0 = n\}$ as $n\to\infty$ is simply false in general: while simple probabilistic models can detect *local* obstructions, they miss further subtleties such as "exceptional square classes" arising from *spinor norm* obstructions. As it turns out, $r_{F_0}(n)$ does satisfy a certain natural asymptotic after adjusting for such obstructions, but the known proofs (see Example 1.1.2) all use automorphic forms or L-functions in an essential way.

Remark 1.1.4. Over global function fields, the analog of GRH is known. But modulo (Langlands and) GRH, the analog of "counting points in natural regions on varieties" appears so far to be comparable in difficulty to the problem over number fields. (There are, however, more techniques available for the analogous softer question of "producing points" [Tia17].)

1.2 Main problems of interest

Though BSD remains wide open in general, one can certainly consider many other interesting questions of a similar local-to-global nature, including both qualitative and quantitative questions about certain cubic equations. In this thesis, we will focus on Examples 1.2.1 and 1.2.3 below.

Example 1.2.1 (The Fermat cubic fourfold)**.** Let $r_3(a) := \#\{(x,y,z)\in\mathbb{Z}^3_{\ge 0} : x^3+y^3+z^3 = a\}$, for $a\in\mathbb{Z}_{\ge 0}$. For real $X\to\infty$, let $M_2(X) := \sum_{0\le a\le X^3} r_3(a)^2$. By Cauchy–Schwarz, $\#\{a\in[0,A] : r_3(a)\ne 0\}\gg A^2/M_2(A^{1/3})$ for real $A\to\infty$.

Let $F(\boldsymbol{x}) = F(x_1,\dots,x_6) := x_1^3+\cdots+x_6^3$. Then $M_2(X) = \#\{\boldsymbol{x}\in\mathbb{Z}^6\cap XK : F(\boldsymbol{x}) = 0\}$ for some fixed compact region $K\subseteq\mathbb{R}^6$. Beginning with Hardy and Littlewood in [HL25] (roughly), many authors, inspired in part by connections to the statistics of *sums of three cubes* (via M_2, for instance), have sought to estimate the number of solutions $\boldsymbol{x}\in\mathbb{Z}^6$ to $F(\boldsymbol{x}) = 0$ in expanding boxes or other regions.

The integral solutions to $F(\boldsymbol{x}) = 0$ are expected to exhibit a randomness-structure dichotomy (along the lines of the Manin conjectures), as we now explain. The purely probabilistic "Hardy–Littlewood model" predicts $M_2(X)\sim c_{\mathrm{HL}}\cdot X^{6-3}$, where the constant

$c_{\mathrm{HL}} := \sigma_{\mathbb{R}} \cdot \prod_p \sigma_p \in \mathbb{R}_{>0}$ is a product of local densities measuring the "local" or "adelic" (i.e. real and p-adic) bias of the equation $F = 0$ (over the regions K and $\mathbb{Z}_p^6$). But $F = 0$ also has $\asymp X^{6/2}$ *special* solutions $\boldsymbol{x} \in \mathbb{Z}^6 \cap XK$, i.e. $\boldsymbol{x}$ with "$x_i + x_j = 0$ in pairs" (e.g. $\boldsymbol{x}$ with $x_1 + x_2 = x_3 + x_4 = x_5 + x_6 = 0$). In fact, Hooley showed that $M_2(X) - \max(c_{\mathrm{HL}} \cdot X^3, \#\{\text{special } \boldsymbol{x} \in \mathbb{Z}^6 \cap XK\}) \gg X^3$ for all sufficiently large $X \gg 1$ [Hoo86a, Theorem 1 and the ensuing sentence], and conjectured that $M_2(X) \sim c_{\mathrm{HL}} \cdot X^3 + \#\{\text{special } \boldsymbol{x} \in \mathbb{Z}^6 \cap XK\}$ [Hoo86a, Conjecture 2]. For convenience, call this conjecture "HLH" (or "HLH for (F, K)").

What is known towards HLH? By Cauchy between "structure" and "randomness" (in four and eight variables, respectively), Hua showed that $M_2(X) \ll_\epsilon X^{7/2+\epsilon}$ [Hua38]. By isolating a new source of randomness ("typical divisors" of integers $x \ll X$), Vaughan gave a more robust proof of Hua's bound, ultimately leading to the improvement $M_2(X) \ll_\epsilon X^{7/2}(\log X)^{\epsilon-5/2}$ [Vau86, Vau20]. Conditionally under a certain "Hypothesis HW" (practically amounting to automorphy and GRH for the Hasse–Weil L-functions associated to smooth projective hyperplane sections of the form $F(\boldsymbol{x}) = \boldsymbol{c} \cdot \boldsymbol{x} = 0$, for $\boldsymbol{c} \in \mathbb{Z}^6$), Hooley established the near-optimal bound $M_2(X) \ll_\epsilon X^{3+\epsilon}$ [Hoo86b, Hoo97], using an "upper-bound precursor" to the delta method of [DFI93, HB96]; Heath-Brown proved (among other things) the same result using the delta method [HB98], independently modulo [Hoo86b].

HLH lies beyond the classical Hardy–Littlewood circle method (according to square-root "pointwise" minor arc considerations), though the "Hardy–Littlewood part" arises satisfactorily from the classical major arcs. But the delta method opens the door to progress on HLH, by harmonically decomposing the true minor arc contribution in a "dual" fashion (via Poisson summation, essentially applied in the form of the Nyquist–Shannon sampling theorem). This is the impetus for [Wan21d, Wan21a], papers to be discussed in Chapters 6–8 below.

Remark 1.2.2. The main results of [Wan21d, Wan21a] apply to diagonal cubic forms in six variables, but to simplify notation, we will mostly focus on the Fermat case. The Fermat case is arguably the most interesting anyways; a simple Hölder argument shows that out of all diagonal cubic forms in six variables, the Fermat cubic has the greatest number, $\int_0^1 d\theta \left(\sum_{x \in \mathbb{Z} \cap [-X,X]} e(\theta x^3)\right)^6$, of integral zeros $\boldsymbol{x} \in [-X, X]^6$.

On the other hand, to extend our work to general cubic forms in six variables (with nonzero discriminant, say), it would take a lot of technical, but significant, work; see Chapters 6–8 below for some remarks on what is presently missing.

Example 1.2.3 (Sums of cubes)**.** For integers $s \geq 1$, let $g_s(y_1, \ldots, y_s) := y_1^3 + \cdots + y_s^3$. In 1825, Ryley proved $\mathbb{Q} \subseteq g_3(\mathbb{Q}^3)$ by explicitly constructing $x, y, z \in \mathbb{Q}(t)$ such that $t = x^3 + y^3 + z^3$. (See Example 1.3.4 below for details.)

How about writing a given integer $a \in \mathbb{Z}$ *integrally* in the form $g_3(\mathbb{Z}^3)$? On the one hand, $g_1(\mathbb{Z}) \subseteq \{0, \pm 1\} \bmod 9$, so $g_3(\mathbb{Z}^3) \subseteq \{0, \pm 1, \pm 2, \pm 3\} \bmod 9$, which *prevents* each integer $a \equiv \pm 4 \bmod 9$ from decomposing as a sum of three integer cubes. On the other hand, as it turns out, there do not exist any other such *local obstructions*. Does every integer $a \not\equiv \pm 4 \bmod 9$ in fact lie in $g_3(\mathbb{Z}^3)$? (Cf. [HB92, p. 623].)

This question seems quite subtle: the space of possible representations seems "relatively sparse on average" over a, and there is no known algorithm that can *provably* determine whether an "arbitrary input" $a \not\equiv \pm 4 \bmod 9$ is represented or not. Even just to give a complete affirmative answer to the question in the finite range $|a| < 100$, Booker and Sutherland had to search quite far (and even further to find a *new* representation of $3 = g_3(1,1,1) = g_3(4,4,-5)$); see Theorems 1.2.5 and 1.2.6 below.

In fact, even the weaker question of proving $\mathbb{Z} \subseteq g_4(\mathbb{Z}^4)$ appears to be open (though it is known that $\{a \not\equiv \pm 4 \bmod 9\} \subseteq g_4(\mathbb{Z}^4)$ [Dem66], and therefore that $\mathbb{Z} \subseteq g_5(\mathbb{Z}^5)$).

The situation is better *on average*: [Dav39] showed that asymptotically 100% of integers $a > 0$ lie in $g_4(\mathbb{Z}^4_{\geq 0})$. (However, for *general* quaternary cubic forms with nonzero discriminant, the analog of [Dav39]'s result is only known conditionally, by [Hoo16].) One of the main goals of this thesis is to summarize our work [Wan21a] proving conditional results of a similar flavor for $g_3(\mathbb{Z}^3_{\geq 0})$ and $g_3(\mathbb{Z}^3)$.

For now, let us note that $g_3(\mathbb{Z}^3_{\geq 0})$ contains $\gg A^{0.91709477}$ integers $a \in [0, A]$, unconditionally, by [Woo95, Woo00, Woo15]; and $\gg_\epsilon A^{1-\epsilon}$ integers $a \in [0, A]$, conditionally on Hypothesis HW from Example 1.2.1, by [Hoo86b, Hoo97]. Both Wooley and Hooley use upper bounds on certain second moments, in the spirit of Example 1.2.1. Wooley uses difficult iterative arguments involving smooth numbers.

Remark 1.2.4. See Observation 1.3.2 and Remark 1.3.3 below for a discussion of possible obstructions (or lack thereof) to the Hasse principle for $g_3(\mathbb{Z}^3)$ and related problems. In general, while rational points can already be very subtle (see e.g. BSD), integral points can be even subtler (see e.g. [Har17, Wil22]).

Theorem 1.2.5 ([Boo19])**.** *Using integers with* 16 *digits,*

$$g_3(8866128975287528, -8778405442862239, -2736111468807040) = 33.$$

Theorem 1.2.6 ([BS21])**.** *Using integers with* 17 *digits,*

$$g_3(-80538738812075974, 80435758145817515, 12602123297335631) = 42.$$

Also, using two integers with 21 *digits, and a* third with only 18 digits*,*

$$g_3(569936821221962380720, -569936821113563493509, -472715493453327032) = 3,$$

thus affirmatively answering a question of Mordell.

1.3 Related background

For Observation 1.3.2 and Example 1.3.4, we need the following technical result:

Proposition 1.3.1 (Cf. [Poo17, Remark 9.4.11])**.** *Fix a smooth projective cubic hypersurface W of dimension $d \geq 2$ over a field K. Then $W(K) \neq \emptyset$ if and only if W is K-unirational; and in this case, we must have $\overline{W(K)} = W$ if K is infinite.*

Proof. The unirationality criterion is due to [Seg43] when $(d, K) = (2, \mathbb{Q})$, and to [Kol02] in general. The rest hinges on the fact that if $\#K = \infty$, then $\mathbb{P}^d_K(K)$ is dense in $\mathbb{P}^d_K$. $\square$

Observation 1.3.2 (Based on [CG66]). Let $h(\boldsymbol{y}) = h(y_1, y_2, y_3) := 5y_1^3 + 12y_2^3 + 9y_3^3$. Then the Hasse principle for $h(\mathbb{Z}^3)$ is *false*. More precisely, the "exceptional set" $\mathcal{E}(h) := \{a \in \mathbb{Z} : h(\boldsymbol{y}) = a \text{ fails Hasse}\}$ is *nonempty*.

Proof. For an inhomogeneous ternary cubic $P \in \mathbb{Q}[\boldsymbol{y}] = \mathbb{Q}[y_1, y_2, y_3]$, let W_P denote the hypersurface in $\mathbb{P}^3_{\mathbb{Q}}$ cut out by $y_4^3 P(\boldsymbol{y}/y_4) = 0$. By [CG66], the surface $W = W_{h+10}$ has $\mathbf{A}_{\mathbb{Q}}$-points, but no $\mathbb{Q}$-points. This can be explained by $W(\mathbf{A}_{\mathbb{Q}})^{\mathrm{Br}}$, the "Brauer–Manin set" (BMS) for W (see [Man86, §47.6] and [Poo17, §9.4.9 and Remark 9.4.30]).

Now for convenience, let *R-soluble* mean "soluble over R". Then by the previous paragraph, $h(\boldsymbol{y}) + 10 = 0$ is $\mathbb{Q}_v$-soluble for each place $v \le \infty$ (by Proposition 1.3.1, applied to $W_{\mathbb{Q}_v}$), but *not* $\mathbb{Q}$-soluble. In particular, there must exist a *constant* $q \in \mathbb{Z} \setminus \{0\}$ such that $h(\boldsymbol{y}) + 10t^3 = 0$ is $\mathbb{Z}_p$-soluble for each $t \in q\mathbb{Z}$ and prime $p \le 5$.

On the other hand, $h(\boldsymbol{y}) = a$ is $\mathbb{Z}_p$-soluble for each $a \in \mathbb{Z}$ and prime $p \ge 7$ (as one can prove using Hensel's lemma and either the Cauchy–Davenport theorem or the Weil conjectures). We conclude that for each $t \in q\mathbb{Z} \setminus \{0\}$, the equation $h(\boldsymbol{y}) = -10t^3$ is $\mathbf{A}_{\mathbb{Z}}$-soluble, but *not* $\mathbb{Z}$-soluble. In other words, $\mathcal{E}(h) \supseteq \{-10t^3 : t \in q\mathbb{Z} \setminus \{0\}\}$. □

Remark 1.3.3. The (rational!) BMS for W_{h-a} "explains" at least part of $\mathcal{E}(h)$. But that for W_{g_3-a} says nothing about the analogous set $\mathcal{E}(g_3)$: even the *integral* ("entière") $\mathbb{Z}$-BMS for $g_3 - a = 0$ can never obstruct $\mathbb{Z}$-solubility [CTW12, p. 1304]. (The BMS for W_{g_3-a} *can* obstruct weak approximation over $\mathbb{Q}$ [HB92, Theorem 1 and ensuing paragraphs]. This is relevant to *approximation questions* for $g_3 - a = 0$, but we have chosen to focus on *Hasse principles* in Example 1.2.3.)

Nonetheless, could there be other obstructions? (For other affine cubic surfaces, there can indeed be obstructions *not* "directly" explained by the $\mathbb{Z}$-BMS; see [CTWX20, Theorem 5.14] and [LM21, Theorem 1.5], both based "indirectly" on the $\mathbb{Z}$-BMS. Cf. [CS20, Corollary 1.1] and [LX15, Corollary 3.10], over $\mathbb{Q}$.)

Let us now recall some background on producing points, related to Examples 1.2.1 and 1.2.3. This background will also illustrate how weaker goals (e.g. producing instead of counting points) sometimes allow for greater flexibility and creativity.

Example 1.3.4. If a cubic equation over a field K has a sufficiently general solution over a quadratic extension L/K, then it has a solution over K. (See also [Bir04, §3]'s discussion of [Hee52] for a similar fact called "Heegner's lemma", for L/K of odd degree and "typical" quartic $y^2 = f(x)$ over K.)

This idea (or alternatively, Proposition 1.3.1) can be used to geometrically derive Ryley's theorem that every rational number is a sum of three rational cubes, starting from a "trivial solution at infinity"; see [Man86, Introduction].

Example 1.3.5. Using ternary *quadratic* forms, [Lin43] proved $G(3) \le 7$, i.e. every sufficiently large integer is a sum of 7 positive cubes. This remains the record for $G(3)$.

The expected asymptotic formula in Waring's problem for 7 cubes remains unproven, but would follow if one knew $M_2(X) \ll X^{3.25-\delta}$ (with $M_2(X)$ defined as in Example 1.2.1). The asymptotic for 8 cubes is "barely" known [Vau86], but an easy proof would follow if one knew $M_2(X) \ll X^{3.5-\delta}$.

Example 1.3.6. Assume the finiteness of the Tate–Shafarevich group $\text{Ш}(E/K)$ for every quadratic extension $K/\mathbb{Q}$ and elliptic curve $E = E_A : X^3 + Y^3 = AZ^3$. Then [SD01] proved, over $\mathbb{Q}$, the Hasse principle for diagonal cubic threefolds $a_1x_1^3 + \cdots + a_5x_5^3 = 0$ in $\mathbb{P}^4$, and for "typical" diagonal cubic surfaces in $\mathbb{P}^3$. The proof uses that diagonal hypersurfaces are

DPC (dominated by a product of curves such as $a_i x_i^3 + a_j x_j^3 + B x_0^3 = 0$), and a version of the *fibration method* (i.e. finding a "nice" B).

Remark 1.3.7. Over $\mathbb{Q}$, the Hasse principle for diagonal cubic hypersurfaces in $\mathbb{P}^{s-1}$ is known unconditionally for $s \geq 7$ [Bak89]. For non-diagonal smooth cubic hypersurfaces in $\mathbb{P}^{s-1}$, it is known unconditionally for $s \geq 9$ [Hoo88].

1.4 Outline and overview of chapters

Recall, from Example 1.2.1, Hooley's conjecture on $M_2(X)$.

Chapter 2 shows that a *slightly deformed* version of Hooley's conjecture would imply that asymptotically 100% of integers $a \not\equiv \pm 4 \bmod 9$ are sums of three cubes. The proof follows [Dia19], up to technical (but important) modifications. There is some flexibility in the choice of deformation; we will show that "HLH for clean pairs $(x_1^3 + \cdots + x_6^3, w)$", in the sense of Definitions 1.4.3 and 1.4.6 below, suffices.

Before proceeding, we first give two convenient general definitions.

Definition 1.4.1. Let $s \in \mathbb{Z}_{\geq 1}$. Given a function $w \colon \mathbb{R}^s \to \mathbb{R}$ and a polynomial $P \in \mathbb{Z}[\boldsymbol{x}] = \mathbb{Z}[x_1, \ldots, x_s]$, let $N_{P,w}(X) := \sum_{\boldsymbol{x} \in \mathbb{Z}^s} w(\boldsymbol{x}/X) \cdot \mathbf{1}_{P(\boldsymbol{x})=0}$ for real $X > 0$. If w is unspecified, we use the *symmetric* box convention $w(\tilde{\boldsymbol{x}}) := \mathbf{1}_{\tilde{\boldsymbol{x}} \in [-1,1]^s}$.

Definition 1.4.2. Let R be a ring. Let $s \in \mathbb{Z}_{\geq 1}$. Given a homogeneous polynomial P that lies in (or maps canonically into) $R[x_1, \ldots, x_s]$, we say P is $\mathbb{P}^{s-1}_R$*-smooth* if its projective zero locus $V_{\mathbb{P}^{s-1}}(P)_{/R}$ is smooth; or equivalently, if P has invertible discriminant in R.

For analytic purposes, the following definition will prove useful:

Definition 1.4.3. Let $s \in \mathbb{Z}_{\geq 1}$. We will always interpret the *support* of a function on $\mathbb{R}^s$ in the *closed* sense. Let $w \colon \mathbb{R}^s \to \mathbb{R}$ be a function; then $\operatorname{Supp} w := \overline{\{\boldsymbol{x} \in \mathbb{R}^s : w(\boldsymbol{x}) \neq 0\}}$. Let $P(x_1, \ldots, x_s)$ be a $\mathbb{P}^{s-1}_{\mathbb{R}}$-smooth homogeneous polynomial. Call the pair (P, w) *smooth* if $\mathbf{0} \notin \operatorname{Supp} w$, and *clean* if $(\operatorname{Supp} w) \cap (\operatorname{hess} P)_{\mathbb{R}}(\mathbb{R}) = \emptyset$; in other words, call (P, w) *smooth* (resp. *clean*) if w is supported away from the locus $x_1 = \cdots = x_s = 0$ in $\mathbb{R}^s$ (resp. the locus $\det(\operatorname{Hess} P) = 0$ in $\mathbb{R}^s$).

Remark 1.4.4. A clean pair (P, w) is smooth, since P is homogeneous.

Example 1.4.5. Let $s \in \mathbb{Z}_{\geq 1}$. Say $P = x_1^3 + \cdots + x_s^3$. Then P is $\mathbb{P}^{s-1}_{\mathbb{R}}$-smooth, and in fact $\mathbb{P}^{s-1}_k$-smooth for every field k of characteristic not dividing 3. Furthermore, (P, w) is clean if and only if w is supported away from the set $\{\boldsymbol{x} \in \mathbb{R}^s : x_1 \cdots x_s = 0\}$.

We now define a singular series, some special loci, some real densities, and a weighted version of HLH. We emphasize that one could attribute the general formulation of HLH (including that in Definition 1.4.6) to [Hoo86a, Conjecture 2], [VW95, Appendix], Manin–Peyre, et al.

Definition 1.4.6. Let $F \in \mathbb{Z}[\boldsymbol{x}] = \mathbb{Z}[x_1, \ldots, x_6]$ be a $\mathbb{P}^5_{\mathbb{Q}}$-smooth 6-variable cubic form. Let $S_{\mathbf{0}}(n) := \sum_{a \in (\mathbb{Z}/n)^\times} \sum_{\boldsymbol{x} \in (\mathbb{Z}/n)^6} e_n(aF(\boldsymbol{x}))$ for $n \in \mathbb{Z}_{\geq 1}$; let $\mathfrak{S}_F := \sum_{n \geq 1} n^{-6} S_{\mathbf{0}}(n)$. Let $C(\mathrm{SSV})$

denote the set of 3-dimensional vector spaces $L/\mathbb{Q}$ such that $F|_L = 0$. Given $L \in C(\mathrm{SSV})$, let $\Lambda := L \cap \mathbb{Z}^6$ denote the unique *primitive* sublattice of $\mathbb{Z}^6$ with $\Lambda \cdot \mathbb{Q} = L$. Then let $L^\perp := \{\boldsymbol{c} \in \mathbb{Q}^6 : \boldsymbol{c} \perp L\}$ and $\Lambda^\perp := \{\boldsymbol{c} \in \mathbb{Z}^6 : \boldsymbol{c} \perp \Lambda\}$ denote the *orthogonal complements* of L and Λ, respectively, with respect to the usual dot product $\boldsymbol{c} \cdot \boldsymbol{x} := c_1x_1 + \cdots + c_6x_6$ (so in particular, $\mathbb{Q} \cdot \Lambda^\perp = L^\perp$). Now choose bases $\boldsymbol{\Lambda}, \boldsymbol{\Lambda}^\perp$ of $\Lambda, \Lambda^\perp$, viewed as 6×3 and 3×6 integer matrices, respectively, so that $\Lambda = \boldsymbol{\Lambda}\mathbb{Z}^3$ and $\Lambda^\perp = \mathbb{Z}^3\boldsymbol{\Lambda}^\perp$ (where we view Λ as a "column space" and $\Lambda^\perp$ as a "row space").

Let $w \in C_c^\infty(\mathbb{R}^6)$ be a smooth compactly supported weight. Let $\sigma_{\infty,F,w} := \lim_{\epsilon \to 0} (2\epsilon)^{-1} \int_{|F(\boldsymbol{x})| \le \epsilon} d\boldsymbol{x}\, w(\boldsymbol{x})$. For each $L \in C(\mathrm{SSV})$, choose $\boldsymbol{\Lambda}^\perp$ as in the previous paragraph, and let $\sigma_{\infty,L^\perp,w} := \lim_{\epsilon \to 0} (2\epsilon)^{-3} \int_{\|\boldsymbol{\Lambda}^\perp \boldsymbol{x}\|_\infty \le \epsilon} d\boldsymbol{x}\, w(\boldsymbol{x})$. Say (F, w) is *Hardy–Littlewood–Hooley* (HLH) if, as $X \to \infty$, we have the asymptotic

$$N_{F,w}(X) = \left(\underbrace{\sigma_{\infty,F,w}\mathfrak{S}_F}_{\text{randomness}} + o_{F,w;X\to\infty}(1) + \sum_{L \in C(\mathrm{SSV})} \underbrace{\sigma_{\infty,L^\perp,w}}_{\text{structure}} \right) \cdot X^3.$$

(Here and elsewhere, we let $o_{S;x \to a}(g)$ denote a quantity f such that the statement "for any real $\epsilon > 0$, we have $|f| \le \epsilon g$ for all x in a neighborhood $I = I(a, \epsilon, S)$ of a" holds. We will find this little-o notation occasionally convenient.)

Remark 1.4.7. Here $\mathfrak{S}_F$ is the usual *singular series*, and $\sigma_{\infty,F,w}, \sigma_{\infty,L^\perp,w}$ are *real densities*, all given in technically convenient forms.[1] Also, in the setting above, $\Lambda = (\Lambda^\perp)^\perp$, i.e. $\Lambda = \{\boldsymbol{x} \in \mathbb{Z}^6 : \boldsymbol{\Lambda}^\perp \boldsymbol{x} = \boldsymbol{0}\}$, so $\sum_{\boldsymbol{x} \in \Lambda} w(\boldsymbol{x}/X) = \sigma_{\infty,L^\perp,w} X^3 + O_{L,w,A}(X^{-A})$, by Poisson summation over Λ (or, at least morally, by the circle method applied to $\boldsymbol{\Lambda}^\perp \boldsymbol{x} = \boldsymbol{0}$). In particular, $\sigma_{\infty,L^\perp,w}$ does not depend on the choice of $\boldsymbol{\Lambda}^\perp$.

Remark 1.4.8. In this thesis, it would be OK to require a *power-saving* error term for HLH in Definition 1.4.6. Also, one could analyze *unweighted* regions (see Appendix B). But at least in Chapter 2, a soft and smooth formulation of HLH has some benefits.

Recall the discussion of (unconditional and conditional) upper bounds on $M_2(X)$, and the connection to the equation $x_1^3 + \cdots + x_6^3 = 0$, from Example 1.2.1. In Chapter 3, we will introduce the delta method, which connects $N_{F,w}(X)$ (for cubic forms F in some generality, at least when (F, w) is smooth) to the local behavior of the intersections $F(\boldsymbol{x}) = \boldsymbol{c} \cdot \boldsymbol{x} = 0$ over $\mathbb{F}_p, \mathbb{Z}_p, \mathbb{R}, \ldots$, as $\boldsymbol{c} \in \mathbb{Z}^6$ and p vary; and especially to certain associated Hasse–Weil L-functions. Chapter 4 discusses two conditional approaches to upper-bounding $M_2(X)$ (or $N_{x_1^3+\cdots+x_6^3,w}(X)$), one based on the delta method and a large-sieve hypothesis (a la Bombieri–Vinogradov), and the other based on a family of ternary quadratic forms; see §4.1 for the former, and §4.2 for the latter. The former is probably more significant, but we include the latter for amusement.

The following remark may provide a holistic view of the rough landscape so far.

Remark 1.4.9 (A cartoon). Let $F_0(\boldsymbol{y}) := y_1^3 + y_2^3 + y_3^3$ first, and $F(\boldsymbol{x}) := x_1^3 + \cdots + x_6^3$ and $\mathcal{W} := V_{\mathbb{A}^6}(F)_{/\mathbb{Z}}$ second. Then the map $\mathbb{A}^1 \leftarrow \mathbb{A}^3$, $F_0(\boldsymbol{y}) \leftarrow\!\shortmid \boldsymbol{y}$ and the diagram

$$\underbrace{\mathbb{A}^3 \xrightarrow{F_0} \mathbb{A}^1 \xleftarrow{F_0} \mathbb{A}^3 \times_{F_0} \mathbb{A}^3}_{\text{Cf. Examples 1.2.1 and 1.2.3}} \cong \underbrace{\mathcal{W} \xleftarrow{\boldsymbol{x}} \{(\boldsymbol{x}, \boldsymbol{c}) \in \mathcal{W} \times \mathbb{A}^6 : \boldsymbol{c} \cdot \boldsymbol{x} = 0\} \xrightarrow{\boldsymbol{c}} \mathbb{A}^6}_{\text{Cf. [Klo26], \ldots, [DFI93,HB96], \ldots}}$$

[1]It is known that $\mathfrak{S}_F$ converges absolutely, and that $\sigma_{\infty,F,w}, \sigma_{\infty,L^\perp,w}$ are finite and well-behaved (though a little care is needed for $\sigma_{\infty,F,w}$ "at the origin" if (F, w) is not smooth).

vaguely depict how when studying certain statistics (in ℓ^1 and ℓ^2), one can "reduce" from the miserly *individual* equations $F_0 = a$ to the more generous *family* of auxiliary *hyperplane sections* $\mathcal{W}_{\boldsymbol{c}} := V_{\mathbb{A}^6}(F, \boldsymbol{c}\cdot\boldsymbol{x})_{/\mathbb{Z}}$ (over $\boldsymbol{c} \in \mathbb{Z}^6 \setminus \{\mathbf{0}\}$)—although the latter will appear "adelically" through $\mathcal{W}_{\boldsymbol{c}}(\mathbf{A}_{\mathbb{Z}})$, rather than "integrally" through $\mathcal{W}_{\boldsymbol{c}}(\mathbb{Z})$.

(The cartoonish "right-hand half" is meant to represent the delta method for the affine zero locus $\mathcal{W} \subseteq \mathbb{A}^6$ of F; see Chapter 3 for details.)

Remark 1.4.10. One could try counting $\mathcal{W}(\mathbb{Z})$ using inclusion-exclusion on $\mathcal{W}_{\boldsymbol{c}}(\mathbb{Z})$ over $\|\boldsymbol{c}\| \ll X^{1/(6-1)} = X^{1/5}$ (in view of Siegel's lemma on linear equations). As far as I can tell, the delta method is distinct—even if it still involves $\mathcal{W}_{\boldsymbol{c}}(\mathbf{A}_{\mathbb{Z}})$.

Chapter 5 discusses, among other things, a near dichotomy between randomness and structure for the point counts of projective cubic threefolds over finite fields (with applications to the aforementioned hyperplane sections $F(\boldsymbol{x}) = \boldsymbol{c}\cdot\boldsymbol{x} = 0$), and raises some further questions in this direction.

Chapter 6 discusses how to extract the main terms of HLH in a natural way, for smooth pairs (F, w) with F diagonal in 6 variables. One of the main inputs is that for $L \in C(\mathrm{SSV})$, the lattices $\Lambda := L \cap \mathbb{Z}^6$ "remain special" in some sense (cf. Chapter 5) for hyperplane sections modulo n.

Chapter 7 discusses [Wan21a]'s new pointwise estimates for exponential sums and oscillatory integrals appearing in the delta method; these estimates hold for various kinds of pairs (F, w). Over primes, the proofs involve Chapter 5; over prime powers, a study of certain arithmetic fourfolds (relative threefolds over $\mathbb{Z}_p$); and over the reals, a critical use of stationary phase beyond that of Hooley and Heath-Brown. The estimates give, for instance (somewhat in the spirit of conjectures of Sarnak and Xue on "naive Ramanujan" failures, but in a different context), a power-saving bound (conditional on limited ranges of the Square-free Sieve Conjecture) on the frequency of certain "square-root cancellation" failures.

Chapter 8 discusses [Wan21a]'s conditional proof that HLH holds for clean pairs (F, w) with F diagonal in 6 variables (and thus, by Chapter 2, that the "100% Hasse principle" holds for sums of three cubes); the proof is conditional on some standard number-theoretic conjectures—the main additions (relative to Hooley and Heath-Brown) being conjectures of Random Matrix Theory (RMT) and Square-free Sieve type. The proof builds on Chapters 6 and 7. By Chapter 6, it suffices to bound the contribution Σ from smooth hyperplane sections in the delta method. We first decompose Σ adelically into discriminant-based pieces; we then conditionally estimate some of these pieces via local calculations and Poisson summation (among other ingredients), and conditionally bound other pieces via Hölder's inequality between good and bad moduli factors (among other ingredients).

Remark 1.4.11. Recall Observation 1.3.2. Our methods for g_3 would surely extend to conditionally show that $\mathcal{E}(h)$ has density 0. These methods do not shed much further light on the true sizes of such sets (a much deeper question). But the statement "$\mathcal{E}(-)$ has density 0" certainly *cannot* be improved to "$\mathcal{E}(-) = \emptyset$" in general.

Chapter 9 discusses questions related to or inspired by the work above.

We will often refer to other papers for details, especially for certain proofs, but not at the expense of the overall story.

Chapter 2

Approximate variances

2.1 Introduction

Definition 2.1.1. Given $a \in \mathbb{Z}$, let $F_a(\boldsymbol{y}) = F_a(y_1, y_2, y_3) := y_1^3 + y_2^3 + y_3^3 - a$, and let $r_3(a) := \#\{\boldsymbol{y} \in \mathbb{Z}_{\geq 0}^3 : F_a(\boldsymbol{y}) = 0\}$, so that $r_3(a) \leq N_{F_a}(a^{1/3})$.

Using the convenient structure $F_a = F_0 - a$, we will first estimate $r_3(a)$ "coarsely" in ℓ^1, ℓ^2 over $a \leq B$, as $B \to \infty$—following a classical strategy (cf. Remark 1.4.9). Later, to obtain a more precise result (Theorem 2.1.8), we will choose "nice" weights $\nu \colon \mathbb{R}^3 \to \mathbb{R}$ (cf. [Hoo16]) and bound $N_{F_a,\nu}(X)$ in "approximate variance" over $a \ll X^3$, as $X \to \infty$—essentially following [GS17, Dia19], up to smoothness.

Recall that [GS17] works with the equations $x^2 + y^2 + z^2 - xyz = a$, which are "critical" just like the equations $x^3 + y^3 + z^3 = a$ considered here and in [Dia19]. It is worth noting that there are significant technical differences between [GS17] and [Dia19] (even though the overall arguments are formally similar), such as the following:

(1) the nature of "exceptional" parametric solutions differs between the two;

(2) delicate issues involving binary quadratic forms arise in [GS17], but not in [Dia19]; and

(3) [GS17] engineers close-to-classical regions (for a family of quadratic equations), while [Dia19] engineers far-from-classical regions (for a single cubic equation).

2.1.1 A moment framework

Recall the well-known *first* moment

$$\sum_{a \leq B} r_3(a) = \#\{\boldsymbol{y} \in \mathbb{Z}_{\geq 0}^3 : F_0(\boldsymbol{y}) \leq B\} \propto B + o_{B \to \infty}(B) \qquad \text{as } B \to \infty,$$

proven by writing $F_a = F_0 - a$, expanding $r_3(a)$ as a "vertical sum" along a fiber of $F_0 \colon \mathbb{Z}^3 \to \mathbb{Z}$, and then approximating the "total space" by a continuous volume.

In particular, $\mathbb{E}_{a \leq B}[r_3(a)] \asymp 1$ holds for all $B > 0$. Now recall that [HL25] formulated the ℓ^∞ hypothesis $r_3(a) \ll_\epsilon a^\epsilon$ for $a > 0$ (*Hypothesis K*); but [Mah36] showed, via the identity $(9u^4)^3 + (3uv^3 - 9u^4)^3 + (v^4 - 9u^3v)^3 = v^{12}$, that $r_3(a) \gg a^{1/12}$ holds for twelfth powers $a = v^{12} > 0$.

Remark 2.1.2. Mahler's identity can be viewed as a "clever specialization" of a "$\mathbb{Q}$-unirational parameterization" of $x^3 + y^3 + z^3 = w^3$. In fact, Mahler's identity can be recovered from [Elk01], which parameterizes *all* $\mathbb{Q}$-points on $x^3 + y^3 + z^3 = w^3$.

By Mahler's construction, $r_3(a)$ can be very large for some *individual* a's, but we can still hope for reasonable statistical behavior as long as $r_3(a)$ is not "too large too often"—a notion most readily formalized by taking higher moments of $r_3(a)$.

It seems difficult at present to rigorously analyze the third moment or higher. (The kth moment is connected to the equation $x_1^3 + y_1^3 + z_1^3 = x_2^3 + y_2^3 + z_2^3 = \cdots = x_k^3 + y_k^3 + z_k^3$, which can be viewed as a system of $k-1$ Diophantine equations in $3k$ variables. In light of Mahler's "special" a's, even *formulating reasonable conjectures* for arbitrarily high moments seems difficult, but see [DHL06] and [Dia19, §4] for plausible random models of sets like $\{x^3 + y^3 + z^3 : (x, y, z) \in \mathbb{Z}^3_{\geq 0}\}$.) But the *second* moment of $r_3(a)$ forms, by double counting, a relatively simple "Diophantine sandwich"

$$\sum_{a \leq B} r_3(a)^2 \leq N_F(B^{1/3}) = \int_{\mathbb{R}/\mathbb{Z}} d\theta\, |T(\theta)|^6$$
$$\ll \int_{\mathbb{R}/\mathbb{Z}} d\theta\, (|T_{>0}(\theta)|^6 + |T_{\leq 0}(\theta)|^6) \ll \sum_{a \leq 3B} r_3(a)^2$$

(unconditionally), where $F := x_1^3 + \cdots + x_6^3$ and $T_S(\theta) := \sum_{|x| \leq B^{1/3}} e(\theta x^3) \cdot \mathbf{1}_{x \text{ satisfies } S}$. And indeed, [HL25] really only applied the ℓ^∞ Hypothesis K through the statement

$$\sum_{a \leq B} r_3(a)^2 \ll_\epsilon B^{1+\epsilon} \qquad \text{as } B \to \infty$$

(known conditionally, as discussed in Example 1.2.1), a weaker ℓ^2 hypothesis (termed *Hypothesis* K^* by [Hoo97]) equivalent to "$N_F(B^{1/3}) \ll_\epsilon B^{1+\epsilon}$ as $B \to \infty$".

Now we can combine the ℓ^1 expectation $\mathbb{E}_{a \leq B}[r_3(a)] \asymp 1$ with the "ℓ^2 data" N_F, to get the following classical result (essentially already mentioned in Example 1.2.1):

Observation 2.1.3 (Second moment method). $\#\{a \leq B : r_3(a) \neq 0\} \gg B^2/N_F(B^{1/3})$.

Proof. By Cauchy, $\sum_{a \leq B} r_3(a)^2 \geq \left(\sum_{a \leq B} r_3(a)\right)^2 / \#\{a \leq B : r_3(a) \neq 0\}$. (If there are *too few* $a \in \mathbb{Z}$ with $r_3(a) \neq 0$, then the second moment is *forced to be large.* This idea comes in many forms, e.g. the probabilistic Chung–Erdős inequality.) □

In particular, if $N_F(B^{1/3}) \ll B$ were known, then Observation 2.1.3 would imply that $\{a \in \mathbb{Z} : r_3(a) \neq 0\} = F_0(\mathbb{Z}^3_{\geq 0})$ has *positive lower density.* Or if more precise estimates for $N_{F,w}(X)$ were known in sufficient generality, then it would follow that $F_0(\mathbb{Z}^3)$ has density 7/9 in $\mathbb{Z}$, essentially by [Dia19]—which we now introduce.

2.1.2 The need for increasingly lopsided regions

To prove, *under our methods* (based on [Dia19]), that $F_0 = a$ "almost always" satisfies the Hasse principle, it will not suffice to statistically analyze the sequence $a \mapsto N_{F_a,\nu}(X)$ (over a range of the form $a \ll_\nu X^3$) for only a *single fixed* weight ν. The following remark more or less explains why.

Remark 2.1.4. The upper density of $\{a \in \mathbb{Z}_{\geq 0} : r_3(a) \neq 0\}$ is at most $\Gamma(4/3)^3/6 \approx 0.12$ [Dav39]. In fact, say for each $a \in \mathbb{Z}$ we restrict to $\boldsymbol{y} \in a^{1/3} \cdot \Omega$, with $\Omega \subseteq \mathbb{R}^3$ *fixed* and bounded. Then as $|a| \to \infty$, the equation $F_a(\boldsymbol{y}) = 0$ *fails Hasse* over a relative density 0.99 subset of $a \in \mathbb{Z}$ lying in *some* arithmetic progression depending on Ω [Dia19, §1].

Thus we will instead let ν *vary*. Specifically, certain *increasingly skewed* regions (cf. the cuspidal regions in [Dia19, p. 26, Remark with pictures]) can help, as we now explain.

Observation 2.1.5. Fix an "inhomogeneity parameter" $A \geq 1$. If $X \gg_A 1$ is sufficiently large, then the number of triples $\boldsymbol{y} \in \mathbb{Z}^3$ with $AX \leq |y_1|, |y_2|, |y_3| \leq 2AX$ and $|F_0(\boldsymbol{y})| \leq X^3$ is $\Theta(X^3)$.

Proof sketch. There are $\asymp (AX)^3$ triples $\boldsymbol{y} \in \mathbb{Z}^3$ with $|y_1|, |y_2|, |y_3| \in [AX, 2AX]$. The event $|F_0(\boldsymbol{y})| \leq X^3$ occurs with probability $\asymp X^3/(AX)^3 = 1/A^3$ among these triples, provided that $AX \gg A^3$ is sufficiently large and $0 \in F_0((1, 2), (1, 2), (-2, -1))$. □

Along these lines, Heath-Brown conjectured the following:

Conjecture 2.1.6 ([HB92, p. 623])**.** *If* $a \not\equiv \pm 4 \bmod 9$*, then* $\lim_{X\to\infty} N_{F_a}(X) = \infty$*.*

Conjecture 2.1.6 for *individual* $a \not\equiv \pm 4 \bmod 9$ might be very hard (if true), even conditionally, so we will content ourselves with a (conditional) *average* study, which we now set up.

2.1.3 A sketch of an approximate variance framework

Given ν, a, X, one can define certain densities $\sigma_{p,F_a}, \sigma_{\infty,F_a,\nu}(X)$. (For details, see §2.2, which begins shortly below.) For $a \in \mathbb{Z}$, informally write $\mathfrak{S}_{F_a} := \prod_{p<\infty} \sigma_{p,F_a}$, for expository purposes. Now consider the naive *Hardy–Littlewood prediction* $\mathfrak{S}_{F_a} \cdot \sigma_{\infty,F_a,\nu}(X)$ for $N_{F_a,\nu}(X)$. Smaller moduli should have a greater effect in $\mathfrak{S}_{F_a}$; furthermore, $\mathfrak{S}_{F_a}$ itself—*as is*—can be subtle (involving the behavior of L-functions at $s = 1$; cf. [DSP90, Theorem 3 and its proof] and [HB96, Theorems 6–7]). So in the following definition, we work with a "restricted" version of $\mathfrak{S}_{F_a} \cdot \sigma_{\infty,F_a,\nu}(X)$.

Definition 2.1.7 (Cf. [Dia19, §3])**.** Fix $\nu\colon \mathbb{R}^3 \to \mathbb{R}$ and let $w(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{z}}) := \nu(\tilde{\boldsymbol{y}})\nu(-\tilde{\boldsymbol{z}})$. Given $M \geq 1$, let $K = K(M) := \prod_{p\leq M} p^{\lfloor \log_p M \rfloor}$ and $s_{F_a}(K) := K^{-2} \cdot \#\{\boldsymbol{y} \in (\mathbb{Z}/K)^3 : F_a(\boldsymbol{y}) = 0\}$. Then define the M*-approximate* ("finite-precision") *variance*

$$\operatorname{Var}(X, M) := \sum_{a\in\mathbb{Z}} [N_{F_a,\nu}(X) - s_{F_a}(K)\sigma_{\infty,F_a,\nu}(X)]^2 =: \Sigma_1 - 2\Sigma_2 + \Sigma_3.$$

The most interesting sum among $\Sigma_1, \Sigma_2, \Sigma_3$ is $\Sigma_1 := \sum_{a\in\mathbb{Z}} N_{F_a,\nu}(X)^2$, which can be rewritten as $N_{x_1^3+\cdots+x_6^3,w}(X)$. In fact, the main ideas of [Dia19, §§2–3] lead to the following result:

Theorem 2.1.8 (Cf. [Dia19, Theorem 3.3])**.** *Suppose* $(x_1^3 + \cdots + x_6^3, w)$ *is HLH (in the sense of Definition 1.4.6) for* every fixed *choice of* $w \in C_c^\infty(\mathbb{R}^6)$ *with* $(x_1^3 + \cdots + x_6^3, w)$ *clean. Then asymptotically* 100% *of integers* $a \not\equiv \pm 4 \bmod 9$ *are sums of three cubes.*

For the proof—differing from [Dia19] only in technical aspects—see §2.2.

2.2 Details

2.2.1 Defining the local densities of individual fibers

Recall the notation $F_a(\boldsymbol{y}) := F_0(\boldsymbol{y}) - a$, where $F_0(\boldsymbol{y}) := y_1^3 + y_2^3 + y_3^3$. We first define certain non-archimedean densities.

Definition 2.2.1. For $a \in \mathbb{Z}_p$, let $\sigma_{p,F_a} := \lim_{l\to\infty}(p^{-2l} \cdot \#\{\boldsymbol{y} \in (\mathbb{Z}/p^l)^3 : F_a(\boldsymbol{y}) = 0\})$.

Next, we define certain real densities (analogously to the p-adic densities σ_{p,F_a}), by "ϵ-thickening" parallel to the 0-level set (i.e. the fiber over 0) of the map $F_a\colon \mathbb{R}^3 \to \mathbb{R}$.

Definition 2.2.2. Fix $\nu \in C_c^\infty(\mathbb{R}^3)$ with (F_0, ν) smooth (in the sense of Definition 1.4.3). Now fix $X \in \mathbb{R}_{>0}$. Then for $(\boldsymbol{y}, a) \in \mathbb{R}^3 \times \mathbb{R}$, write $\tilde{\boldsymbol{y}} := \boldsymbol{y}/X$ and $\tilde{a} := a/X^3$. Also, for all $a \in \mathbb{R}$, let

$$\sigma_{\infty,F_a,\nu}(X) := \lim_{\epsilon\to 0} (2\epsilon)^{-1} \int_{|F_a(\boldsymbol{y})/X^3|\le\epsilon} d(\boldsymbol{y}/X)\, \nu(\boldsymbol{y}/X).$$

Observation 2.2.3. Here $\sigma_{\infty,F_a,\nu}(X) = \sigma_{\infty,F_{\tilde{a}},\nu}(1)$. (Indeed, by definition, we have $\sigma_{\infty,F_a,\nu}(X) = \lim_{\epsilon\to 0} (2\epsilon)^{-1} \int_{|F_0(\tilde{\boldsymbol{y}})-\tilde{a}|\le\epsilon} d\tilde{\boldsymbol{y}}\, \nu(\tilde{\boldsymbol{y}}) = \sigma_{\infty,F_{\tilde{a}},\nu}(1)$.)

Remark 2.2.4. In particular, for $a \ne 0$ at least, $\sigma_{\infty,F_a,\nu}(a^{1/3}) = \sigma_{\infty,F_1,\nu}(1)$ is *constant*. Essentially for this reason, perhaps, many references do not seem to treat real densities very thoroughly.

Before proceeding, we make two conceptual remarks on real densities, for completeness.

Remark 2.2.5. The function $\tilde{a} \mapsto \sigma_{\infty,F_a,\nu}(X)$ is supported on a bounded range $\tilde{a} \ll_\nu 1$, on which the ϵ-limit converges *uniformly* at a rate depending only on F_0, ν. Generally, if α is "nice" and $\alpha_\epsilon := \alpha(t/\epsilon)$ (e.g. $\alpha_\epsilon(t) = \mathbf{1}_{|t|\le\epsilon}$ above), then

$$\frac{\int_{\mathbb{R}^3} d\tilde{\boldsymbol{y}}\, \nu(\tilde{\boldsymbol{y}})\alpha_\epsilon(F_0(\tilde{\boldsymbol{y}}) - \tilde{a})}{\int_{\mathbb{R}} dt\, \alpha_\epsilon(t)} = \sigma_{\infty,F_a,\nu}(X) + O_{F_0,\nu,\alpha}(\epsilon).$$

Remark 2.2.6. Here $d\tilde{a}\ \sigma_{\infty,F_a,\nu}(X) = (F_0)_*(d\tilde{\boldsymbol{y}}\ \nu(\tilde{\boldsymbol{y}}))$ as a pushforward measure, so that $\int_{\mathbb{R}} d\tilde{a}\, \sigma_{\infty,F_a,\nu}(X) g(\tilde{a}) = \int_{\mathbb{R}^3} d\tilde{\boldsymbol{y}}\, \nu(\tilde{\boldsymbol{y}}) g(F_0(\tilde{\boldsymbol{y}}))$ holds for all "nice" $g\colon \mathbb{R} \to \mathbb{R}$.

For convenience (when working with $\sigma_{\infty,F_a,\nu}(X)$), we now observe the following:

Observation 2.2.7. Define $\pi_1\colon \mathbb{R}^3 \to \mathbb{R}$ by $\boldsymbol{y} \mapsto y_1$. Fix $\nu \in C_c^\infty(\mathbb{R}^3)$ such that $0 \notin \pi_1(\operatorname{Supp}\nu)$. Now fix $(a, X) \in \mathbb{R} \times \mathbb{R}_{>0}$. Then a change of variables from $\tilde{y}_1$ to $F_0 := F_0(\tilde{\boldsymbol{y}})$ proves

$$\sigma_{\infty,F_a,\nu}(X) = \int_{\mathbb{R}^2} d\tilde{y}_2\, d\tilde{y}_3\, \nu(\tilde{\boldsymbol{y}}) \cdot \underbrace{|\partial\tilde{y}_1/\partial F_0|_{F_0=\tilde{a}}}_{\text{local } \tilde{y}_1\text{-density}} = \int_{\mathbb{R}^2} d\tilde{y}_2\, d\tilde{y}_3\, \nu(\tilde{\boldsymbol{y}}) \cdot \underbrace{|\partial F_0/\partial\tilde{y}_1|}_{\text{escape rate}}{}_{F_0=\tilde{a}}^{-1},$$

where $\partial F_0/\partial\tilde{y}_1 = 3\tilde{y}_1^2 \gg_\nu 1$ over the support of the integrand.

Proof. Cf. [HB96, proof of Lemma 11]. $\square$

At least in the absence of better surface coordinates, the earlier "ϵ-thickening" still provides greater intuition, while the surface integral allows for effortless rigor.

The following technical bound will come up later, when integrating by parts.

Proposition 2.2.8. *Fix $\nu \in C_c^\infty(\mathbb{R}^3)$ with (F_0, ν) smooth. Then for integers $k \geq 0$, we have $\partial_{\tilde{a}}^k[\sigma_{\infty,F_a,\nu}(X)] \ll_{k,\nu} 1$, uniformly over $(a, X) \in \mathbb{R} \times \mathbb{R}_{>0}$.*

Proof when $0 \notin \pi_1(\operatorname{Supp}\nu)$. Consider the surface integral representation of $\sigma_{\infty,F_a,\nu}(X)$ from Observation 2.2.7. The integrand vanishes unless $\tilde{a}, \tilde{y}_2, \tilde{y}_3 \ll_\nu 1$. Fix $\tilde{y}_2, \tilde{y}_3 \ll_\nu 1$, and let $\tilde{y}_1$ vary with $\tilde{a}$ according to $F_0(\tilde{\boldsymbol{y}}) = \tilde{a}$. Then $\partial_{\tilde{a}}[\tilde{y}_1] = (3\tilde{y}_1^2)^{-1} \ll_\nu 1$. Now repeatedly apply $\partial_{\tilde{a}}$ to the integrand (using Leibniz and the chain rule). □

Proof in general. Use a suitable partition of unity. □

2.2.2 Interpreting the approximate variance

Let $\boldsymbol{y}, \boldsymbol{z}$ denote 3-vectors, and define $x_i := y_i$ for $i \in [3]$ and $x_{i+3} := z_i$ for $i \in [3]$. Let $\tilde{\boldsymbol{y}} := \boldsymbol{y}/X$, etc. Now fix $\nu \in C_c^\infty(\mathbb{R}^3)$ with (F_0, ν) smooth. Recall, from Definition 2.1.7, the definitions of $w, K(M), s_{F_a}(K), \operatorname{Var}(X, M), \Sigma_1, \Sigma_2, \Sigma_3$.

By double counting (and the "symmetric" ν-factorization of w), the ℓ^2 moment $\Sigma_1 := \sum_{a\in\mathbb{Z}} N_{F_a,\nu}(X)^2$ equals $N_{F,w}(X)$, where $F := x_1^3 + \cdots + x_6^3$.

Also, by analogy with standard probability theory, we expect an ℓ^1 calculation to show that $\Sigma_2 \approx \Sigma_3$. However, a rigorous proof (of such a fact) takes some nontrivial work, since $s_{F_a}(K)\sigma_{\infty,F_a,\nu}(X)$ *varies* with a. To begin, write

$$\Sigma_3 := \sum_{a\in\mathbb{Z}} [s_{F_a}(K)\sigma_{\infty,F_a,\nu}(X)]^2 = \sum_{b \bmod K} s_{F_b}(K)^2 \sum_{a\equiv b \bmod K} \sigma_{\infty,F_a,\nu}(X)^2$$

by collecting along fibers of $\mathbb{Z} \to \mathbb{Z}/K$, and write

$$\begin{aligned}\Sigma_2 &:= \sum_{a\in\mathbb{Z}} N_{F_a,\nu}(X) s_{F_a}(K)\sigma_{\infty,F_a,\nu}(X) \\ &= \sum_{\boldsymbol{z}\in\mathbb{Z}^3} \nu(\boldsymbol{z}/X) s_{F_{F_0(\boldsymbol{z})}}(K)\sigma_{\infty,F_{F_0(\boldsymbol{z})},\nu}(X) \\ &= \sum_{\boldsymbol{d} \bmod K} s_{F_{F_0(\boldsymbol{d})}}(K) \sum_{\boldsymbol{z}\equiv\boldsymbol{d} \bmod K} \nu(\boldsymbol{z}/X)\sigma_{\infty,F_{F_0(\boldsymbol{z})},\nu}(X)\end{aligned}$$

by expanding $N_{F_a,\nu}(X)$ along fibers of $F_0\colon \mathbb{Z}^3 \to \mathbb{Z}$ and collecting along $\mathbb{Z}^3 \to (\mathbb{Z}/K)^3$.

To simplify Σ_2, Σ_3 further, we will use Poisson summation, as well as the following observation:

Observation 2.2.9. The "pure L^2 moment" $\int_{a\in\mathbb{R}} d\tilde{a}\, \sigma_{\infty,F_a,\nu}(X)^2$ and the "mixed L^1 moment" $\int_{\tilde{\boldsymbol{z}}\in\mathbb{R}^3} d\tilde{\boldsymbol{z}}\, \nu(\tilde{\boldsymbol{z}})\sigma_{\infty,F_{F_0(\boldsymbol{z})},\nu}(X)$ both simplify to $\sigma_{\infty,F,w}$.

Proof when $0 \notin \pi_1(\operatorname{Supp}\nu)$. First, $F_a := F_0 - a$, so $\int_{a\in\mathbb{R}} d\tilde{a}\, \sigma_{\infty,F_a,\nu}(X)^2$ expands (via Observation 2.2.7) to

$$\int_{\mathbb{R}^4} d\tilde{y}_2 \cdots d\tilde{z}_3 \int_{\tilde{y}_1\in\mathbb{R}} dF_0(\tilde{\boldsymbol{y}})\, \frac{\nu(\tilde{\boldsymbol{y}})\nu(\tilde{\boldsymbol{z}})}{\partial_1 F_0|_{\tilde{\boldsymbol{y}}}} (\partial_1 F_0|_{\tilde{\boldsymbol{z}}})^{-1}|_{F_0(\tilde{\boldsymbol{y}})=F_0(\tilde{\boldsymbol{z}})},$$

which simplifies to $\int_{\mathbb{R}^5} d\tilde{y}_2 \cdots d\tilde{y}_1\, w(\tilde{\boldsymbol{x}}) \cdot (\partial_1 F_0|_{\tilde{\boldsymbol{z}}})^{-1}|_{F(\tilde{\boldsymbol{x}})=0} = \sigma_{\infty,F,w}$.

Second, $F_{F_0(\boldsymbol{z})} = F_0(\boldsymbol{y}) - F_0(\boldsymbol{z})$, so by Observation 2.2.7,

$$\int_{\tilde{\boldsymbol{z}}\in\mathbb{R}^3} d\tilde{\boldsymbol{z}}\, \nu(\tilde{\boldsymbol{z}})\sigma_{\infty,F_{F_0(\boldsymbol{z})},\nu}(X) = \int_{\mathbb{R}^3\times\mathbb{R}^2} d\tilde{\boldsymbol{z}}\, d\tilde{y}_2\, d\tilde{y}_3\, \nu(\tilde{\boldsymbol{y}})\nu(\tilde{\boldsymbol{z}}) \cdot (\partial_1 F_0|_{\tilde{\boldsymbol{y}}})^{-1}|_{F_0(\tilde{\boldsymbol{y}})=F_0(\tilde{\boldsymbol{z}})},$$

which again simplifies to $\sigma_{\infty,F,w}$. □

Proof in general. Argue in terms of ϵ-thickenings. Alternatively, generalize to a "bilinear" statement (involving *two* weights ν_1, ν_2, rather than just one); then reduce the bilinear statement to a surface integral computation (based on a general version of Observation 2.2.7), after taking suitable partitions of unity. □

Proposition 2.2.10 (Cf. [Dia19, proof of Lemma 3.1])**.** *Uniformly over X, K and $b, \boldsymbol{d} \bmod K$ with $X \geq K \geq 1$, we have*

$$\sum_{a\equiv b \bmod K} \sigma_{\infty,F_a,\nu}(X)^2 = K^{-1}X^3\sigma_{\infty,F,w} + O_{j,\nu}((X^3/K)^{-j})$$

and

$$\sum_{\boldsymbol{z}\equiv \boldsymbol{d} \bmod K} \nu(\boldsymbol{z}/X)\sigma_{\infty,F_{F_0(\boldsymbol{z})},\nu}(X) = K^{-3}X^3\sigma_{\infty,F,w} + O_{j,\nu}((X/K)^{-j}).$$

Proof sketch. By Poisson summation over $K\cdot\mathbb{Z}$ and $K\cdot\mathbb{Z}^3$, the sums are approximately

$$K^{-1}\int_{a\in\mathbb{R}} da\, \sigma_{\infty,F_a,\nu}(X)^2 = K^{-1}X^3\sigma_{\infty,F,w}$$

and

$$K^{-3}\int_{\boldsymbol{z}\in\mathbb{R}^3} d\boldsymbol{z}\, \nu(\boldsymbol{z}/X)\sigma_{\infty,F_{F_0(\boldsymbol{z})},\nu}(X) = K^{-3}X^3\sigma_{\infty,F,w},$$

respectively (where we have simplified the integrals using the previous observation), up to "errors" (i.e. "off-center contributions") of $\ll_{\nu,j} K^{-1}X^3(X^3/K)^{-j}$ and $\ll_{\nu,j} K^{-3}X^3(X/K)^{-j}$, respectively. (For proof, bound the first "off-center contribution" *absolutely* by

$$\sum_{c\neq 0} K^{-1}\left|\int_{a\in\mathbb{R}} da\, \sigma_{\infty,F_a,\nu}(X)^2 e(-c\cdot a/K)\right|;$$

then plug in $a = X^3\tilde{a}$, repeatedly integrate by parts in $\tilde{a}$, and invoke Proposition 2.2.8. The second "off-center contribution" is similar.) □

It follows from the previous proposition (and the trivial bound $|s_{F_b}(K)| \leq K$) that

$$\Sigma_3 = \mathcal{S}(K)\cdot\sigma_{\infty,F,w}X^3 + K^3\cdot O_{j,\nu}((X^3/K)^{-j}),$$

where $\mathcal{S}(K) := K^{-1}\sum_{b \bmod K} s_{F_b}(K)^2 = K^{-5}\cdot\#\{\boldsymbol{x}\in(\mathbb{Z}/K)^6 : F(\boldsymbol{x}) = 0\}$.

Similarly,

$$K^{-3}\sum_{\boldsymbol{d} \bmod K} s_{F_{F_0(\boldsymbol{d})}}(K) = K^{-3}\sum_{b \bmod K}[K^2\cdot s_{F_b}(K)]\cdot s_{F_b}(K) = \mathcal{S}(K)$$

by collecting along fibers of $F_0\colon (\mathbb{Z}/K)^3 \to \mathbb{Z}/K$, so

$$\Sigma_2 = \mathcal{S}(K)\cdot\sigma_{\infty,F,w}X^3 + K^4\cdot O_{j,\nu}((X/K)^{-j}).$$

Lemma 2.2.11 (Cf. [Dia19, Lemma 3.1]). *Uniformly over $X, M \ge 1$ with $X \ge K(M) \ge 1$, we have $\mathcal{S}(K) = \mathfrak{S}_F + O_\epsilon(M^{-2/3+\epsilon})$ and*

$$\operatorname{Var}(X, M) = N_{F,w}(X) - \mathcal{S}(K) \cdot \sigma_{\infty,F,w} X^3 + K^4 \cdot O_{j,\nu}((X/K)^{-j}).$$

Proof. For the second part, use $\operatorname{Var}(X, M) = \Sigma_1 - 2\Sigma_2 + \Sigma_3$. For the first part, note that $p^{6v}\mathcal{S}(p^v) = \sum_{a \in \mathbb{Z}/p^v} \sum_{\boldsymbol{x} \in (\mathbb{Z}/p^v)^6} e_{p^v}(aF(\boldsymbol{x})) = \sum_{l \in [0,v]} p^{6(v-l)} S_{\mathbf{0}}(p^l)$ (cf. [Dav05, Lemma 5.3]), whence $\mathcal{S}(K) = \sum_{n \mid K} n^{-6} S_{\mathbf{0}}(n)$. Yet $\mathfrak{S}_F = \sum_{n \ge 1} n^{-6} S_{\mathbf{0}}(n)$ (by Definition 1.4.6), and $n \mid K$ for all $n \in [1, M]$, so

$$|\mathcal{S}(K) - \mathfrak{S}_F| \le \sum_{n > M} n^{-6} |S_{\mathbf{0}}(n)| \ll_\epsilon M^{(4-6)/3+\epsilon} = M^{-2/3+\epsilon}$$

by standard bounds (e.g. Lemma 3.4.1 below), as desired. □

Consequently, *if* (F, w) is HLH (in the sense of Definition 1.4.6), then

$$\begin{aligned}\operatorname{Var}(X, M) &= [N_{F,w}(X) - \mathfrak{S}_F \cdot \sigma_{\infty,F,w} X^3] + O(M^{-20/31} \cdot \sigma_{\infty,F,w} X^3) + \frac{O_\nu(K^4)}{(X/K)^{100}} \\ &= \left[o_{\nu; X \to \infty}(1) + \sum_{L \in C(\mathrm{SSV})} \sigma_{\infty,L^\perp,w}\right] X^3 + O\left(\frac{\sigma_{\infty,F,w} X^3}{M^{20/31}}\right) + \frac{O_\nu(K^4)}{(X/K)^{100}},\end{aligned}$$

the most interesting term being the diagonal-type contribution $\sum_{L \in C(\mathrm{SSV})} \sigma_{\infty,L^\perp,w} X^3$.

2.2.3 Applying increasingly cuspidal weights

The preceding analysis applies to arbitrary $\nu \in C_c^\infty(\mathbb{R}^3)$ with (F_0, ν) smooth. Now we finally choose specific ν's. Fix a *nonnegative, even* weight $w_0 \in C_c^\infty(\mathbb{R})$ with $w_0|_{[-1,1]} \ge 1$, and a *nonnegative* weight $D \in C_c^\infty(\mathbb{R}_{>0})$ with $D|_{[1,2]} \ge 1$.

Definition 2.2.12. Given $A_0 \in \mathbb{R}_{\ge 1}$, set

$$\nu(\tilde{\boldsymbol{y}}) = \nu_{w_0,D,A_0}(\tilde{\boldsymbol{y}}) := w_0(F_0(\tilde{\boldsymbol{y}})) \int_{A \in [1,A_0]} d^\times A \prod_{i \in [3]} D(|\tilde{y}_i|/A),$$

where $d^\times A := dA/A$. Then set $w(\tilde{\boldsymbol{x}}) := \nu(\tilde{\boldsymbol{y}})\nu(-\tilde{\boldsymbol{z}}) = \nu(\tilde{\boldsymbol{y}})\nu(\tilde{\boldsymbol{z}})$.

Remark 2.2.13. The integral over A *enlarges the search space* for representations of numbers by F_0. Essentially, we search among $\boldsymbol{y}$ with $X \ll |y_1| \asymp |y_2| \asymp |y_3| \ll A_0 X$ and $|F_0(\boldsymbol{y})| \ll X^3$. It is important that the special locus $y_1 + y_2 = 0$ (with $\tilde{y}_3^3 = F_0(\tilde{\boldsymbol{y}}) \in \operatorname{Supp} w_0$) does not accumulate when integrated over A. Cf. [Dia19, discussion on p. 24]. (By "switching" $\int, \prod$, we could instead consider all $\boldsymbol{y}$ with $X \ll |y_1|, |y_2|, |y_3| \ll A_0 X$ and $|F_0(\boldsymbol{y})| \ll X^3$, but then we would also need to "manually" restrict to $|y_1 + y_2| \gg X$, etc. Cf. [Dia19, p. 24, definition of R_N^*].)

Remark 2.2.14. In view of Example 1.4.5, the use of the "dyadic" weight D ensures that the pairs (F_0, ν) and (F, w) are not only smooth, but also clean.

Remark 2.2.15. Pointwise, $|\nu(\tilde{\boldsymbol{y}})| \ll \int_{\mathbb{R}_{>0}} d^\times A\, D(|\tilde{y}_1|/A) = \|D \circ \log\|_{L^1(\mathbb{R})} \ll 1$, and in general $\|\nu\|_{k,\infty} \ll_k 1$ (i.e. the derivatives of ν of order $\le k$ are uniformly bounded). On the other hand, $\operatorname{vol}(\operatorname{Supp}\nu) \asymp \log A_0$ truly *grows* with A_0. (To prove $\operatorname{vol}(\operatorname{Supp}\nu) \ll \log A_0$, fix $A \in [1, A_0]$ and $|\tilde{y}_1|, |\tilde{y}_2| \in [A, 2A]$, and note that $\{\tilde{y}_3 \in \pm[A, 2A] : F_0(\tilde{\boldsymbol{y}}) \in \operatorname{Supp} w_0\}$ has length $\ll 1/A^2$. To prove $\operatorname{vol}(\operatorname{Supp}\nu) \gg \log A_0$, use similar but more careful ideas, based on the fact that $w_0|_{[-1,1]} \ge 1$ and $D|_{[1,2]} \ge 1$.)

Why require $w_0|_{[-1,1]} \ge 1$ and $D|_{[1,2]} \ge 1$? Recall that

$$\sigma_{\infty,F_a,\nu}(X) = \sigma_{\infty,F_{\tilde{a}},\nu}(1) = \lim_{\epsilon\to 0} (2\epsilon)^{-1} \int_{|F_0(\tilde{\boldsymbol{y}})-\tilde{a}|\le\epsilon} d\tilde{\boldsymbol{y}}\, \nu(\tilde{\boldsymbol{y}}).$$

Now fix ϵ, plug in the definition of ν, switch the order of $\tilde{\boldsymbol{y}}, A$, and fix $A \in [1, A_0]$. If $|\tilde{a}| \le 1$ and $\epsilon \le 0.1$, then $w_0(F_0(\tilde{\boldsymbol{y}})) \prod_{i\in[3]} D(|\tilde{y}_i|/A) \ge 1$ certainly holds on the set

$$S_{\epsilon,A,\tilde{a}} := \{\tilde{\boldsymbol{y}} \in \mathbb{R}^3 : |\tilde{y}_i| \in [A, 2A] \text{ and } F_0(\tilde{\boldsymbol{y}}) \in [\tilde{a}-\epsilon, \tilde{a}+\epsilon] \cap [-1, 1]\},$$

and furthermore, $\operatorname{vol}(S_{\epsilon,A,\tilde{a}}) \gg A^3 \cdot (\epsilon/A^{\deg F_0}) = \epsilon$ (as one can prove by restricting attention to $\tilde{y}_1, \tilde{y}_2 \in [A, 1.1A]$, for instance). Thus for $|a| \le X^3$, we have

$$\sigma_{\infty,F_a,\nu}(X) \gg \lim_{\epsilon\to 0} (2\epsilon)^{-1} \int_{A\in[1,A_0]} d^\times A\, \epsilon \gg \log A_0.$$

This proves the first part of the following observation:

Observation 2.2.16 (Cf. [Dia19, §2's analysis]). Uniformly over $A_0 \ge 1$, we have

(1) $\sigma_{\infty,F_a,\nu}(X) \gg \log A_0$ uniformly over $|a| \le X^3$, while

(2) $\sigma_{\infty,L^\perp,w} \ll \log A_0$ for all $L \in C(\mathrm{SSV})$.

Remark 2.2.17. The first part implies $\sigma_{\infty,F,w} = \int_{a\in\mathbb{R}} d\tilde{a}\, \sigma_{\infty,F_a,\nu}(X)^2 \gg (\log A_0)^2$. In fact, one can show that $\sigma_{\infty,F_a,\nu}(X) \ll \log A_0$ holds for *all* $a \in \mathbb{R}$, whence $\sigma_{\infty,F,w} \ll (\log A_0)^2$ as well.

Proof of second part. Fix L. By symmetry, there are really only two cases:

(1) L is cut out by $y_i + z_i = 0$, i.e. $x_i + x_{i+3} = 0$, over $i \in \{1, 2, 3\}$.

(2) L is cut out by $y_1 + y_2 = z_1 + z_2 = y_3 + z_3 = 0$, i.e. $x_1 + x_2 = x_4 + x_5 = x_3 + x_6 = 0$.

In the first case,

$$\sigma_{\infty,L^\perp,w} = \lim_{\epsilon\to 0} (2\epsilon)^{-3} \int_{|\tilde{y}_i+\tilde{z}_i|\le\epsilon} d\tilde{\boldsymbol{x}}\, w(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{z}}) = \int_{\mathbb{R}^3} d\tilde{\boldsymbol{y}}\, w(\tilde{\boldsymbol{y}}, -\tilde{\boldsymbol{y}}) \ll \operatorname{vol}(\operatorname{Supp}\nu) \ll \log A_0.$$

In the second case, $\sigma_{\infty,L^\perp,w} = \int_{\mathbb{R}^3} d\tilde{y}_1\, d\tilde{z}_1\, d\tilde{y}_3\, w(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{z}})|_L \ll 1$, because every point $(\tilde{y}_1, \tilde{z}_1, \tilde{y}_3) \in \operatorname{Supp}(w|_L)$ must satisfy the following three conditions: (a) $w_0(\tilde{y}_3^3) \ne 0$, whence $\tilde{y}_3 \ll 1$; (b) $|\tilde{y}_1| \asymp |\tilde{y}_3|$; and (c) $|\tilde{z}_1| \asymp |\tilde{z}_3| = |\tilde{y}_3|$. $\square$

Finally, we prove Theorem 2.1.8.

Proof of Theorem 2.1.8. The hypothesis of Theorem 2.1.8 implies, in particular, that our pair (F, w) (constructed above, given A_0) is HLH for *every fixed* choice of A_0.

Now, towards the conclusion of Theorem 2.1.8, note that we may restrict attention to $a \geq 0$, by symmetry. To proceed, we follow [Dia19, proof of Theorem 3.3], who first shows (via local calculations sketched later) that

$$K^{-1} \sum_{1 \leq a \leq K}' \frac{1}{s_{F_a}(K)} \leq \prod_{p=3} O(1) \prod_{p \equiv 2 \bmod 3} (1 + O(p^{-3})) \prod_{p \equiv 1 \bmod 3} (1 + O(p^{-3/2})) \ll 1,$$

where we restrict to $a \not\equiv 4, 5 \bmod 9$ (i.e. locally *admissible* integers).

Now fix $\delta > 0$ arbitrarily small. Suppose $9 \mid K$ and $X \geq K \geq 1$. The previous display implies that $s_{F_a}(K) \geq \delta$ for all but an $O(\delta)$ *fraction* of admissible residues $a \bmod K$. But by construction, $\sigma_{\infty, F_a, \nu}(X) \gg \log A_0$ uniformly over $|a| \leq X^3$. Thus $\rho_{a,\nu}(X) := s_{F_a}(K) \sigma_{\infty, F_a, \nu}(X) \gg \delta \log A_0$ holds for all but an $O(\delta)$ fraction of admissible integers $a \leq X^3$. So if $A_0 \geq 2$, say, then

$$\begin{aligned} \star &:= \#\{\text{admissible } a \leq X^3 : N_{F_a}(\infty) = 0\} \\ &\leq \#\{\text{admissible } a \leq X^3 : |N_{F_a,\nu}(X) - \rho_{a,\nu}(X)| \geq \rho_{a,\nu}(X)/2\} \\ &\leq O\left(\delta X^3 + \frac{\operatorname{Var}(X, M)}{\delta^2 (\log A_0)^2}\right). \end{aligned}$$

(The preceding display—a variant of Chebyshev's inequality—may be quite far from the truth. But without higher moments, we cannot say more.) By HLH and the bound $\sigma_{\infty, L^\perp, w} \ll \log A_0$, though,

$$\operatorname{Var}(X, M) \leq [o_{\nu; X \to \infty}(1) + O(\log A_0)] X^3 + O(X^3 \sigma_{\infty, F, w} M^{-20/31}) + K^4 O_\nu((X/K)^{-100}).$$

To finish, fix $A_0 \gg_\delta 1$ so that $O(\delta^{-2} (\log A_0)^{-2}) \cdot O(\log A_0) \leq \delta$. Recall that A_0 determines ν, and fix $M \gg_{\delta, A_0} 1$ so that $O(\delta^{-2} (\log A_0)^{-2}) \cdot O(\sigma_{\infty, F, w} M^{-20/31}) \leq \delta$. Then $\star \leq O(\delta X^3)$ for all sufficiently large $X \gg_{\delta, A_0, K} 1$. Since $\delta > 0$ was arbitrary, we are done. $\square$

Remark 2.2.18. If we assumed a power saving in HLH, with "$O(A_0^{O(1)} X^{3 - \Omega(1)})$" in place of $o_{\nu; X \to \infty}(X^3)$, then we would likely be able to let A_0 grow as a small power of X, and M, δ as small powers of $(\log X)^{\pm 1}$. We would then likely get "$\star \ll X^3 / (\log X)^{\Omega(1)}$" as $X \to \infty$. For proof, we would need the following ingredients (for fixed w_0, D), not all proven above:

(1) $K \ll_\epsilon X^\epsilon$ (true for $M = o(\log X)$, since in general, $\log K(M) \sim M$ as $M \to \infty$);

(2) $(\log A_0)^{-2} \cdot \sigma_{\infty, F, w} \asymp 1$ (a fact essentially remarked earlier); and

(3) replacing "$K^4 O_\nu((X/K)^{-100})$" with "$O(\operatorname{diam}(\operatorname{Supp} \nu)^{O(1)} K^4 (X/K)^{-100})$ (by analyzing the ν-dependence in our error estimate from Poisson summation, stemming from Proposition 2.2.8).

Sketch of local calculations. See [Dia19, pp. 10–11 (proof of Theorem 1.4(ii)) and pp. 32–34] for details, including the necessary Hensel lifting to moduli $p^{\geq 2}$ at each prime p. The 3-adic densities, in particular, require a bit more lifting work than the densities for $p \neq 3$ do.

At $p \equiv 1 \bmod 3$, the local calculations mostly boil down to a finite linear combination of cubic characters $\chi \bmod p$ (evaluated at certain points depending on F_0, a), since $F_0 = a$ is an affine *diagonal* cubic surface over $\mathbb{F}_p$ when $a \neq 0$. (Unlike with CM elliptic curves, Hecke characters of infinite order, e.g. signed normalized cubic Gauss sums $-\tilde{g}(\chi) = -p^{-1/2} \sum_{x \in \mathbb{F}_p} \chi(x) e_p(x)$, do not arise, except in secondary terms related to $F_0 = 0$.) Ultimately here,

$$\begin{aligned}\sum_{a \in \mathbb{F}_p} \frac{1}{s_{F_a}(p)} &= \sum_{a \in \mathbb{F}_p} \frac{p^2}{\#\{\boldsymbol{y} \in \mathbb{F}_p^3 : F_a(\boldsymbol{y}) = 0\}} \\ &= 1 + O(p^{-1/2}) + \frac{(p-1)/3}{1 + 6/p + O(p^{-3/2})} + \frac{2(p-1)/3}{1 - 3/p + O(p^{-3/2})} \\ &= 1 + O(p^{-1/2}) + (p-1)[1 - (1/3)(6/p) + (2/3)(3/p)] = p + O(p^{-1/2}).\end{aligned}$$

(The precise "$O(p^{-1/2})$" is roughly proportional to $\frac{1}{p}(a_p(E) + O(1))$, where $E := V_{\mathbb{P}^2}(F_0)_{/\mathbb{Q}}$.)

At primes $p \equiv 2 \bmod 3$, the local calculations are easier (as if F_0 were linear), due to bijectivity of the function $x \mapsto x^3$ on $\mathbb{F}_p$. $\square$

Chapter 3

Review of the delta method

3.1 The basic setup

Let $m \in \mathbb{Z}_{\geq 3}$. Let $F \in \mathbb{Z}[\boldsymbol{x}] = \mathbb{Z}[x_1, \ldots, x_m]$ be an m-variable cubic form with nonzero discriminant. Let $\mathcal{V} := V_{\mathbb{P}^{m-1}}(F)_{/\mathbb{Z}}$ and $V := \mathcal{V}_{\mathbb{Q}} = V_{\mathbb{P}^{m-1}}(F)_{/\mathbb{Q}}$ be cut out by $F = 0$. Then V, the generic fiber of $\mathcal{V}$, is a smooth projective hypersurface in $\mathbb{P}^{m-1}_{\mathbb{Q}}$. For $\boldsymbol{c} \in \mathbb{Z}^m$, we define hyperplane sections $\mathcal{V}_{\boldsymbol{c}}, V_{\boldsymbol{c}}$ via the following convenient general definition:

Definition 3.1.1. Given a ring R, an m-tuple $\boldsymbol{c}$ that "makes sense" in R^m (or more precisely, an m-tuple $\boldsymbol{c}$ that maps canonically into R^m), and a closed subscheme $W \subseteq \mathbb{P}^{m-1}_R$, define $W_{\boldsymbol{c}}$ to be the scheme-theoretic intersection $W \cap \{\boldsymbol{c} \cdot \boldsymbol{x} = 0\}$.

Remark 3.1.2. If $\boldsymbol{c} \in \mathbb{Q}^m$, then $V_{\boldsymbol{c}}$ is a hypersurface in $(\mathbb{P}^{m-1}_{\mathbb{Q}})_{\boldsymbol{c}}$ (since F is irreducible), where $(\mathbb{P}^{m-1}_{\mathbb{Q}})_{\boldsymbol{c}} \cong \mathbb{P}^{m-2}_{\mathbb{Q}}$ if $\boldsymbol{c} \neq \boldsymbol{0}$.

Definition 3.1.3. If $n \in \mathbb{Z}_{\geq 1}$, then for each m-tuple $\boldsymbol{c}$ that "makes sense" in $(\mathbb{Z}/n)^m$, let $S_{\boldsymbol{c}}(n) := \sum_{a \in (\mathbb{Z}/n)^\times} \sum_{\boldsymbol{x} \in (\mathbb{Z}/n)^m} e_n(aF(\boldsymbol{x}) + \boldsymbol{c} \cdot \boldsymbol{x})$ and $\widetilde{S}_{\boldsymbol{c}}(n) := n^{-(m+1)/2} S_{\boldsymbol{c}}(n)$. Then for each $\boldsymbol{c} \in \mathbb{Z}^m$, let $\Phi(\boldsymbol{c}, s) := \sum_{n \geq 1} \widetilde{S}_{\boldsymbol{c}}(n) n^{-s}$.

Now fix $w \in C_c^\infty(\mathbb{R}^m)$ with (F, w) smooth (in the sense of Definition 1.4.3), and suppose we are interested in the weighted zero count $N_{F,w}(X)$ (from Definition 1.4.1) for real $X > 0$. We begin by choosing "standard cutoff parameters" in the delta method, following [DFI93, HB96, HB98].

Definition 3.1.4. Set $Y := X^{(\deg F)/2} = X^{3/2}$ to be, roughly, the largest modulus used in the delta method—and given $\epsilon_0 \in (0, 10^{-10}]$, set $Z = Z_{\epsilon_0} := Y/X^{1-\epsilon_0} = X^{1/2+\epsilon_0}$. (To have correct epsilon management, we have named this epsilon ϵ_0 to be safe.)

Before proceeding, let us recall the standard intuition behind the delta method—intuition one can formalize via eq. (3.1), Proposition 3.1.6, and Lemma 3.1.7 below.

Remark 3.1.5 (Intuition). If $\boldsymbol{x} \ll X$, then $F(\boldsymbol{x}) \ll X^3$. So it is natural to try using a total of $\asymp X^3$ harmonics to "detect" the condition $F(\boldsymbol{x}) = 0$ over $\boldsymbol{x} \ll X$. Since the delta method morally uses $\asymp Y^2$ harmonics (corresponding to proper reduced fractions with denominator $\ll Y$), this suggests setting $Y^2 \asymp X^3$.

(Smaller choices of Y could in principle also be worth considering; see the quartic analysis of [MV19]. See §9.2 for some discussion.)

One then encounters certain "pseudo-exponential sums to modulus $n \ll Y$" over $\boldsymbol{x} \ll X$. These sums are typically "incomplete" (since $X = o(Y)$). By "completing" these sums using Poisson summation, we end up with "dual sums" morally of length $\lessapprox Y/X \asymp X^{1/2}$.

By [HB96, Theorem 2, (1.2)] (based on [DFI93])—or rather, [HB96, (1.2), up to easy manipulations from §3 involving the formula (3.3) and the switching of $n, \boldsymbol{c}$]—we have (uniformly over $X > 0$)

$$(1 + O_A(Y^{-A})) \cdot N_{F,w}(X) = Y^{-2} \sum_{n\geq 1} \sum_{\boldsymbol{c}\in\mathbb{Z}^m} n^{-m} S_{\boldsymbol{c}}(n) I_{\boldsymbol{c}}(n), \tag{3.1}$$

where $S_{\boldsymbol{c}}(n)$ is defined as in Definition 3.1.3, and where (in terms of a certain function $h(-,-)$ typically left in the background; see e.g. [HB98, eq. (2.3)] for the definition)

$$I_{\boldsymbol{c}}(n) := \int_{\boldsymbol{x}\in\mathbb{R}^m} d\boldsymbol{x}\, w(\boldsymbol{x}/X) h(n/Y, F(\boldsymbol{x})/Y^2) e_n(-\boldsymbol{c}\cdot\boldsymbol{x}),$$

for $\boldsymbol{c} \in \mathbb{Z}^m$. Dimensional analysis suggests the normalization

$$\widetilde{I}_{\boldsymbol{c}}(n) := X^{-m} I_{\boldsymbol{c}}(n) = \int_{\tilde{\boldsymbol{x}}\in\mathbb{R}^m} d\tilde{\boldsymbol{x}}\, w(\tilde{\boldsymbol{x}}) h(n/Y, F(\tilde{\boldsymbol{x}})) e_{n/X}(-\boldsymbol{c}\cdot\tilde{\boldsymbol{x}})$$

(using $Y^2 = X^{\deg F}$ to get $F(\boldsymbol{x})/Y^2 = F(\tilde{\boldsymbol{x}})$ for $\tilde{\boldsymbol{x}} := \boldsymbol{x}/X$). Eq. (3.1) then becomes

$$(1 + O_A(X^{-A})) \cdot N_{F,w}(X) = X^{m-3} \sum_{n\geq 1} \sum_{\boldsymbol{c}\in\mathbb{Z}^m} n^{-(m-1)/2} \widetilde{S}_{\boldsymbol{c}}(n) \widetilde{I}_{\boldsymbol{c}}(n). \tag{3.2}$$

We have referred to the switching of $n, \boldsymbol{c}$ as "easy" because the compact support of w guarantees that $I_{\boldsymbol{c}}(n)$ is supported on finitely many moduli n, and is rapidly decaying in $\|\boldsymbol{c}\|$ for each n. These *qualitative* facts—important for [HB96, §3, proof of Theorem 2] (which at one point switches $n, \boldsymbol{c}$ to get to [HB96, (1.2)]), and for us—follow from the following two standard results, which will soon begin to play an important *quantitative* role as well.

Proposition 3.1.6 (Vanishing for large n)**.** *The functions $n \mapsto I_{\boldsymbol{c}}(n)$ are supported on $n \ll_{F,w} Y$, uniformly over $\boldsymbol{c} \in \mathbb{Z}^m$.*

Proof. See e.g. [HB96, par. 1 of §7]. The vanishing of $I_{\boldsymbol{c}}(n)$ for $n \gg_{F,w} Y$ sufficiently large is a consequence of the choice $Y \asymp X^{(\deg F)/2}$ and the definition of $h(-,-)$. □

As a sanity check, note that if $X \ll_{F,w} 1$ is sufficiently small, then Proposition 3.1.6 yields $I_{\boldsymbol{c}}(n) = 0$ for all $n \geq 1$, whence the right-hand side of eq. (3.1) vanishes. This is consistent with the fact that the "true" factor of $1 + O_A(Y^{-A})$ on the left-hand side of eq. (3.1) vanishes for $Y < 1$ (cf. [HB96, left-hand side of (3.3) for $Q < 1$]).

The case $X \leq 1$, say, is similarly uninteresting: if X is bounded, then both sides of eq. (3.1) are trivially bounded as well. So from now on, we assume $X \geq 1$.

Lemma 3.1.7 (Decay for large $\boldsymbol{c}$)**.** *If $\|\boldsymbol{c}\| \geq Z$ and $n \geq 1$, then $I_{\boldsymbol{c}}(n) \ll_{\epsilon_0,A} \|\boldsymbol{c}\|^{-A}$.*

Proof. See e.g. [HB98, Lemma 3.2, (3.9)]. □

Proposition 3.1.6 and Lemma 3.1.7 easily (and comfortably) imply the *absolute* bound

$$Y^{-2}\sum_{n\geq 1}\sum_{\|\boldsymbol{c}\|\geq Z} n^{-m}|S_{\boldsymbol{c}}(n)|\cdot|I_{\boldsymbol{c}}(n)| \ll_{F,w,\epsilon_0,A} X^{-A}$$

(even if we only use the trivial bound $|S_{\boldsymbol{c}}(n)| \leq n^{1+m}$). To analyze $N_{F,w}(X)$, as $X\to\infty$, via eq. (3.2), it thus precisely remains to understand (for arbitrarily small ϵ_0) the quantity

$$X^{m-3}\sum_{n\geq 1}\sum_{\|\boldsymbol{c}\|\leq Z} n^{-(m-1)/2}\widetilde{S}_{\boldsymbol{c}}(n)\widetilde{I}_{\boldsymbol{c}}(n). \tag{3.3}$$

(Here $\widetilde{I}_{\boldsymbol{c}}(n) = \widetilde{I}_{\boldsymbol{c}}(n)\cdot \mathbf{1}_{n\ll Y}$ for a suitable factor $\mathbf{1}_{n\ll Y}$. But it is more convenient to keep the factor implicit, to allow for more flexible technique.)

3.2 Exponential sums and L-functions

The sums $S_{\boldsymbol{c}}(n)$ have some nice properties originating from the *homogeneity* of F (crucial in our analysis, and likewise in [Hoo86b, Hoo97, HB98]).

(1) The function $n\mapsto S_{\boldsymbol{c}}(n)$ is *multiplicative*, i.e. $S_{\boldsymbol{c}}(n_1n_2) = S_{\boldsymbol{c}}(n_1)S_{\boldsymbol{c}}(n_2)$ if $(n_1,n_2)=1$.

(2) The function $(F,\boldsymbol{c})\mapsto S_{\boldsymbol{c}}(n)$ is *scale-invariant*, e.g. $S_{\boldsymbol{c}}(n) = S_{\lambda\boldsymbol{c}}(n)$ if $\lambda\in(\mathbb{Z}/n)^\times$.

In particular, (2) suggests that $S_{\boldsymbol{c}}$ might only depend on *homogeneous geometric* invariants of $(F,\boldsymbol{c})$. This is indeed the case. We now recall some of the main invariants involved.

Definition 3.2.1. Let $m_* := m-3$. For a prime power q, and an m-tuple $\boldsymbol{c}$ that "makes sense" in $\mathbb{F}_q^m$, let $\rho(q),\rho_{\boldsymbol{c}}(q)$ be the respective $\mathbb{F}_q$-point counts of $\mathcal{V}_{\mathbb{F}_q}, (\mathcal{V}_{\mathbb{F}_q})_{\boldsymbol{c}}$. Normalize the "errors" $E(q) := \rho(q) - (q^{m-1}-1)/(q-1)$ and $E_{\boldsymbol{c}}(q) := \rho_{\boldsymbol{c}}(q) - (q^{m-2}-1)/(q-1)$ to get $\widetilde{E}(q) := q^{-(1+m_*)/2}E(q)$ and $\widetilde{E}_{\boldsymbol{c}}(q) := q^{-m_*/2}E_{\boldsymbol{c}}(q)$.

Remark 3.2.2. Perhaps somewhat confusingly, $E_{\mathbf{0}}(q)\neq E(q)$.

Proposition-Definition 3.2.3 (Classical)**.** *Up to scaling, there is a* unique $F^\vee\in\mathbb{Z}[\boldsymbol{c}]\setminus\{0\}$ *of degree* $3\cdot 2^{m-2}$ *such that if* $\boldsymbol{c}\in\mathbb{C}^m\setminus\{\mathbf{0}\}$*, then* $F^\vee(\boldsymbol{c})=0$ *if and only if* $(V_{\mathbb{C}})_{\boldsymbol{c}}$ *is singular.*

Now fix $F^\vee$*. Then* $F^\vee$ *is homogeneous, and irreducible over* $\mathbb{C}$*. Informally, we call* $F^\vee$ *a (geometric)* discriminant form*. Furthermore, we may choose* $F^\vee$ *so that for all* $\boldsymbol{c}\in\mathbb{Z}^m$ *and primes* $p\nmid F^\vee(\boldsymbol{c})$*, the special fiber* $(\mathcal{V}_{\boldsymbol{c}})_{\mathbb{F}_p}$ *is smooth of dimension* m_**.*

Proof. In general, see [Wan22, Appendix A] (for a discriminant-based perspective) or [Wan21d, Remark A.3] (for a perspective based on dual varieties). When F is diagonal, see [Wan21c, Proposition-Definition 1.8] for an explicit treatment; for example, if $F = x_1^3+\cdots+x_m^3$, then we may take $F^\vee = 3\prod(c_1^{3/2}\pm c_2^{3/2}\pm\cdots\pm c_m^{3/2})$. □

For $\boldsymbol{c}\in\mathbb{Z}^m$, recall that $S_{\boldsymbol{c}}$ is multiplicative. So $\widetilde{S}_{\boldsymbol{c}}$ is too. Now we recall some standard formulas at prime powers.

Proposition 3.2.4. *Say $p \nmid \boldsymbol{c}$. Then $S_{\boldsymbol{c}}(p) = p^2 E_{\boldsymbol{c}}(p) - pE(p)$.*

Proof sketch. Although [Hoo86b, p. 69, (47)] assumes $p \nmid F^\vee(\boldsymbol{c})$ and $F = x_1^3 + \cdots + x_6^3$, the underlying arguments work more generally; cf. [Hoo14, Lemma 7]. Because F is homogeneous, we have $S_{\boldsymbol{c}}(p) = S_{\lambda\boldsymbol{c}}(p)$ for all $\lambda \in \mathbb{F}_p^\times$. So

$$\begin{aligned}(p-1)S_{\boldsymbol{c}}(p) &= \sum_{\lambda, a \in \mathbb{F}_p^\times} \sum_{\boldsymbol{x} \in \mathbb{F}_p^m} e_p(aF(\boldsymbol{x}) + \lambda \boldsymbol{c} \cdot \boldsymbol{x}) \\ &= \sum_{\boldsymbol{x} \in \mathbb{F}_p^m} (p \cdot \mathbf{1}_{p \mid F(\boldsymbol{x})} - 1)(p \cdot \mathbf{1}_{p \mid \boldsymbol{c} \cdot \boldsymbol{x}} - 1) = p^2(p-1)E_{\boldsymbol{c}}(p) - p(p-1)E(p)\end{aligned}$$

if $p \nmid \boldsymbol{c}$. (The factor $p^2(p-1)$ comes from summing over λ, a and taking an affine cone over $(\mathcal{V}_{\mathbb{F}_p})_{\boldsymbol{c}}$.) □

Remark 3.2.5. If $p \mid \boldsymbol{c}$, then $S_{\boldsymbol{c}}(p) = S_{\boldsymbol{0}}(p) = p^2 E(p) - pE(p)$ instead.

In particular, $\widetilde{S}_{\boldsymbol{c}}(p) = \widetilde{E}_{\boldsymbol{c}}(p) - p^{-1/2}\widetilde{E}(p)$ at primes $p \nmid F^\vee(\boldsymbol{c})$. Here $\widetilde{E}(p) \ll 1$, by the Weil conjectures (after "absorbing" bad primes for F into the implied constant).

Proposition 3.2.6. *Say $p \nmid F^\vee(\boldsymbol{c})$. Then $S_{\boldsymbol{c}}(p^l) = 0$ for all integers $l \geq 2$.*

Proof sketch. Although [Hoo86b, Lemma 7] assumes $F = x_1^3 + \cdots + x_6^3$, the underlying proof immediately generalizes; see [Hoo14, Lemma 10]. This time, scalar symmetry in $\boldsymbol{c}$ (using homogeneity of F) gives

$$\phi(p^l) S_{\boldsymbol{c}}(p^l) = \sum_{\boldsymbol{x} \in (\mathbb{Z}/p^l)^m} [-p^{l-1} \cdot \mathbf{1}_{p^{l-1} \mid \boldsymbol{c} \cdot \boldsymbol{x}} + p^l \cdot \mathbf{1}_{p^l \mid \boldsymbol{c} \cdot \boldsymbol{x}}][-p^{l-1} \cdot \mathbf{1}_{p^{l-1} \mid F(\boldsymbol{x})} + p^l \cdot \mathbf{1}_{p^l \mid F(\boldsymbol{x})}].$$

So $S_{\boldsymbol{c}}(p^l) = 0$ is equivalent to certain statements about point counts. One can prove these statements by Hensel lifting; the lifting calculus follows dimension predictions, because $p \nmid F^\vee(\boldsymbol{c})$ implies (by Proposition-Definition 3.2.3) that the $\mathbb{F}_p$-variety $(\mathcal{V}_{\mathbb{F}_p})_{\boldsymbol{c}}$ is smooth of codimension 2. □

Consequently, we know that

(1) $\widetilde{S}_{\boldsymbol{c}}(p) = \widetilde{E}_{\boldsymbol{c}}(p) - p^{-1/2}\widetilde{E}(p)$ if $p \nmid \boldsymbol{c}$ (e.g. if $p \nmid F^\vee(\boldsymbol{c})$),

(2) $\widetilde{S}_{\boldsymbol{c}}(p) = \widetilde{S}_{\boldsymbol{0}}(p) = (p^{1/2} - p^{-1/2})\widetilde{E}(p)$ if $p \mid \boldsymbol{c}$, and

(3) $\widetilde{S}_{\boldsymbol{c}}(p^l) = 0$ for $l \geq 2$ if $p \nmid F^\vee(\boldsymbol{c})$.

But if $\boldsymbol{c} \in \mathbb{Z}^m$ and $p \nmid F^\vee(\boldsymbol{c})$, then $(\mathcal{V}_{\boldsymbol{c}})_{\mathbb{F}_p}$ is a smooth complete intersection in $\mathbb{P}^{m-1}_{\mathbb{F}_p}$ of dimension m_* and multi-degree $(3, 1)$. By the theory of ℓ-adic cohomology (including the Grothendieck–Lefschetz fixed-point theorem and the resolution of the Weil conjectures), we can thus make the following definition:

Definition 3.2.7. Fix $\boldsymbol{c} \in \mathbb{Z}^m$ with $F^\vee(\boldsymbol{c}) \neq 0$, i.e. with $V_{\boldsymbol{c}}$ smooth of dimension m_*. Then for each prime $p \nmid F^\vee(\boldsymbol{c})$, define the *analytically normalized local factor*

$$L_p(s, V_{\boldsymbol{c}}) := \exp\left((-1)^{m_*} \sum_{r \geq 1} \widetilde{E}_{\boldsymbol{c}}(p^r) \frac{(p^{-s})^r}{r}\right)$$
$$= \prod_{j=1}^{\dim_{m_*}} (1 - \tilde{\alpha}_{\boldsymbol{c},j}(p) p^{-s})^{-1} =: \sum_{l \geq 0} \tilde{\lambda}_{\boldsymbol{c}}(p^l) p^{-ls}$$

of degree $\dim_{m_*} := \operatorname{rank}(H^{m_*}_{\mathrm{sing}}(V_{\boldsymbol{c}}(\mathbb{C}), \mathbb{Z}) / H^{m_*}_{\mathrm{sing}}(\mathbb{P}^{m-1}(\mathbb{C}), \mathbb{Z})) = A_{3, m_*+2}$, where

$$A_{d,s} := \#\{\boldsymbol{a} \in [d-1]^s : a_1 + \cdots + a_s \equiv 0 \bmod d\} = \frac{(d-1)^s + (-1)^s(d-1)}{d}$$

for integers $d, s \geq 1$ (following [Wei49, p. 506]). Here the $\tilde{\alpha}_{\boldsymbol{c},j}(p)$ denote certain "normalized" Frobenius eigenvalues (known to satisfy $|\tilde{\alpha}_{\boldsymbol{c},j}(p)| = 1$).

Remark 3.2.8. Each $V_{\boldsymbol{c}}$ above is a subvariety of $\mathbb{P}^{m-1}_{\mathbb{Q}}$, so $H^{m_*}_{\mathrm{sing}}(\cdots)/H^{m_*}_{\mathrm{sing}}(\cdots)$ is well-defined. Also, each $L_p(s, V_{\boldsymbol{c}})$ above is well-defined: for any $\boldsymbol{c}, \boldsymbol{c}' \in \mathbb{Z}^m$ with $V_{\boldsymbol{c}} = V_{\boldsymbol{c}'}$, one can show (e.g. by smooth proper base change) that $E_{\boldsymbol{c}}(q) = E_{\boldsymbol{c}'}(q)$ holds for all prime powers q coprime to $F^\vee(\boldsymbol{c})F^\vee(\boldsymbol{c}')$. (To avoid discussing the two points above, we could write $-\mathbf{1}_{2 \mid m_*} + \operatorname{rank} H^{m_*}_{\mathrm{sing}}(V_{\boldsymbol{c}}(\mathbb{C}), \mathbb{Z})$ in place of $\operatorname{rank}(\cdots/\cdots)$, and $L_p(s, \boldsymbol{c})$ in place of $L_p(s, V_{\boldsymbol{c}})$. But the current notation is more transparent and suggestive.)

In particular, if $p \nmid F^\vee(\boldsymbol{c})$, then $(-1)^{m_*} \widetilde{E}_{\boldsymbol{c}}(p) = \sum_j \tilde{\alpha}_{\boldsymbol{c},j}(p) = \tilde{\lambda}_{\boldsymbol{c}}(p)$, so $\widetilde{E}_{\boldsymbol{c}}(p) \ll_m 1$. (Similarly, we have $\widetilde{E}(p) \ll_F 1$ uniformly over *all* primes p.) So for each *fixed* $\boldsymbol{c} \in \mathbb{Z}^m$ with $F^\vee(\boldsymbol{c}) \neq 0$, we roughly expect an approximation of the form

$$\Phi(\boldsymbol{c}, s) \approx \prod_{p \nmid F^\vee(\boldsymbol{c})} L_p(s, V_{\boldsymbol{c}})^{(-1)^{m_*}} \qquad \text{("to leading order").}$$

Indeed, the works [Hoo86b, Hoo97, HB98, Wan21c] are based on the intuitive notion of a "first-order approximation" of a Dirichlet series. One could imagine many different precise definitions—perhaps useful for different purposes. We now make the following convenient (but not necessarily comprehensive or all-purpose) definition:

Definition 3.2.9. Fix a family of Dirichlet series $\Psi_1(\boldsymbol{c}, s)$ indexed by $\{\boldsymbol{c} \in \mathbb{Z}^m : F^\vee(\boldsymbol{c}) \neq 0\}$. For each $\boldsymbol{c}$, let $b_{\boldsymbol{c}}(n), a_{\boldsymbol{c}}(n), a'_{\boldsymbol{c}}(n)$ be the nth coefficients of the Dirichlet series $\Psi_1, \Psi_1^{-1}, \Phi(\boldsymbol{c}, s)/\Psi_1$, respectively. Suppose that

(1) $b_{\boldsymbol{c}}$ is multiplicative, i.e. Ψ_1 has a (formal) Euler product;

(2) $b_{\boldsymbol{c}}(n), a'_{\boldsymbol{c}}(n) \ll_\epsilon n^\epsilon \sum_{d \mid n} |\widetilde{S}_{\boldsymbol{c}}(d)|$ holds uniformly over $\boldsymbol{c}, n$; and

(3) $a'_{\boldsymbol{c}}(p) \cdot \mathbf{1}_{p \nmid F^\vee(\boldsymbol{c})} \ll p^{-1/2}$ holds uniformly over $\boldsymbol{c}, p$ with p prime.

Then we call Ψ_1 a *(one-sided, first-order, Euler-product) approximation of* Φ.

Example 3.2.10. Suppose $m \in \{4,6\}$ and for each $\boldsymbol{c}$, we let $\Psi_1(\boldsymbol{c}, s) := L(s, V_{\boldsymbol{c}})^{-1}$, with the standard Hasse–Weil L-function $L(s, V_{\boldsymbol{c}})$, defined as in [Hoo86b] for $F = x_1^3 + \cdots + x_6^3$ (though the definition readily generalizes; see e.g. [HB98]); for primes $p \nmid F^\vee(\boldsymbol{c})$ the local factor agrees with the $L_p(s, V_{\boldsymbol{c}})$ defined earlier. Then Ψ_1 is an approximation of Φ, in the sense of Definition 3.2.9.

Proof sketch. Say $p \nmid F^\vee(\boldsymbol{c})$. Then (if $\mu_{\boldsymbol{c}}(-)$ denote the coefficients of $1/L(s, V_{\boldsymbol{c}})$)

$$a'_{\boldsymbol{c}}(p) := \widetilde{S}_{\boldsymbol{c}}(p) + \mu_{\boldsymbol{c}}(p) = \widetilde{S}_{\boldsymbol{c}}(p) - \widetilde{E}_{\boldsymbol{c}}(p) = -p^{-1/2}\widetilde{E}(p) \ll p^{-1/2}.$$

Also, $a'_{\boldsymbol{c}}(p^k) = \mu_{\boldsymbol{c}}(p^k) + \widetilde{S}_{\boldsymbol{c}}(p)\mu_{\boldsymbol{c}}(p^{k-1}) \ll_\epsilon p^{k\epsilon}$ for all $k \geq 2$, since $\widetilde{S}_{\boldsymbol{c}}(p^l) = 0$ for $l \geq 2$. More care is needed to handle primes $p \mid F^\vee(\boldsymbol{c})$; we need the general bound $b_{\boldsymbol{c}}(n), a_{\boldsymbol{c}}(n) \ll_\epsilon n^\epsilon$, which is luckily known for $m \in \{4,6\}$. $\square$

Remark 3.2.11. The aforementioned works [Hoo86b, Hoo97, HB98] are based on a certain "Hypothesis HW" (mentioned in Example 1.2.1) for the Hasse–Weil L-functions $L(s, V_{\boldsymbol{c}})$. Fix $\boldsymbol{c}$, let $1/L(s, V_{\boldsymbol{c}}) =: \sum_{n\geq 1} \mu_{\boldsymbol{c}}(n)n^{-s}$, and let $q(V_{\boldsymbol{c}})$ denote a certain "conductor" associated to $V_{\boldsymbol{c}}$. In lieu of precisely recalling Hypothesis HW—which amounts to *certain Selberg-type axioms* (e.g. suitable analytic continuation), *plus GRH*—we simply record the following conjectures, each known to imply the next:

(1) Certain Langlands-type conjectures, plus GRH, for $L(s, V_{\boldsymbol{c}})$.

(2) Hypothesis HW for $L(s, V_{\boldsymbol{c}})$.

(3) A certain standard elementary "uniform square-root cancellation" bound—namely $\sum_{n\leq N} \mu_{\boldsymbol{c}}(n) \ll_{m,3,\epsilon} q(V_{\boldsymbol{c}})^\epsilon N^{1/2+\epsilon}$ (with an implied constant depending only on $m, 3, \epsilon$).

Here, chief among the "Langlands-type conjectures" in (1) is *automorphy*, i.e. (a general form of) *Langlands reciprocity*—a statement generalizing the *modularity of elliptic curves*.

Remark 3.2.12. Later, in Chapter 8, we will work *directly* with (1), rather than with "avatars" like (2). This is natural, in view of "analytic-to-automorphic" converse theorems (in other natural families of Dirichlet series)—and because (1), modulo GRH, not only lies in a more conceptual framework, but also offers the only known approach to proving (2), modulo GRH.

3.3 Main unconditional general pointwise bounds

For the rest of Chapter 3, assume F is diagonal. For technical reasons, we will analyze the $\boldsymbol{c}$'s in groups depending on how many coordinates are zero. It would be interesting to find a similar notion for non-diagonal forms F.

Definition 3.3.1. Given $\mathcal{I} \subseteq [m]$ of size $r \leq m$, we call the set $\mathcal{R} \subseteq [-Z, Z]^m$ of $\boldsymbol{c} \in \mathbb{Z}^m$ with $c_j \in [-Z, Z] \setminus \{0\}$ for $j \in \mathcal{I}$ and $c_j = 0$ for $j \notin \mathcal{I}$ a (uniform) r-dimensional *deleted box*.

Let $\mathcal{R} \subseteq [-Z, Z]^m$ be a deleted box with $|\mathcal{I}| = r \in [0, m]$. Then we can bound the sums $\widetilde{S}_{\boldsymbol{c}}(n) := n^{-(m+1)/2}S_{\boldsymbol{c}}(n)$ using the somewhat crude but *general* pointwise bound given by the next result (available since F is diagonal).

Definition 3.3.2. For an integer $n \ge 1$, let $\operatorname{sq}(n) := \prod_{p^2 \mid n} p^{v_p(n)}$ denote the *square-full part* of n, and $\operatorname{cub}(n) := \prod_{p^3 \mid n} p^{v_p(n)}$ the *cube-full part* of n.

Proposition 3.3.3 ([Hoo86b, HB98]). *For all $\boldsymbol{c} \in \mathbb{Z}^m$ and integers $n \ge 1$, we have*

$$n^{-1/2}|\widetilde{S}_{\boldsymbol{c}}(n)| \ll_F O(1)^{\omega(n)} \prod_{j \in [m]} \gcd(\operatorname{cub}(n)^{1/6}, \gcd(\operatorname{cub}(n), \operatorname{sq}(c_j))^{1/4}).$$

Here we interpret $\gcd(-,-)$ formally in terms of exponents that are allowed to be rational.

Proof. In general $\widetilde{S}_{\boldsymbol{c}}(n) = n^{-(m+1)/2} S_{\boldsymbol{c}}(n)$ by definition, and for diagonal F we have

$$|S_{\boldsymbol{c}}(p^l)| \ll (p^\infty, O_F(1))^{O(1)} \cdot p^{l(1+m/2)} \prod_{j \in [m]} \gcd(\operatorname{cub}(p^l)^{1/6}, \gcd(\operatorname{cub}(p^l), \operatorname{sq}(c_j))^{1/4})$$

by [HB98, p. 682, (5.1)–(5.2)] for $l \ge 2$ and [HB83, Lemma 11] for $l = 1$. $\square$

Remark 3.3.4. We have stated Proposition 3.3.3 uniformly over $\boldsymbol{c} \in \mathbb{Z}^m$. But given $\mathcal{R}$, Proposition 3.3.3 *implies* that for all $\boldsymbol{c} \in \mathcal{R}$ and integers $n \ge 1$, we have

$$n^{-1/2}|\widetilde{S}_{\boldsymbol{c}}(n)| \ll_F O(1)^{\omega(n)} \operatorname{cub}(n)^{(m-r)/6} \prod_{j \in \mathcal{I}} \gcd(\operatorname{cub}(n), \operatorname{sq}(c_j))^{1/4}.$$

It is this simpler statement that we typically use.

Now we turn to the *integrals* $\widetilde{I}_{\boldsymbol{c}}(n) := X^{-m} I_{\boldsymbol{c}}(n)$, assuming $r \ge 1$. (We cover $\boldsymbol{c} = \boldsymbol{0}$ in §3.4.) The statement [HB98, p. 678, Lemma 3.2] covers the essential cases $k = 0, 1$ in the next result (which, like Proposition 3.3.3, is available since F is diagonal). We state a generalization to all $k \ge 0$, just in case it comes in handy for some (future) smoothing purposes.

Lemma 3.3.5 (Main n-aspect bounds). *Uniformly over $\boldsymbol{c} \in \mathcal{R}$ and $n \in [1/2, \infty)$, we have*

$$n^k |\partial_n^k \widetilde{I}_{\boldsymbol{c}}(n)| \ll_{k,\epsilon} X^\epsilon \left(\frac{X\|\boldsymbol{c}\|}{n}\right)^{1-(m+r)/4} \prod_{i \in \mathcal{I}} \left(\frac{\|\boldsymbol{c}\|}{|c_i|}\right)^{1/2} \quad \text{for } k = 0, 1, 2, \ldots.$$

Furthermore, if $B \in C_c^\infty(\mathbb{R}_{>0})$ is supported on $[1/2, 1]$, then $n^{k+1}|\partial_n^k[y^{-1}B(n/y)\widetilde{I}_{\boldsymbol{c}}(n)]|$ satisfies the same bound for all $n \in (0, \infty)$, uniformly as $y \ge 1$ varies.

Proof. By [Wan21g, Remark 4.1 and Lemma 4.9], we know that (for all $\boldsymbol{c} \in \mathbb{R}^m \setminus \{\boldsymbol{0}\}$)

$$n^k \cdot \partial_n^k I_{\boldsymbol{c}}(n) \ll_{k,\epsilon} \left(\frac{X\|\boldsymbol{c}\|}{n}\right) X^{m+\epsilon} \prod_{i=1}^m \min[(n/X|c_i|)^{1/2}, (n/X\|\boldsymbol{c}\|)^{1/4}],$$

and that $n^{k+1} \cdot \partial_n^k[y^{-1}B(n/y)I_{\boldsymbol{c}}(n)]$ satisfies the same bound. Now "replace" $\min[-,-]$ with $(n/X|c_i|)^{1/2}$ for each $i \in \mathcal{I}$, and with $(n/X\|\boldsymbol{c}\|)^{1/4}$ for each $i \in [m] \setminus \mathcal{I}$; this bounds the right-hand side by $X^{m+\epsilon}(X\|\boldsymbol{c}\|/n)^{1-(m-r)/4} \prod_{i \in \mathcal{I}} (n/X|c_i|)^{1/2}$, which (after dividing by X^m) simplifies to what we want (since $\widetilde{I}_{\boldsymbol{c}}(n) := X^{-m} I_{\boldsymbol{c}}(n)$). $\square$

Remark 3.3.6. We should emphasize that the above bounds on $I_{\boldsymbol{c}}(n)$ are likely only valid (as written) for the usual (and present) setting of the parameter $Y \asymp X^{(\deg F)/2} = X^{3/2}$.

Remark 3.3.7. If (F, w) is clean, then $n^k |\partial_n^k \widetilde{I}_{\boldsymbol{c}}(n)| \ll_{k,\epsilon} X^\epsilon \min(1, (X\|\boldsymbol{c}\|/n)^{1-m/2}) \ll X^\epsilon (X\|\boldsymbol{c}\|/n)^{1-m/2}$; cf. [Hoo14, p. 252].

Remark 3.3.8. It would be interesting to know the optimal asymptotics, e.g. whether or not the bound remains true with a smaller power of $X\|\boldsymbol{c}\|/n$. At least for generic $\boldsymbol{c}$, one might expect to reduce the exponent $1-m/2$ using a deeper stationary phase analysis (implemented to some extent in Chapter 7).

3.4 Contribution from the central terms

Here we address $\boldsymbol{c} = \boldsymbol{0}$ in (3.3), using the theory of $I_{\boldsymbol{0}}(n)$ developed in [HB96]. This section is standard, and does not require F to be diagonal (see [Wan21d, Appendix B]), but we keep the diagonality assumption for convenience. We begin with a slight variant of [Vau97, Lemma 4.9].

Lemma 3.4.1. *If $N > 0$, then $\sum_{n \asymp N} n^{-m} |S_{\boldsymbol{0}}(n)| \ll_\epsilon N^{(4-m)/3+\epsilon}$.*

Proof. F is diagonal, so by Proposition 3.3.3 (or [Vau97, Lemma 4.7]) and Hölder, $n^{-m} S_{\boldsymbol{0}}(n) \ll O(1)^{\omega(n)} n^{1-m/2} \operatorname{cub}(n)^{m/6}$. If $n_3 := \operatorname{cub}(n)$ (so that n_3 is cube-full), then it follows that

$$\sum_{n \asymp N} n^{-m} |S_{\boldsymbol{0}}(n)| \ll_\epsilon \sum_{n_3 \ll N}{}' (N/n_3) \cdot N^{1-m/2+\epsilon} n_3^{m/6} \asymp \sum_{N_3 \ll N}{}' N_3^{1/3} \cdot N^{2-m/2+\epsilon} N_3^{m/6-1},$$

where N_3 ranges over $\{1, 2, 4, 8, \dots\}$. Since $1/3 + (m/6 - 1) = (m-4)/6 \geq 0$, the right-hand side is $\ll_\epsilon N^{(m-4)/6+\epsilon} \cdot N^{2-m/2+\epsilon} = N^{(4-m)/3+2\epsilon}$, as desired. $\square$

In particular, by Lemma 3.4.1 (or [Vau97, proof of Lemma 4.9]), the singular series

$$\mathfrak{S} := \sum_{n \geq 1} n^{-m} S_{\boldsymbol{0}}(n)$$

converges absolutely for $m \geq 5$. On the other hand, the real density

$$\sigma_{\infty,w} := \lim_{\epsilon \to 0} (2\epsilon)^{-1} \int_{|F(\boldsymbol{x})| \leq \epsilon} d\boldsymbol{x}\, w(\boldsymbol{x})$$

is $O_{F,w}(1)$ (as one can show using [HB96, Theorem 3], for instance).

Yet by [HB96, Lemma 13], $\widetilde{I}_{\boldsymbol{0}}(n) = \sigma_{\infty,w} + O_A((n/Y)^A)$ for all sufficiently small $n \ll Y$—hence for all $n \geq 1$ (since $\widetilde{I}_{\boldsymbol{0}}(n) \ll 1$ always, by [HB96, Lemma 16]). On the other hand, $\widetilde{I}_{\boldsymbol{0}}(n) = 0$ for all sufficiently large $n \gg Y$ (by Proposition 3.1.6). So by Lemma 3.4.1, the sum $\sum_{n \geq 1} n^{-m} S_{\boldsymbol{0}}(n) \widetilde{I}_{\boldsymbol{0}}(n)$ simplifies to

$$\left(\mathfrak{S} - \sum_{N \gg Y}{}' O_\epsilon(N^{(4-m)/3+\epsilon})\right) \sigma_{\infty,w} + \sum_{N \ll Y}{}' O_{A,\epsilon}\left(N^{(4-m)/3+\epsilon}(N/Y)^A\right),$$

where N ranges over $\{1, 2, 4, 8, \dots\}$. For $A = (m-3)/3$, both N-sums are geometric series peaking at $N \asymp Y$ (since we have assumed $X \geq 1$). Since $n^{-m} S_{\mathbf{0}}(n) = n^{-(m-1)/2} \widetilde{S}_{\mathbf{0}}(n)$, it follows that if $m \geq 5$, then

$$X^{m-3} \sum_{n \geq 1} n^{-(m-1)/2} \widetilde{S}_{\mathbf{0}}(n) \widetilde{I}_{\mathbf{0}}(n) = X^{m-3} \cdot [\sigma_{\infty,w} \mathfrak{S} + O_\epsilon(X^{(4-m)/2+\epsilon})].$$

On the other hand, for all $m \geq 4$,

$$X^{m-3} \sum_{n \geq 1} n^{-(m-1)/2} \widetilde{S}_{\mathbf{0}}(n) \widetilde{I}_{\mathbf{0}}(n) \ll X^{m-3} \sum_{n \geq 1} n^{-m} |S_{\mathbf{0}}(n)| \cdot \mathbf{1}_{n \ll Y} \ll_\epsilon X^{m-3+\epsilon}$$

by Hölder and [Vau97, Lemma 4.9].

Remark 3.4.2. One may well do better by analyzing the Dirichlet series $\sum_{n \geq 1} n^{-s} S_{\mathbf{0}}(n)$, using the derivative bound $\partial_n^k I_{\mathbf{0}}(n) \ll_k n^{-k} X^m$ (stated in [HB96, Lemma 16] for $k = 0, 1$, but valid for $k \geq 0$ with little change in proof). It would be interesting to extend the above analysis to the case $m = 4$ (with powers of $\log X$ expected to occur for certain F's); cf. [Bro09, §8.3.3's heuristic analysis of the Fermat cubic surface].

Chapter 4

Using hypotheses on average

4.1 Using large-sieve hypotheses

Let $m := 6$ and $F := x_1^3 + \cdots + x_6^3$. Recall, from Example 1.2.1, the conditional near-optimal bound $M_2(X) \ll_\epsilon X^{3+\epsilon}$, and the underlying Hypothesis HW (practically amounting to "automorphy and GRH") for the standard Hasse–Weil L-functions $L(s, V_{\boldsymbol{c}})$ associated to *smooth* hyperplane sections $V_{\boldsymbol{c}}$ (see Example 3.2.10).

Since automorphy (in full generality)—and especially GRH—might not be proven for some time, it is therefore natural to ask whether "Hypothesis HW" can be weakened. To quote [Hoo86b, pp. 51–52]: "The removal of the dependence of our work on the Riemann hypothesis is an obvious desideratum. Some weakening of the hypothesis is certainly possible either by substituting some form of zero density requirement or by insisting merely that the zeros of the Hasse-Weil L-functions be to the left of some vertical line lying to the right of the critical line $\sigma = 2$. Yet is has not seemed worthwhile to explore such developments here because the principles of the method would be obscured and because we cannot predict the precise form of the first serviceable alternative to the Riemann hypothesis that might subsequently be established." (Regarding the question of automorphy, see Appendix A for more details in the specific setting of $L(s, V_{\boldsymbol{c}})$'s above.)

In fact, by [Wan21g], it suffices to assume automorphy, along with a "density hypothesis for zeros of height $\lessapprox 1$"—an upper-bound statement on certain exceptional counts $N(\sigma, \mathcal{F}, T) := \#\{\pi \in \mathcal{F} : L(s, \pi) \text{ has a zero in } [\sigma, 1] \times [-T, T]\}$—of the form (cf. [Wan21g, §6's density hypothesis, with $l(\sigma) := 2(1-\sigma) + \epsilon$])

$$N(\sigma, \mathcal{F}, T) \ll_\epsilon T^{O(1)} |\mathcal{F}|^{2(1-\sigma)+\epsilon}, \quad \text{uniformly over } T, |\mathcal{F}| \geq 1 \text{ and } \sigma \in [1/2, 1] \tag{4.1}$$

(optimal, up to ϵ, among positive linear-in-σ exponents in the $|\mathcal{F}|$-aspect; but "arbitrarily polynomially poor" in the T-aspect). In general, (4.1) is morally provable under an "optimal" large sieve inequality over $\mathcal{F}$, i.e. "approximate ℓ^2 orthogonality" of the matrix $\mathbb{R}^{|I|} \to \mathbb{R}^{|\mathcal{F}|}$ defined by $n \mapsto (\lambda_\pi(n))_{\pi \in \mathcal{F}}$—but currently a "Lindelöf on average" hypothesis over $\mathcal{F}$ also plays a role in rigorous proof [Wan21b]. For instance, (4.1) is "close" to known for the Bombieri–Vinogradov family $\{\chi \bmod q : q \leq Q\}$.

Yet eventually, I realized that a large sieve *by itself*—if true—would suffice for the original goal of "recovering" the main results of [Hoo86b, Hoo97, HB98]. See [Wan21c] for details. Let us roughly compare the argument to that of Hooley and Heath-Brown.

Outline of conditional $M_2(X) \ll_\epsilon X^{3+\epsilon}$ proofs. Let $Y := X^{3/2}$ and $Z := X^{1/2+\epsilon_0}$ (as in Definition 3.1.4). For a suitable weight w, it suffices to bound the expression (3.3). The contribution to (3.3) from the locus $F^\vee(\boldsymbol{c}) = 0$ can be unconditionally bounded by $O_{\epsilon_0}(X^{3+O(\epsilon_0)})$ (and in fact, it captures the main terms of HLH; see Chapter 6), so we focus on the locus $F^\vee(\boldsymbol{c}) \neq 0$. Given $\boldsymbol{c} \in \mathbb{Z}^m$ with $F^\vee(\boldsymbol{c}) \neq 0$, recall the Dirichlet series $\Phi(\boldsymbol{c}, s)$ from Definition 3.1.3, and let $\Psi_1 := 1/L(s, V_{\boldsymbol{c}})$ and $\Psi_2 := \Phi L$, so that $\Phi = \Psi_1\Psi_2$. Let $a'_{\boldsymbol{c}}(n) := [n^s]\Psi_2$ denote the nth coefficient of the "error" Ψ_2. Then the following hold (and can be proven using Example 3.2.10 and Proposition 3.3.3):

[B1'] For positive $N \leq Z^3$, the first moment $\sum'_{\boldsymbol{c} \ll Z} \sum_{n \asymp N} |a'_{\boldsymbol{c}}(n)|$ is $O_\epsilon(Z^{m+\epsilon} N^{1/2})$.

(To first order, this follows from [Hoo86b, pp. 78–79, analysis of $Q(\boldsymbol{m}; k_2)$], or alternatively from [HB98, Lemma 5.2].)

[B2'] For positive $N \leq Z^3$, the second moment $\sum'_{\boldsymbol{c} \ll Z} (\sum_{n \asymp N} |a'_{\boldsymbol{c}}(n)|)^2$ is $O_\epsilon(Z^{m+\epsilon} N)$.

(This follows directly from [Wan21c, §2.4, Proposition 2.16]. This is a surprisingly delicate point; see Remark 4.1.3 below.)

The arguments in [Hoo86b, Hoo97] and [HB98] can then loosely be interpreted as

(H1) using partial summation over $n \asymp N \ll Y$ to "factor out" $\widetilde{K} := n^{-(m-1)/2} \widetilde{I}_{\boldsymbol{c}}(n)$ from $\sum_n \widetilde{S}\widetilde{K}$ (though in fact, partial summation is more like a certain "weighted decoupling" over n), and then bounding the $\widetilde{K}$-contribution in $\ell^\infty(\{n \asymp N\})$;

(H2) expanding $\widetilde{S} = \widetilde{S}_{\boldsymbol{c}} = \mu_{\boldsymbol{c}} * a'_{\boldsymbol{c}}$ using $\Phi = \Psi_1\Psi_2$;

(H3) using GRH to bound the Ψ_1-contribution in $\ell^\infty(\{\boldsymbol{c} \ll Z\})$; and

(H4) using [B1'] *afterwards*, to bound the Ψ_2-contribution in $\ell^1(\{\boldsymbol{c} \ll Z\})$.

Routine *dyadic bookkeeping* then bounds (3.3) by $O_{\epsilon_0}(X^{3+O(\epsilon_0)})$, under Hypothesis HW. But in fact, by [Wan21c], one can replace Hypothesis HW with a clean "elementary GRH on average in ℓ^2" statement, which can in turn be reduced to an "optimal" large sieve inequality for the L-function family $\boldsymbol{c} \mapsto L(s, V_{\boldsymbol{c}})$. □

Remark 4.1.1. In the aforementioned *dyadic bookkeeping*, each dyadic range of *moduli* n contributes *roughly equally* to the final bound $O_{\epsilon_0}(X^{3+O(\epsilon_0)})$. Thus in Chapter 8 we will need new integral bounds that decay, as $n \to 0$, fairly uniformly over $\boldsymbol{c}$. We will develop these (and other estimates) in Chapter 7.

Remark 4.1.2. Roughly speaking, [Wan21c] uses partial summation specifically to deduce, for a suitable probability measure $\nu = \nu_N$ supported on $[N, 2N]$, that

$$\left| \sum_{n \in [N,2N)} \widetilde{S}(n)\widetilde{K}(n) \right| \ll \left(\sup_{n \in [N,2N]} \left(|\widetilde{K}(n)|, N \cdot |\partial_n \widetilde{K}(n)| \right) \right) \int_{x \in [N,2N]} d\nu(x)\, |B([N,x))|,$$

where $B(J) := \sum_{n \in J} \widetilde{S}(n)$ for intervals J. The actual argument is subtler, but it builds on this idea. The point is that the large sieve only accepts uniform vectors; yet our "initially given" vectors (in the delta method) are only approximately uniform over $\boldsymbol{c}$, due to variation in the archimedean component $\widetilde{K}$, *and* in the error factor a'.

Remark 4.1.3. What is [B2'] really saying? Say n is *good* if $n \perp F^\vee(\boldsymbol{c})$, and *purely bad* if $n \mid F^\vee(\boldsymbol{c})^\infty$; then see Table 4.1 below for a breakdown of [B2']. (Note that $\mathbb{E}_{c\ll Z}[\operatorname{sq}(c)^{1/2}] \ll_\epsilon Z^\epsilon$; this is surprisingly delicate [Wan21c, Remark 2.17].)

Table 4.1: Types of contributions to $\sum'_{\boldsymbol{c}\ll Z}(\sum_{n\asymp N}|a'_{\boldsymbol{c}}(n)|)^2$ in [B2']

If n is (frequency)	then $a'_{\boldsymbol{c}}(n)$ is $O_\epsilon(n^\epsilon)\cdot\boxed{?}$	contributing $\ll$ (up to ϵ)
good, sq.-free (common)	$n^{-1/2}$	$Z^m(N\cdot N^{-1/2})^2 = Z^mN$
good, sq.-full (rare)	1	$Z^m(N^{1/2}\cdot 1)^2 = Z^mN$
purely bad (very rare)	$n^{1/2}\prod_{j\in[m]}\operatorname{sq}(c_j)^{1/4}$	$Z^m(1\cdot N^{1/2})^2 = Z^mN$
arbitrary (factor mixture)	mixture	mixture

Looking ahead to Chapters 7–8, the first two rows ("sources of ϵ") in Table 4.1 will inspire a better approximation of Φ, while the third source (from bad n) will present a tougher challenge (hence the "B" in [B1']–[B2']).

Let us now state [Wan21c]'s hypotheses more precisely. We first state a hypothesis that can morally be thought of *either as* an elementary GRH-on-average statement (in a certain range of parameters), *or as* an elementary manifestation of a zero-density hypothesis. Let Ψ_1 be an approximation of Φ, in the sense of Definition 3.2.9.

Definition 4.1.4. The *second-moment hypothesis* for Ψ_1 refers to the statement that for all $(Z,\beta)\in\mathbb{R}^2_{>0}$, for all (positive reals) $Y\le Z^3$, for all (positive reals) $N\le\beta Y$, and for all (real) intervals $I\subseteq[N/2,2N]$, we have

$$\sideset{}{'}\sum_{\boldsymbol{c}\in[-Z,Z]^m}\left|\sum_{n\in I}b_{\boldsymbol{c}}(n)\right|^2 \ll_{\beta,\epsilon} Z^\epsilon\max(Z^m,Y)\cdot N \qquad \text{(uniformly over } Z,Y,N,I),$$

where we restrict the $\boldsymbol{c}$-sum to $\{\boldsymbol{c}\in\mathbb{Z}^m : F^\vee(\boldsymbol{c})\ne 0\}$. (Here $Z^m\gg Y$, since $m\ge 3$. But we write $\max(Z^m,Y)$ in connection with the large sieve to be discussed soon.)

Remark 4.1.5 (Counting, or "density", interpretation). Since $m\ge 4$, the family $\mathcal{F}\approx[-Z,Z]^m$ is in fact *strictly* larger than the range of moduli, Y. Given N,I in Definition 4.1.4, we tolerate roughly $|\mathcal{F}|/N^{2\sigma-1}$ contributions of size $N^{\sigma-1/2}$. In particular, we are OK with roughly $|\mathcal{F}|/N$ "extremely exceptional" $\boldsymbol{c}$'s with $|\sum_{n\in I}b_{\boldsymbol{c}}(n)|\gg N$ matching the "heuristic trivial bound".[1] Contrast with "almost covering" problems (about e.g. small primes modulo c), where—in analogy with the "Sarnak–Xue density philosophy"—an "expander property" (e.g. a sufficiently strong prime number theorem) is often needed near $\sigma=1$.

Here $b_{\boldsymbol{c}}(n)$ "morally contains" a Möbius factor $\mu(n)$. But at least in favorable situations, the second-moment hypothesis for the Dirichlet series $\Psi_1=\sum_{n\ge1}b_{\boldsymbol{c}}(n)n^{-s}$ can be reduced to a large-sieve hypothesis involving the family of "friendlier" ("Möbius-free") Dirichlet coefficient vectors $(a_{\boldsymbol{c}}(n))_{n\ge1}$.

[1] In the full generality of Definition 3.2.9, we cannot always "actually trivially bound" $|\sum_{n\in I}b_{\boldsymbol{c}}(n)|$ for approximations Ψ_1 of Φ.

Definition 4.1.6. Call Ψ_1 *standard* if $b_{\boldsymbol{c}}(n), a_{\boldsymbol{c}}(n) \ll_\epsilon n^\epsilon$ holds uniformly over $\boldsymbol{c}, n$.

Example 4.1.7. Say Ψ_1 is defined as in Example 3.2.10. Then Ψ_1 is standard.

We now come to our "main" hypothesis: a large sieve in certain ranges.

Definition 4.1.8. If Ψ_1 is standard, then let $\gamma := \mathbf{1}_{n=\mathrm{rad}(n)} \cdot a$, or let $\gamma := b$; if Ψ_1 is non-standard, then let $\gamma := b$. We define the *large-sieve hypothesis* for γ to be the statement that for all $(Z, \beta) \in \mathbb{R}^2_{>0}$, and for all (positive reals) $Y \le Z^3$, we have

$$\sideset{}{'}\sum_{\boldsymbol{c} \in [-Z,Z]^m} \left| \sum_{n \le 2\beta Y} v_n \cdot \gamma_{\boldsymbol{c}}(n) \right|^2 \ll_{\beta,\epsilon} Z^\epsilon \max(Z^m, Y) \cdot \sum_{n \le 2\beta Y} |v_n|^2 \quad \text{for all } \boldsymbol{v} \in \mathbb{R}^{\lfloor 2\beta Y \rfloor},$$

uniformly over Z, Y. (Here $Z^m \gg Y$, since $m \ge 3$. But we write $\max(Z^m, Y)$ to avoid potential confusion with large sieves in other ranges of parameters.)

Remark 4.1.9. The large-sieve hypothesis for $\gamma := a$ (though excluded from Definition 4.1.8 for expository convenience) would directly imply that for $\gamma := \mathbf{1}_{n=\mathrm{rad}(n)} \cdot a$. The factor $\mathbf{1}_{n=\mathrm{rad}(n)}$ is simply "restriction to square-free moduli n".

Remark 4.1.10. For concreteness, say Ψ_1 is defined as in Example 3.2.10. Say $\gamma := \mathbf{1}_{n=\mathrm{rad}(n)} \cdot a$. Then $\boldsymbol{c} \mapsto (\gamma_{\boldsymbol{c}}(n))_{n \ge 1}$ is genuinely a family of Hasse–Weil coefficients, up to $\mathbf{1}_{n=\mathrm{rad}(n)}$. For Ψ_1, γ, it remains open to prove the associated large sieve (or satisfactory partial results towards it, if true). The difficulty (of the problem) seems unclear. One approach might be to dualize and try to adapt [Lou14], at least over $\mathbb{Z}[\zeta_3]$. Alternatively, one could try elementary geometric and analytic arguments.

For a discussion of large sieves in general (and what one can or cannot expect to be true), we refer the reader to [IK04, §7], [Dja12, §1], and [DR21, §1.4]. From the point of view of *geometric* families, a nice example of a large sieve is [HB95]'s quadratic-character large sieve ("optimal" if one restricts to square-free moduli, as is OK for most applications), applied in [PP97] to quadratic twist families of elliptic curves. (Interestingly, the conductors in the latter setting grow significantly faster than those in the former, but [HB95] applies equally well to the two settings, as far as the large sieve is concerned.)

Optimistically, one might hope that a "natural" geometric family—especially one with "big monodromy" (in some sense)—would have a good chance of satisfying an "optimal" large sieve. For a discussion of the (closely related) "expected random matrix symmetry type" of our present $V_{\boldsymbol{c}}$ families (with $2 \mid m$), we refer the reader to Chapter 8.

Remark 4.1.11. The philosophy behind [Wan21c] should also apply to other problems; see §9.1 for some potential examples in the spirit of the problems discussed so far. In geometric settings in particular, one might also wonder about the structure of the conjecturally relevant automorphic objects; Appendix A discusses a basic preliminary question in this direction.

4.2 Using average quadratic hypotheses

Throughout §4.2, let $n := 3$, and let $Q_{\boldsymbol{h}} := \sum_{i \in [n]} h_i y_i^2 \in \mathbb{Z}[\boldsymbol{y}]$ for each $\boldsymbol{h} \in \mathbb{Z}^n$. Also let $F_0(\boldsymbol{y}) := y_1^3 + \cdots + y_3^3$ and $F(\boldsymbol{x}) := x_1^3 + \cdots + x_6^3$.

4.2.1 A uniform conjecture on ternary quadratic equations

Conjecture 4.2.1. Uniformly *over reals* $X, H \geq 1$ *and pairs* $(\boldsymbol{h}, k) \in \mathbb{Z}^3 \times \mathbb{Z}$ *with* $\|\boldsymbol{h}\| \in [H, 2H]$ *and* $-h_1h_2h_3k \notin \{0\} \cup (\mathbb{Q}^\times)^2$*, we have*

$$\#\{\boldsymbol{y} \in \mathbb{Z}^3 \cap [-X, X]^3 : Q_{\boldsymbol{h}}(\boldsymbol{y}) = k\} \ll_\epsilon (XH)^\epsilon \cdot \left(\frac{X}{|h_1h_2h_3|^{1/3}} + \left(X^2H\right)^{1/4}\right).$$

Remark 4.2.2. In the positive-definite case $\boldsymbol{h} \in \mathbb{Z}^3_{>0}$, [Gol96, before the statements of the main results] mentions a similar "quite plausible" conjecture asserting the $O(k^{1/4})$-boundedness of the kth coefficient of the cusp form of weight $3/2$ corresponding to $Q_{\boldsymbol{h}}$, uniformly over $\boldsymbol{h}$.

Remark 4.2.3. Using [DFI93, HB96]'s delta method along the lines of [BKS19], one can likely give a *heuristic* argument for Conjecture 4.2.1, assuming "generic square-root cancellation" over the modulus q (for "generic" $\boldsymbol{c}$'s). Some details, including the analysis of the oscillatory integral for small q's, would likely require some care to flesh out (cf. [BKS19]), because $k \neq 0$.

(The "homogeneous case" $k = 0$ would be easier to analyze—at least under standard hypotheses—but we have excluded it from Conjecture 4.2.1 for simplicity.)

Remark 4.2.4. The $X/|h_1h_2h_3|^{1/3}$ comes from a loose real-density bound: note that

$$\begin{aligned}
&(2\epsilon)^{-1} \operatorname{vol}\left\{\boldsymbol{y} \in \prod_{i\in[3]}[X_i, 2X_i] : Q_{\boldsymbol{h}}(\boldsymbol{y}) \in [k-\epsilon, k+\epsilon]\right\} \\
&\ll \frac{X_1X_2X_3}{\max_{i\in[3]}(|h_i|X_i^2)} \leq \frac{X_1X_2X_3}{\prod_{i\in[3]}(|h_i|X_i^2)^{1/3}} = \frac{(X_1X_2X_3)^{1/3}}{|h_1h_2h_3|^{1/3}} \ll \frac{X}{|h_1h_2h_3|^{1/3}}
\end{aligned}$$

holds uniformly over $(\boldsymbol{X}, k, \epsilon) \in (0, X]^3 \times \mathbb{R} \times (0, 1]$, say. (For "lopsided" $\boldsymbol{h}$, the bound $X/|h_1h_2h_3|^{1/3}$ can likely be improved at the cost of cleanliness.)

Remark 4.2.5. Fix $(\boldsymbol{h}, k) \in \mathbb{Z}^3 \times \mathbb{Z}$ with $h_1 + h_2 + h_3 = 0$ and $k = -3F_0(\boldsymbol{h})$. Then $k = -9h_1h_2h_3$, so $-h_1h_2h_3k = (3h_1h_2h_3)^2$ is a square. And in fact, here $Q_{\boldsymbol{h}}(\boldsymbol{y}) = k$ does have many "trivial" solutions $\boldsymbol{y} \in \mathbb{Z}^3$, including $(t + 3h_2, t - 3h_1, t)$ for each $t \in \mathbb{Z}$. So in the range $X^\delta \leq \|\boldsymbol{h}\| \leq \delta X$, for instance, $Q_{\boldsymbol{h}}(\boldsymbol{y}) = k$ would have $\gg X$ solutions $\boldsymbol{y} \in [-X, X]^3$ as $X \to \infty$.

Thus in Conjecture 4.2.1, we need *some* nontrivial requirement on $(\boldsymbol{h}, k)$, even if $h_1h_2h_3 \neq 0$. Conjecture 4.2.1's assumption "$-h_1h_2h_3k \notin \{0\} \cup (\mathbb{Q}^\times)^2$" is thus quite natural.

4.2.2 A cubic application via differencing

To bound $N_F(X)$, we will first "dilute and fiber $N_F(X)$ into ternary affine quadrics" using multidimensional van der Corput differencing (inspired by [MV19, Lemma 3.3], which similarly reduced a quartic problem—in some ranges—to a problem about mixed cubic-quartic exponential sums) and then apply Conjecture 4.2.1 to bound most of the resulting fibers.

Remark 4.2.6. The basic idea of studying "quadratic fibers" of cubic Diophantine problems also appears in [Gol96, Woo13], for instance, and dates back at least to [Lin43]'s proof that $G(3) \leq 7$. (For more details on $G(3)$ and its history, see [VW02].)

Since Conjecture 4.2.1 does not apply to all $(\boldsymbol{h}, k) \in \mathbb{Z}^3 \times \mathbb{Z}$, we will also need the following conjecture:

Conjecture 4.2.7. *Uniformly over reals $H > 0$, we have*

$$\#\{(\boldsymbol{h}, z) \in (\mathbb{Z}^3 \setminus \{\boldsymbol{0}\}) \times \mathbb{Z} : \|\boldsymbol{h}\| \le H \text{ and } 3h_1h_2h_3(h_1^3 + h_2^3 + h_3^3) = z^2\} \ll_\epsilon H^{2+\epsilon}.$$

Remark 4.2.8. For fixed h_1, h_2 with $h_1h_2(h_1^3 + h_2^3) \neq 0$, the equation $z^2 = 3h_1h_2h_3F_0(\boldsymbol{h})$ in h_3, z cuts out a genus 1 affine plane curve, which contains the points $(0, 0)$ and $(-h_1 - h_2, \pm 3h_1h_2h_3)$—and thus defines an elliptic (in fact, a Mordell) curve, namely

$$y^2 = 3h_1h_2\left((h_1^3 + h_2^3)x^3 + 1\right), \quad \text{where } (x, y) := (1/h_3, z/h_3^2).$$

In particular, we may view $\{z^2 = 3h_1h_2h_3F_0(\boldsymbol{h})\}$ as an "elliptic fibration" over h_1, h_2. Conjecture 4.2.7 would thus follow from a certain standard hypothesis on the rank growth of elliptic curves over $\mathbb{Q}$ (which would itself follow from BSD and GRH, by [IK04, Proposition 5.21]); cf. [BB21, discussion after Theorem 1.1].

What can be said unconditionally towards Conjecture 4.2.7? (The fact that we are interested only in certain integral points, rather than all rational points, may help.)

Proposition 4.2.9. *Assume Conjectures 4.2.1 and 4.2.7. Then $N_F(X) \ll_\epsilon X^{45/13+\epsilon}$ as $X \to \infty$. (Here $45/13 = 3.4615\ldots < 7/2$.)*

Proof. Fix a small absolute constant $c > 0$ to be specified later. Then fix a nonzero nonnegative weight $\nu \in C_c^\infty(\mathbb{R}^n)$ supported on $[1, 1 + c]^n$. (Recall that $n := 3$.)

Now let $w(\boldsymbol{y}, \boldsymbol{z}) := \nu(\boldsymbol{y})\nu(-\boldsymbol{z})$. Then by Hölder and dyadic decomposition, it certainly suffices to show that $N_{F,w}(X) \ll_{\epsilon,\nu} X^{45/13+\epsilon}$. But

$$N_{F,w}(X) = \sum_{a \in \mathbb{Z}} N_{F_a,\nu}(X)^2.$$

Now fix a van der Corput differencing set $\mathcal{H} \subseteq \mathbb{Z}^n \cap [0, X]^n$, let $\nu_X := \nu \circ X^{-1}$ and $w_X := w \circ X^{-1}$ for convenience, and "dilute" $N_{F_a,\nu}(X)$ by $\mathcal{H}$ to get

$$N_{F_a,\nu}(X) := \sum_{\boldsymbol{x}_0 \in \mathbb{Z}^n} \nu_X(\boldsymbol{x}_0)\mathbf{1}_{F_0(\boldsymbol{x}_0)=a} = \sum_{\boldsymbol{x}_0 \in \mathbb{Z}^n} |\mathcal{H}|^{-1} \sum_{\boldsymbol{h} \in \mathcal{H}} \nu_X(\boldsymbol{x})\mathbf{1}_{F_0(\boldsymbol{x})=a},$$

where $\boldsymbol{x} := \boldsymbol{x}_0 + \boldsymbol{h}$. Then by Cauchy over $\boldsymbol{x}_0 \ll X$, it follows uniformly over $a \in \mathbb{Z}$ that

$$\begin{aligned} N_{F_a,\nu}(X)^2 &\ll X^n \sum_{\boldsymbol{x}_0 \in \mathbb{Z}^n} |\mathcal{H}|^{-2} \sum_{\boldsymbol{h}_1, \boldsymbol{h}_2 \in \mathcal{H}} \nu_X(\boldsymbol{x}_2)\mathbf{1}_{F_0(\boldsymbol{x}_2)=a} \cdot \nu_X(\boldsymbol{x}_1)\mathbf{1}_{F_0(\boldsymbol{x}_1)=a} \\ &= X^n|\mathcal{H}|^{-2} \sum_{\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{Z}^n} D(\boldsymbol{x}_2 - \boldsymbol{x}_1) \cdot \nu_X(\boldsymbol{x}_2)\mathbf{1}_{F_0(\boldsymbol{x}_2)=a} \cdot \nu_X(\boldsymbol{x}_1)\mathbf{1}_{F_0(\boldsymbol{x}_1)=a} \\ &= X^n|\mathcal{H}|^{-2} \sum_{\boldsymbol{h} \in \mathbb{Z}^n} D(\boldsymbol{h}) \sum_{\boldsymbol{x}_1 \in \mathbb{Z}^n} \nu_{X,\boldsymbol{h}}(\boldsymbol{x}_1)\mathbf{1}_{F_0(\boldsymbol{x}_1+\boldsymbol{h})=F_0(\boldsymbol{x}_1)=a}, \end{aligned}$$

where $D(\boldsymbol{h}) := \#\{(\boldsymbol{h}_1, \boldsymbol{h}_2) \in \mathcal{H}^2 : \boldsymbol{h} = \boldsymbol{h}_2 - \boldsymbol{h}_1\} \le \#\mathcal{H}$ and $\nu_{X,\boldsymbol{h}}(\boldsymbol{y}) := \nu_X(\boldsymbol{y}+\boldsymbol{h})\nu_X(\boldsymbol{y})$. Summing over $a \in \mathbb{Z}$, and letting $\boldsymbol{y} := \boldsymbol{x}_1$, leads to the bound

$$\begin{aligned} N_{F,w}(X) &\ll X^n |\mathcal{H}|^{-2} \sum_{\boldsymbol{h} \in \mathbb{Z}^n} D(\boldsymbol{h}) \sum_{\boldsymbol{y} \in \mathbb{Z}^n} \nu_{X,\boldsymbol{h}}(\boldsymbol{y}) \mathbf{1}_{F_0(\boldsymbol{y}+\boldsymbol{h}) = F_0(\boldsymbol{y})} \\ &\ll X^n |\mathcal{H}|^{-1} \sum_{\boldsymbol{h} \in \mathcal{H}-\mathcal{H}} \sum_{\boldsymbol{y} \in \mathbb{Z}^n} \mathbf{1}_{\boldsymbol{y}+\boldsymbol{h}, \boldsymbol{y} \in [X,(1+c)X]^n} \mathbf{1}_{F_{\boldsymbol{h}}(\boldsymbol{y}) = 0}, \end{aligned}$$

where $F_{\boldsymbol{h}}(\boldsymbol{y}) := F_0(\boldsymbol{y}+\boldsymbol{h}) - F_0(\boldsymbol{y})$.

Inspired by [MV19], we now let $K := cX^\theta$ for an exponent $\theta \in (0,1]$ to be specified, and let $\mathcal{H} := \{\boldsymbol{d} \in [0,cX] \times [0,cK]^{n-1} : 6 \mid \boldsymbol{d}\}$. The key point is that *if* $(\boldsymbol{h}, \boldsymbol{y}) \in (\mathcal{H}-\mathcal{H}) \times [X,(1+c)X]^n$ and $|h_1| \ge K$, then $|h_2|, |h_3| \le cK \le c|h_1|$, so the mean value theorem for $y \mapsto y^3$ implies

$$|F_{\boldsymbol{h}}(\boldsymbol{y})| \ge |h_1| \cdot (3 - o_{c\to 0}(1))X^2 - (|h_2|+|h_3|) \cdot (3 + o_{c \to 0}(1))X^2 \ge |h_1| X^2$$

(provided we chose c to be sufficiently small), whence $F_{\boldsymbol{h}}(\boldsymbol{y}) \ne 0$. Thus

$$N_{F,w}(X) \ll X^n |\mathcal{H}|^{-1} \sum_{\boldsymbol{h} \in \mathbb{Z}^n} \mathbf{1}_{\boldsymbol{h} \in [-K,K]^n} \mathbf{1}_{6 \mid \boldsymbol{h}} \sum_{\boldsymbol{y} \in \mathbb{Z}^n} \mathbf{1}_{\boldsymbol{y} \in [X,(1+c)X]^n} \mathbf{1}_{F_{\boldsymbol{h}}(\boldsymbol{y}) = 0}.$$

For each $\boldsymbol{h}, \boldsymbol{y}$ contributing to the sum above, let $\boldsymbol{h}' := \boldsymbol{h}/6 \in \mathbb{Z}^n \cap [-K/6, K/6]^n$ and $\boldsymbol{y}' := \boldsymbol{y} + 3\boldsymbol{h}' \in \mathbb{Z}^n \cap [(1-c)X, (1+2c)X]^n$; then

$$F_{\boldsymbol{h}}(\boldsymbol{y}) = \sum_{i \in [n]} [(y_i + h_i)^3 - y_i^3] = \sum_{i \in [n]} [3h_i(y_i + h_i/2)^2 + h_i^3/4] = 18 Q_{\boldsymbol{h}'}(\boldsymbol{y}') + 54 F_0(\boldsymbol{h}').$$

Let $k_{\boldsymbol{h}'} := -3F_0(\boldsymbol{h}')$; then by Conjectures 4.2.7 and 4.2.1, and the "trivial bound"

$$\begin{aligned} &\#\{\boldsymbol{y}' \in \mathbb{Z}^n \cap [-2X, 2X]^n : Q_{\boldsymbol{h}'}(\boldsymbol{y}') = k_{\boldsymbol{h}'}\} \\ &\ll X^n \mathbf{1}_{\boldsymbol{h}' = \mathbf{0}} + X^{n-1} \left(\sum_{\pi \in S_3} \mathbf{1}_{h'_{\pi(1)} = h'_{\pi(2)} = 0} + \sum_{\pi \in S_3} \mathbf{1}_{h'_{\pi(1)} = h'_{\pi(2)} + h'_{\pi(3)} = 0} \right) + O_\epsilon(X^{n-2+\epsilon}) \end{aligned}$$

(proven later, soon below), we conclude that $N_{F,w}(X)$ is

$$\begin{aligned} &\ll_{c,\epsilon} \frac{X^n (XK)^\epsilon}{1 + XK^{n-1}} \left(X^n + KX^{n-1} + K^2 X^{n-2} + \sum_{\boldsymbol{h}' \ll K}{}' \left(\frac{X}{|h'_1 h'_2 h'_3|^{1/3}} + \left(X^2 K\right)^{1/4} \right) \right) \\ &\ll_{c,\epsilon} \frac{X^{n+2\epsilon}}{1 + XK^{n-1}} \left(X^n + KX^{n-1} + K^2 X^{n-2} + K^{2n/3} X + K^n \left(X^2 K\right)^{1/4} \right), \end{aligned}$$

where the sum over $\boldsymbol{h}'$ is restricted to $\{-h'_1 h'_2 h'_3 k_{\boldsymbol{h}'} \notin \{0\} \cup (\mathbb{Q}^\times)^2\}$. Finally, recall that $n := 3$, and set $\theta := (4/13) \cdot (5/2) = 10/13$ to obtain the desired bound $N_{F,w}(X) \ll_\epsilon X^{45/13+\epsilon}$. $\square$

Loose ends. To derive the "trivial bound" for $\#\{\boldsymbol{y}' \ll X : Q_{\boldsymbol{h}'}(\boldsymbol{y}') = k_{\boldsymbol{h}'}\}$, note that

(1) uniformly over $(a,b,t) \in (\mathbb{Z}\setminus\{0\})^2 \times \mathbb{Z}$, we have

$$\#\{(u,v) \in [-2X,2X]^2 : au^2 + bv^2 = t\} \ll_\epsilon X \cdot \mathbf{1}_{t=0} + (|ab|X)^\epsilon,$$

by using $\tau(|t|) \cdot \mathbf{1}_{t\neq 0} \ll_\epsilon |t|^\epsilon$ if $-ab$ is a square, and by working in $\mathbb{Q}(\sqrt{-ab})$ otherwise (noting that p^e has at most $(e+1)^2 = \tau(p^e)^2$ ideal divisors in $\mathbb{Q}(\sqrt{-ab})$, and that if $-ab > 0$ then the fundamental unit ε is $\geq \min(1+\sqrt{2}, \frac{1+\sqrt{5}}{2})$);

(2) if $h'_1 h'_2 h'_3 \neq 0$, then $k_{\boldsymbol{h}'} - h'_3(y'_3)^2$ is *nonzero* for all but at most $O(1)$ integers y'_3;

(3) if $\pi \in S_3$ and $h'_{\pi(1)} = 0$, but $h'_{\pi(2)} h'_{\pi(3)} (h'_{\pi(2)} + h'_{\pi(3)}) \neq 0$, then $h'_{\pi(2)} h'_{\pi(3)} k_{\boldsymbol{h}'} \neq 0$; and

(4) if $\pi \in S_3$ and $h'_{\pi(1)} = h'_{\pi(2)}(h'_{\pi(2)} + h'_{\pi(3)}) = 0$, but $\boldsymbol{h}' \neq \mathbf{0}$, then $h'_{\pi(3)} \neq 0$, whence for each $y'_{\pi(1)}, y'_{\pi(2)}$ there are $\leq O(1)$ integers $y'_{\pi(3)}$ with $Q_{\boldsymbol{h}'}(\boldsymbol{y}') = k_{\boldsymbol{h}'}$.

The desired "trivial bound" follows upon combining (1)–(4). □

4.2.3 Commentary on the proof of the proposition

Remark 4.2.10. One can certainly relax the assumptions of Proposition 4.2.9. For example, an "ℓ^1-average" version of Conjecture 4.2.1 (over $\boldsymbol{h} \ll K$) would suffice in place of Conjecture 4.2.1. Also, we could relax the exponent in Conjecture 4.2.7 from $2+\epsilon$ to $2.6+\epsilon$.

In a more qualitative direction, note that $y \mapsto y^3$ is increasing. So if $\boldsymbol{h} \neq \mathbf{0}$, and $\boldsymbol{h} \geq 0$ (or $\boldsymbol{h} \leq 0$), then $F_{\boldsymbol{h}}(\boldsymbol{y}) > 0$ (resp. $F_{\boldsymbol{h}}(\boldsymbol{y}) < 0$) holds for all $\boldsymbol{y} \in \mathbb{R}^n$. Thus we only need Conjecture 4.2.1 in the *indefinite* case (i.e. when h_1, h_2, h_3 do not all have the same sign)—and in fact, we may further assume that $\max(\boldsymbol{h}) \asymp |\min(\boldsymbol{h})|$.

Remark 4.2.11. At the beginning, we localized to $x \in [X, (1+c)X]$. In fact, we could have localized to $|x - X| \leq X/\log X$ or $|x - X| \leq X^{1-\epsilon_0}$, say, but it is unclear if this would help.

Remark 4.2.12. For approximate lattices $\mathcal{H}$, the estimate $D(\boldsymbol{h}) \ll \#\mathcal{H}$ is close to the truth.

Remark 4.2.13. The "lopsided" choice of $\mathcal{H}$ in Proposition 4.2.9 saves a factor of roughly K/X over what we would have gotten from a more "uniform" choice of $\mathcal{H}$. This is essentially the observation behind the "averaged van der Corput differencing" in [MV19, §3.2]—which perhaps originated in work of Heath-Brown [HB07, §4]. The point is, in order for something like $h_1x_1^2 + h_2x_2^2 + h_3x_3^2$ to vanish (or be small at all), two of the $h_ix_i^2$ must be comparable. Above we only used a real (archimedean) version of this observation; is there a useful way to also introduce divisibility (non-archimedean) conditions on $\mathcal{H}$?

And is there a more conceptual justification for our choice of $\mathcal{H}$ above? Choosing $\mathcal{H}$ still seems to be an art. What is the "best" choice of $\mathcal{H}$? It may be worth trying $\mathcal{H} \approx [cX]^2 \times [cK]$, for instance—or more generally, $\mathcal{H} \approx [cX] \times [cK_2] \times [cK_3]$ with $X \gg K_2 \gg K_3$.

Remark 4.2.14. Instead of applying van der Corput differencing to $N_{F_a,\nu}(X)$ (once for each $a \in \mathbb{Z}$), we could have applied van der Corput differencing to $T(\theta)^n = T(\theta)^3$ (once for each $\theta \in \mathbb{R}/\mathbb{Z}$), where $T(\theta)$ denotes a weighted cubic Weyl sum. (Recall that $N_F(X)$ is the sixth moment of a certain cubic Weyl sum.) However, at least with the approach above, the end results do not seem to differ.

Remark 4.2.15. If one wanted to "amplify" the Hua problem (for $n = 3$) to some $n \geq 4$—and perhaps replace Proposition 4.2.9 with a question about n-variable quadratic forms (instead of ternary ones)—one would need to "restrict to minor arcs" (though the precise definition, including the initial choice of Dirichlet covering of $\mathbb{R}/\mathbb{Z}$, might require a little care).

Remark 4.2.16. Recall that Hua proved $N_F(X) \ll_\epsilon X^{7/2+\epsilon}$ by interpolating between (unconditional) results in L^4 and L^p for some $p \in [8, \infty]$. The simplest such "ingredients" might be Hua's lemma in L^4 and L^8—which can be proven by applying Weyl differencing once and twice, respectively. There are more elaborate arguments for the fourth and eight moments (roughly due to geometry and smooth numbers, respectively) that remove Hua's ϵ's, or even give log-power savings for $N_F(X)$ (see [Vau20, around Theorem 1.2] for details), but here we focus on classical differencing.

Thus in some sense, Hua used Weyl differencing "strictly between 1 and 2 times" for the sixth moment. And Proposition 4.2.9 can be viewed as an attempt to "perturb Hua" using a single (*more general*) van der Corput differencing *directly applied* to the sixth moment.

4.2.4 Final questions and remarks

Remark 4.2.17. [Mah36] used the identity $(9u^4)^3 + (3uv^3 - 9u^4)^3 + (v^4 - 9u^3v)^3 = v^{12}$ to show that $r_3(N) \gg N^{1/12}$ for $N = v^{12} > 0$. Although "localizing" to $x \approx X$ would rule out such solutions (one would need $9u^4 \approx 3uv^3 - 9u^4$, i.e. $v^3 \approx 6u^3$, but then $v^4 - 9u^3v \approx (6u^3 - 9u^3)v < 0$—a contradiction), this example suggests that some bases or fibrations are more biased than others. How biased is the "differenced" basis used in Proposition 4.2.9?

Remark 4.2.18. Suppose one applied the delta method directly on $r_3(N)$ to see what one gets assuming the most optimistic square-root cancellation over the modulus q (but no further cancellation over the dual variable $\boldsymbol{c}$). If $X \asymp N^{1/3}$, then one should expect roughly $r_3(N) \ll_\epsilon X^{3s/4-3/2+\epsilon} = X^{3/4+\epsilon}$, by extrapolating [HB98]'s work for $s = 4, 6$ to $s = 3$.

The "total heuristic bound" $\sum_{\boldsymbol{y} \ll X} r_3(F_0(\boldsymbol{y})) \ll_\epsilon X^{3.75+\epsilon}$ seems to arise naturally in other *exact*[2] "complete geometric exponential sum" approaches to bounding $N_F(X)$ as well:

(1) If one applies Deligne rather than GRH in [Hoo97, HB98], one gets $N_F(X) \ll_\epsilon X^{3.75+\epsilon}$ (rigorously, in fact—but unfortunately, this is worse than the Hua bound).

(2) The delta method heuristic for Conjecture 4.2.1 has an error term of $(X^2H)^{1/4+\epsilon}$,[3] which for $H \asymp X$ gives $X^{3/4+\epsilon}$ for each $\boldsymbol{h}$.

Is the $X^{3/4+\epsilon}$ commonality a red herring? It might be interesting to see if there is a deeper connection between the cubic hyperplane section Hasse–Weil L-functions of [Hoo86b, HB98] on the one hand, and the various Kloosterman–Salié sums and modular forms associated to quadratic problems arising from Weyl or van der Corput differencing on the other.

Remark 4.2.19. It may be interesting to try attacking other problems—including some of those listed in §9.1—by differencing or fibering.

[2] i.e. "dilution-free" (no nontrivial van der Corput differencing), etc.

[3] the heuristic being uniform in k, though $k = -3F_0(\boldsymbol{h})$ is what is relevant to us

Chapter 5

Biases in finite-field point counts

5.1 Introduction

The Weil conjectures imply in particular that point counts of smooth projective complete intersections over finite fields satisfy a certain randomness heuristic of “square-root cancellation” type. As [Hoo91, second paragraph after Theorem 2] notes, the same heuristic fails for *some* singular complete intersections, and it would be nice to have a “satisfactory criterion” to determine when, but “on the scanty evidence at present available, all we can say as yet is that it seems as if there were only a minority of singular varieties for which the result of the theorem cannot be improved.” Yet [Hoo91, Theorem 2] itself is silent on this issue (see Remark 5.1.7 below), and the full truth is far from known in general. Let us now summarize positive results of [Wan22] in this direction.

Given a base field k, an integer $m \geq 3$, and a homogeneous polynomial $F \in k[\boldsymbol{x}] = k[x_1, \ldots, x_m]$, let V denote $V_{\mathbb{P}^{m-1}}(F)$, and let $V_{\boldsymbol{c}}$ denote the intersection $V_{\mathbb{P}^{m-1}}(F, \boldsymbol{c} \cdot \boldsymbol{x})$ for $\boldsymbol{c} \in k^m$ (a hyperplane section of V if $\boldsymbol{c} \neq \boldsymbol{0}$). (We will repeatedly use this setup with $k, m, F, V, V_{\boldsymbol{c}}$; call it the *Main Setup* for convenience.) Our main general result, Theorem 5.1.14 below, shows that if k is finite, V is smooth, and $\deg F \geq 3$, then under mild conditions, the $V_{\boldsymbol{c}}$'s satisfy “square-root cancellation” for all $\boldsymbol{c}$ away from an explicit locus of codimension two. This result does not extend to $\deg F = 2$ in general (see Proposition 5.1.21), but its truth for $\deg F = 3$ is significant for Chapters 7–8.

Theorem 5.1.14 directly leads to general progress on Problem 5.1.2 below. In special cases, one can do better. The quadratic case is more or less understood (in odd characteristic, at the very least), so we focus on cubics. Corollary 5.1.31 shows that if k is finite, V is smooth, $\deg F = 3$, and $m \in \{4, 6\}$, then $V_{\boldsymbol{c}}$ fails “square-root cancellation” if and only if $(V_{\boldsymbol{c}})_{\overline{k}}$ contains a certain kind of subvariety of dimension $(m-2)/2$. This is at least morally significant for Chapter 6 on special subvarieties in Manin-type conjectures. This also leads to the following (vague) question:

Question 5.1.1. To what extent do special subvarieties in Manin's conjectures correlate with special subvarieties in the sense of the present chapter? For example, it would be interesting to determine whether the special quadratic locus $x_1 + x_2 + x_3 = x_4 + x_5 + x_6 = (x_1^2 + x_2^2 + x_3^2) - (x_4^2 + x_5^2 + x_6^2) = 0$ on the 6-variable quartic $x_1^4 + x_2^4 + x_3^4 = x_4^4 + x_5^4 + x_6^4$ (which I learned from a talk of Wooley; see [Woo19]) remains special for hyperplane sections of the quartic over finite fields.

Problem 5.1.2. Fix $G \in \mathbb{Z}[\boldsymbol{x}]$, homogeneous of degree $d \geq 2$ in $m \geq 3$ variables, with nonzero discriminant. Let $M_{d,m} := 72(3+2d)^m$. Given a prime p and integer $r \geq 1$, let $q := p^r$. Given $\boldsymbol{c} \in \mathbb{F}_q^m$, let $E_{\boldsymbol{c}}(q) := \#V_{\mathbb{P}^{m-1}}(G, \boldsymbol{c}\cdot\boldsymbol{x})(\mathbb{F}_q) - \#\mathbb{P}^{m-3}(\mathbb{F}_q)$. Let

$$N_{G,d,m}(q) := \#\{\boldsymbol{c} \in \mathbb{F}_q^m \setminus \{\boldsymbol{0}\} : |E_{\boldsymbol{c}}(q)| > M_{d,m} \cdot q^{(m-3)/2}\}.$$

As G, d, m vary, estimate the *exponent* $\sigma_{G,d,m} := \limsup_{p\to\infty} \sup_{r\to\infty} \log_q(N_{G,d,m}(q))$.

Remark 5.1.3. Problem 5.1.2 is close in spirit to [Lin20, Theorem 1.3 and the line after], who studies a different aspect (with q, m fixed and $d \to \infty$), works with a different family (a universal family of hypersurfaces), and works with a related cohomological problem instead of concrete point counts.

Remark 5.1.4. Morally, $\sigma_{G,d,m}$ is the "numerical dimension" of the *locus of $\boldsymbol{c}$'s for which square-root cancellation fails.* Most other natural measures of failure (e.g. those suggested by Proposition 5.1.9) are harsher, i.e. $\geq \sigma_{G,d,m}$. In any case, the Weil conjectures (plus Lang–Weil) give the bound $\sigma_{G,d,m} \leq m-1$ (as does [Hoo91, Theorem 2]), which we improve to $\sigma_{G,d,m} \leq m-2$ for $d \geq 3$ in general (see Remark 5.1.15), and to $\sigma_{G,d,m} \leq m-3$ in many cases with $(m,d) = (6,3)$ (see e.g. Example 5.1.32, where $m \in \{4,6\}$ and $\sigma_{F,3,m} = m/2$). Based on this, it would be reasonable to conjecture that $\sigma_{G,d,m} \leq m-3$, or perhaps even more, holds when $d \geq 3$ and $m \geq 6$ (if one believes that "randomness" should increase with d, m).

We now introduce some relevant conventions and notions.

Definition 5.1.5. Let k be a base field. A *k-scheme* is a scheme equipped with a morphism to $\operatorname{Spec} k$. A *variety over k* (or *k-variety* for short) is a separated k-scheme of finite type (not necessarily reduced or irreducible). In the context of projective varieties, and especially their singular loci, let $\dim(\emptyset) := -1$ and $\mathbb{P}_k^{-1} := \emptyset$. If X is a projective k-variety and k is finite, let $E(X) := \#X(k) - \#\mathbb{P}^{\dim X}(k)$. (The base field k of X is essential to this definition.)

For a k-variety X of pure dimension $d \geq 0$, let X_{sing} denote the *singular subscheme* of X, i.e. the closed subscheme of X cut out by the dth Fitting ideal of the cotangent sheaf $\Omega_{X/k}$ (or informally, "by the Jacobian criterion"), following [Sta22, Tag 0C3H]. (Most important to us is $X_{\mathrm{sing}}(\overline{k})$, the set of singular $\overline{k}$-points of X.) For a scheme Y, we let $|Y|$ denote the underlying topological space; thus, for instance, $|X_{\mathrm{sing}}|$ denotes a topological space, while $|X_{\mathrm{sing}}(\overline{k})| = \#X_{\mathrm{sing}}(\overline{k})$ denotes an integer.

Our main concern is the notion of *error-goodness* (or *$|E|$-goodness* for short), which we now define alongside some related notions (related by Proposition 5.1.9).

Definition 5.1.6. Let k denote an arbitrary finite field, and X a projective k-variety.

(1) Given $f \in \{|E|, +E, -E\}$, say X is *f-good* (with constant C) if there exists $C \in \mathbb{R}_{>0}$ such that for all finite extensions k'/k, we have $f(X_{k'}) \leq C|k'|^{(\dim X)/2}$. Say X is *f-bad* if it is not f-good.

(2) Given a property *blah* in (1), say X is *potentially* (resp. *stably*) *blah* if $X_{k'}$ is blah for some (resp. for every) finite extension k'/k.

Remark 5.1.7. Fix a projective complete intersection Y/k. Then [Hoo91, Theorem 2] gives $|E(Y_{k'})| \le O(|k'|^{(1+\dim(Y_{\mathrm{sing}})+\dim Y)/2})$ as $[k':k] \to \infty$. Like the Weil conjectures, this only proves that Y is $|E|$-good if $\dim(Y_{\mathrm{sing}}) = -1$, i.e. Y is smooth.

For reference, we recall a useful principle of Zak:

Theorem 5.1.8 (See e.g. [Hoo91, Katz's Appendix, Theorem 2])**.** *Let $k, m, F, V, V_{\boldsymbol{c}}$ be as in the Main Setup. If V is smooth, then $\dim((V_{\boldsymbol{c}})_{\mathrm{sing}}) \le 0$ for all $\boldsymbol{c} \in k^m \setminus \{\boldsymbol{0}\}$.*

We now state a useful amplification-type result. The proof (see [Wan22, §2]) uses general foundational results due to Katz and others (see especially Theorem 5.2.8).

Proposition 5.1.9. *Let k be a finite field. Let X be a projective k-variety of the form $V(F_1, \dots, F_r) \subseteq \mathbb{P}^n_k$, with $n, r \ge 1$ and $\max_{i \in [r]} \deg F_i \le d$. Consider the following four conditions:*

(1) X is $|E|$-good with constant $18(3+rd)^{n+1}2^r$;

(2) X is $|E|$-good;

(3) X is potentially $|E|$-good;

(4) X is potentially $(-1)^{1+\dim X}E$-good;

In general, (1)–(2) are equivalent, (2) implies (3), and (3) implies (4). If $\operatorname{codim} X = r$ and $\dim(X_{\mathrm{sing}}) \le 0$, then (1)–(4) are equivalent.

Remark 5.1.10. In particular, a projective complete intersection Y/k with $\dim(Y_{\mathrm{sing}}) \le 0$ is $|E|$-good if and only if it is potentially $|E|$-good. It would be nice to have this more generally.

To state our main general result, Theorem 5.1.14, we need two definitions.

Definition 5.1.11. For a variety of pure dimension $n-1$ over $K = \overline{K}$, an A_1 *singularity* is a point at which the completed local ring is $\cong K[[z_1, \dots, z_n]]/(z_1^2 + \cdots + z_n^2)$.

Remark 5.1.12. In characteristic $p \ne 2$, the following notions coincide: A_1 singularity, non-degenerate double point, and ordinary double point. (See e.g. [PS20, §4].)

Definition 5.1.13. Given integers $d \ge 2$ and $m \ge 3$, let $\operatorname{disc}(P, \boldsymbol{a})$ be a discriminant polynomial associated to the "universal" intersection $V(P, \boldsymbol{a} \cdot \boldsymbol{x}) \subseteq \mathbb{P}^{m-1}$ defined by the "universal" homogeneous polynomials $P(\boldsymbol{x}), \boldsymbol{a} \cdot \boldsymbol{x}$ of respective degrees $d, 1$ in $\boldsymbol{x} = (x_1, \dots, x_m)$. (See e.g. [Ter18, §1.1], or [Wan22, Appendix A], for details.)

Theorem 5.1.14. *In the Main Setup with $k, m, F, V, V_{\boldsymbol{c}}$, suppose V is smooth, $d := \deg F \ge 3$, and k is finite of characteristic $p \ne 2$. Let $\boldsymbol{c} \in k^m \setminus \{\boldsymbol{0}\}$, and suppose $V_{\boldsymbol{c}}$ is $|E|$-bad (see Definition 5.1.6 and Proposition 5.1.9). Then the following hold:*

(1) Either $V_{\boldsymbol{c}} \times \overline{k}$ has ≥ 2 singularities, or it has a non-A_1 singularity.

(2) If $p \nmid d(d-1)$, then $\boldsymbol{c}$ is a singular zero of the polynomial $\operatorname{disc}(F, -)$.

Remark 5.1.15. The significance of (2) comes from its codimension-two nature; when combined with Proposition 5.1.9 and Lang–Weil, it gives $\sigma_{G,d,m} \le m-2$ in Problem 5.1.2 for $d \ge 3$. Though (2) is easier to use in certain applications, (1) can also be useful in view of [PS20, Theorem 1.1].

Remark 5.1.16 (Sawin). One could perhaps prove a less explicit version of Theorem 5.1.14(2) using the perversity-based strategy of [GS21, proof of Lemma 3.1].

Remark 5.1.17. Theorem 5.1.14(2) may have an analog for the universal family of hypersurfaces of a given degree ≥ 3 and dimension ≥ 0. We have not checked.

The proof of Theorem 5.1.14 (see [Wan22, §3]) mainly combines results on discriminants, duality, and ℓ-adic cohomology (especially [Lin20, Theorem 1.2]). When $(m,d)=(6,3)$, one can replace the ingredient [Lin20, Theorem 1.2] with the following proposition, proven later (after Lemma 5.2.10) using some geometry (going back to Clemens–Griffiths) specific to cubic threefolds.

Proposition 5.1.18. *Let X be a projective cubic threefold over a finite field k. Assume that $X_{\mathrm{sing}}(\overline{k}) = \{x\}$, with $x \in X(\overline{k})$ either nodal or mildly cuspidal (i.e. either of analytic type $z_1z_2 + z_3z_4 = 0$ or $z_1z_2 + z_3^2 + z_4^3 = 0$; see e.g.* [CML09, Corollary 3.2(i)–(ii)] *and* [vdGK10, paragraph before Proposition 2.1] *for the equivalence). Then x is defined over k, and X is $|E|$-good.*

Theorem 5.1.14 has the following corollary (proven in [Wan22, §4] by point counting), which when $2 \mid m$ reproves a consequence of the Deligne–Katz equidistribution theorem. (The case $2 \nmid m$ may or may not involve "exceptional monodromy", which could complicate an attempt to use Deligne–Katz.) It suggests that there may be a deeper connection between monodromy, moments, and Problem 5.1.2.

Corollary 5.1.19. *Let $G, d, m, E_{\boldsymbol{c}}(q)$ be as in Problem 5.1.2. If $d, m \ge 3$, then as $p \to \infty$, the following hold (with implied constants depending only on G):*

(1) $\mathbb{E}_{\boldsymbol{c}\in\mathbb{F}_p^m}[E_{\boldsymbol{c}}(p)\cdot \mathbf{1}_{p\nmid \mathrm{disc}(G,\boldsymbol{c})}] = p^{-1}\cdot E(V_{\mathbb{P}^{m-1}}(G)_{/\mathbb{F}_p}) + p^{-1}\cdot O(p^{(m-3)/2})$.

(2) $\mathbb{E}_{\boldsymbol{c}\in\mathbb{F}_p^m}[E_{\boldsymbol{c}}(p^2)\cdot \mathbf{1}_{p\nmid \mathrm{disc}(G,\boldsymbol{c})}] = (1+O(p^{-1/2}))\cdot p^{m-3}$.

(3) $\mathbb{E}_{\boldsymbol{c}\in\mathbb{F}_p^m}[E_{\boldsymbol{c}}(p)^2\cdot \mathbf{1}_{p\nmid \mathrm{disc}(G,\boldsymbol{c})}] = (1+O(p^{-1/2}))\cdot p^{m-3}$.

The next result shows that the assumption $d \ge 3$ in Theorem 5.1.14 is essential when $2 \mid m$; take $\dim(X) = m-3 \equiv 1 \bmod 2$ and $d = \dim(X)-1$ to see why.

Definition 5.1.20. Let k be a base field. A *cone* is a projective cone over a projective k-variety, with vertex a k-point. (Informally, an embedded projective k-variety is a cone if and only if it is "missing a variable" after some k-linear change of coordinates. For algebraic convenience, we will consider a hypersurface $X \subseteq \mathbb{P}^1_k$ with $\#X(k) = \#X(\overline{k}) = 1$ to be a cone.) An *iterated cone* is obtained by taking cones one or more times.

Proposition 5.1.21 (Quadric dichotomy)**.** *Let X be a projective quadric of dimension ≥ 0 over a finite field k. Then X is stably $|E|$-bad if and only if $X_{\overline{k}}$ is an iterated cone over a smooth projective quadric of dimension $d \in [0, \dim(X)-1]$ with $2 \mid d$.*

Remark 5.1.22. If $2 \nmid |k|$, then the adverb "stably" can be removed, and the cone condition is equivalent to "X is of the form $V(Q)$ with $2 \mid \operatorname{rank}(Q) \in [2, \dim(X)+1]$".

Proposition 5.1.21 is essentially classical. We will sketch a proof using the following general fact:

Proposition 5.1.23 (Routine)**.** *Let k be a finite field. If $C(Y)$ is a cone over a projective k-variety Y, then $E(C(Y)) = |k| \cdot E(Y)$.*

Proof. Proof by calculation (one could probably alternatively use cohomology somehow): $|C(Y)(k)| = 1 + |k| \cdot |Y(k)|$ ("vertex + lines through vertex") and likewise $|\mathbb{P}^{\dim C(Y)}| = |C(\mathbb{P}^{\dim Y})| = 1 + |k| \cdot |\mathbb{P}^{\dim Y}|$; now subtract. $\square$

Proof sketch for Proposition 5.1.21. The key ingredients are the following:

(1) If X is smooth of dimension $d \geq 0$, then $E(X) = 0$ if $2 \nmid d$, and $E(X) = \pm |k|^{d/2}$ if $2 \mid d$. (One can prove this using the Weil conjectures. If $2 \nmid |k|$, one can alternatively diagonalize and then follow e.g. [Wei49].)

(2) If X is non-reduced, then $E(X) = 0$. On the other hand, a reduced projective $\overline{k}$-quadric (of dimension ≥ 0) is singular if and only if it is an iterated cone over a smooth projective quadric (of dimension ≥ 0).

By (1)–(2) and Proposition 5.1.23, X is potentially $|E|$-good if and only if $X_{\overline{k}}$ is either smooth or non-reduced, or an iterated cone over a smooth projective quadric of dimension $d \in [0, \dim(X) - 1]$ with $2 \nmid d$. The result follows from (2). $\square$

The following results show that Theorem 5.1.14 is far from the full truth in general.

Definition 5.1.24. A $\boldsymbol{d}$-*CI* is a (projective) complete intersection of multi-degree $\boldsymbol{d}$.

Definition 5.1.25. Let $d \geq 1$ be an integer. Over $K = \overline{K}$, a *cubic d-scroll* is an embedded projective d-fold $\Sigma \subseteq \mathbb{P}^n$, integral of degree 3, with $\dim(\operatorname{Span}(\Sigma)) = d + 2$. (See Proposition 5.2.4 below, and its proof, for some background on cubic scrolls.)

Theorem 5.1.26. *Let $n \in \{2, 4\}$. Let X be a (3)-CI in $\mathbb{P}^n$ over a finite field k. Suppose $X_{\overline{k}}$ is not a cone over a cone. If $n = 2$, then X is $|E|$-bad if and only if $X_{\overline{k}}$ contains a line. Now suppose $n = 4$, and consider the following four conditions:*

(1) X is stably E-bad (see Definition 5.1.6 and Proposition 5.1.9);

(2) X is stably $|E|$-bad;

(3) $X_{\overline{k}}$ contains a plane or a singular cubic 2-scroll in $\mathbb{P}^4_{\overline{k}}$; and

(4) there exists a $(2,2)$-CI of the form $V(Q_1, Q_2) \subseteq \mathbb{P}^4_{\overline{k}}$, with $V(Q_1)_{\mathrm{sing}} \cap V(Q_2)_{\mathrm{sing}} \neq \emptyset$, such that $X_{\overline{k}}$ contains a nonempty open subscheme of $V(Q_1, Q_2)$.

In general, (1)–(2) are equivalent, (2) implies (3), and (3)–(4) are equivalent. If $\dim(X_{\mathrm{sing}}) \leq 0$, then (1)–(4) are equivalent.

Remark 5.1.27. The case when $X_{\overline{k}}$ is a cone over a cone can still be fully analyzed (using Proposition 5.1.23), but it is less interesting than the opposite case.

Remark 5.1.28. The equivalence of (1)–(2) suggests that (stable) $|E|$-badness might be explained by "excess" points from "special" subvarieties, which we have tried to pinpoint in (3)–(4). Theorem 5.1.26 is close to a complete dichotomy. To "complete" it (for $n = 4$), one would need to analyze the case $\dim(X_{\mathrm{sing}}) \geq 1$, which might be tricky (see e.g. [Wan22, proof of Lemma 5.5]).

Remark 5.1.29. Our methods can also be used to show that a projective cubic surface X/k is stably $|E|$-bad if and only if $X_{\overline{k}}$ is either reducible, or a cone over a smooth cubic curve. We omit this from Theorem 5.1.26 because it has a different flavor.

The proof of Theorem 5.1.26 (see [Wan22, §5]) uses (for $n = 4$) base change and some situation-specific geometry, including some classification results over $\overline{k}$ (see e.g. Proposition 5.2.1 and Lemma 5.2.2 below). It would be interesting to find different proofs of Theorem 5.1.26 that work directly over k, or that generalize naturally. Condition (4) in Theorem 5.1.26 seems especially suggestive as to what one might try more generally. One might also try using mixed Hodge theory, in the spirit of e.g. [Dim90, Klo22].

Remark 5.1.30. Originally we sought to prove Theorem 5.1.26 (for $n = 4$) using (an extension of) [BSD67]'s "conic bundle" method; see §5.3 below (and specifically, Remark 5.3.16). This approach inspired condition (4). But our present approach is overall more efficient in the singular case.

Corollary 5.1.31. *In the Main Setup, suppose F is a cubic form in $m \in \{4, 6\}$ variables over a finite field k, and V is smooth. Let $\boldsymbol{c} \in k^m \setminus \{\mathbf{0}\}$. Then $V_{\boldsymbol{c}}$ is $|E|$-bad if and only if $(V_{\boldsymbol{c}})_{\overline{k}}$ contains an $(m-2)/2$-plane or a singular cubic 2-scroll in $\mathbb{P}^{m-1}_{\overline{k}}$.*

Proof. Combine Theorem 5.1.8, Proposition 5.1.9, and Theorem 5.1.26. $\square$

Example 5.1.32. Let $m \in \{4, 6\}$ and $\boldsymbol{F} = (F_1, \ldots, F_m) \in (\mathbb{Z} \setminus \{0\})^m$. Suppose $F = F_1x_1^3 + \cdots + F_mx_m^3$, and assume the characteristic of k is sufficiently large in terms of $\boldsymbol{F}$. Then by [Wan22, Proposition B.3] (proven by a calculation—a singularity analysis—involving, among other things, 3×3 Vandermonde determinants arising from diagonality), the phrase "or a singular cubic 2-scroll" in Corollary 5.1.31 is unnecessary for F. Furthermore, the $(m-2)/2$-planes on $V_{\overline{k}}$ are known to be cut out by systems of equations of the form "$F_ix_i^3 + F_jx_j^3 = 0$ in pairs". Thus a given $V_{\boldsymbol{c}}$ is $|E|$-bad if and only if "$c_i^3/F_i = c_j^3/F_j$ in pairs".

Our proof of [Wan22, Proposition B.3] makes use of the following technical fact (though it would be nice to know how much the hypotheses can be weakened; cf. [Bri15]):

Proposition 5.1.33. *Let X_1, X_2 be subvarieties of $\mathbb{P}^n$ of pure dimension d over $K = \overline{K}$, where $d, n \geq 1$. Let $X := X_1 \cup X_2$. Assume the following hypotheses:*

(1) $\dim(X_1 \cap X_2) = 0$*; and*

(2) for each $x \in X_1 \cap X_2$, there exist a subvariety Y of $\mathbb{P}^n$ of pure dimension $2d$, and an open neighborhood U of x in $\mathbb{P}^n$, such that $X_1 \cap U$ and $X_2 \cap U$ are Cohen–Macaulay, $Y \cap U$ is smooth, and $Y \cap U \supseteq X \cap U$.

Then $X_{\mathrm{sing}} \supseteq X_1 \cap X_2$.

Proof. The statement is local, so we may assume $X_1 \cap X_2$ is supported on a singleton $\{x\}$. Now let Y, U be as in hypothesis (2); by shrinking U if necessary, we may assume that U is affine and that X_1, X_2, Y are closed subschemes of U. Say $U = \operatorname{Spec} R$, and let $I_1, I_2, J \subseteq R$ be the ideals defining X_1, X_2, Y, respectively. Then X_1, X_2 are Cohen–Macaulay, Y is regular, and $J \subseteq I_1 \cap I_2$. By [Spe10] and [Ser00, Proposition 11 in §IV.B.1, and Corollary to Theorem 4 in §V.B.6], it follows that $(I_1/J) \cap (I_2/J) = (I_1/J) \cdot (I_2/J)$ in R/J (cf. [Dao22]). So if $f \in I_1 \cap I_2$, then $f \equiv h \bmod I_1 I_2$ for some $h \in J$, whence $Df \equiv Dh \bmod (I_1, I_2)$ for all derivations $D \colon R \to R$. But $\Omega_{Y/K}$ is locally free of rank $2d \geq d+1$, so the dth Fitting ideal of $\Omega_{Y/K}$ is 0. Thus $V_U(I_1 \cap I_2)_{\mathrm{sing}} \supseteq V_U(I_1, I_2)$, i.e. $X_{\mathrm{sing}} \supseteq X_1 \cap X_2$. □

Problem 5.1.2 and Corollary 5.1.31 motivate the following question:

Question 5.1.34. Given a smooth cubic hypersurface $X \subseteq \mathbb{P}^5_{\mathbb{C}}$, let $S \subseteq (\mathbb{P}^5_{\mathbb{C}})^\vee$ parameterize hyperplane sections of X containing a singular cubic 2-scroll in $\mathbb{P}^5_{\mathbb{C}}$. What is the best possible upper bound on the dimension of the Zariski closure of S?

Remark 5.1.35. It is known that $|S| = 1$ for sufficiently general X containing a (not necessarily singular) cubic 2-scroll; cf. especially [Has96, proof of Lemma 2.11, and the subsequent dimension counting] and [HT10, Propositions 3.3 and 6.1]. And S may well be finite for $X = V_{\mathbb{P}^5}(x_1^3 + \cdots + x_6^3)_{/\mathbb{C}}$, though our analysis in [Wan22, Appendix B] falls short of a proof (due to interference from the planes on X).

5.2 Miscellaneous writeups

The proof of Theorem 5.1.26 (see [Wan22, §5]) begins with a classical "rationality"-type idea (cf. [Dol16, §1]).

Proposition 5.2.1. *Let* $X \subseteq \mathbb{P}^n$ *be a* (3)*-CI over a finite field* k*, where* $n \geq 2$*. Assume* $[0 : \cdots : 0 : 1] \in X_{\mathrm{sing}}$*. Then* $X = V(f_2 x_{n+1} + f_3)$ *for some* $f_i \in k[x_1, \ldots, x_n]$*, homogeneous of degree* i*, with* $(f_2, f_3) \neq (0,0)$*. Furthermore,* $E(X) = |k| \cdot E(V_{\mathbb{P}^{n-1}}(f_2, f_3)) - E(V_{\mathbb{P}^{n-1}}(f_2)) + |k|^{n-1} \cdot \mathbf{1}_{\dim V(f_2,f_3) = \dim V(f_2)}$*.*

Proof. The first part is clear. So there are two kinds of points $[\boldsymbol{x}] \in X(k)$: (i) those with $f_2 \neq 0$ and $x_{n+1} = -f_3/f_2$, and (ii) those with $f_2 = 0$ and $f_3 = 0$. Therefore $|X(k)| = |(\mathbb{P}^{n-1} \setminus V(f_2))(k)| + |C(V_{\mathbb{P}^{n-1}}(f_2, f_3))(k)|$. Proposition 5.1.23, and casework on $\dim V(f_2) - \dim V(f_2, f_3) \in \{0, 1\}$, then lead to the desired equality. □

To study $E(X)$ using Proposition 5.2.1, one needs to analyze low-degree complete intersections in some detail. [Wan22, §5] repeatedly uses the following lemma describing the low-degree components of non-integral $(2,2)$-CI's and $(2,3)$-CI's:

Lemma 5.2.2. *Let* $n \geq 2$ *and* $K = \overline{K}$*. Let* Y *be a* (d,e)*-CI of the form* $V(A,B) \subseteq \mathbb{P}^n_K$*, with* $d = 2$ *and* $e \in \{2, 3\}$*. Then* Y *is non-integral if and only if it has an irreducible component of degree* ≤ 3*. Let* Z *be any such component, equipped with the reduced induced scheme structure. Then the following dichotomy holds:*

(1) If $\deg Z \le 2$*, or* $e = 3$ *and* A *is reducible, then* $\dim(\mathrm{Span}(Z)) \le n-1$.

(2) If $\deg Z = 3$*, and* $e = 2$ *or* A *is irreducible, then* $\dim(\mathrm{Span}(Z)) = n$.

Furthermore, $\dim(\mathrm{Span}(Z)) \le n-1$ *if and only if* Z *is a* $(1, \deg Z)$*-CI.*

Proof. The first part is clear, since $\deg Y = de \le 6$. Now fix Z. Since $\dim Z = n-2$, and Z is integral, the final sentence is clear: both conditions are equivalent to "Z lies in an $(n-1)$-plane (scheme-theoretically)". So it remains to prove (1)–(2).

If $\deg Z \le 2$, then $\dim(\mathrm{Span}(Z)) \le \dim(Z) + 1 = n-1$ by [EH87, Proposition 0]. This proves (1)–(2) when $\deg Z \le 2$.

Now suppose $\deg Z = 3$. If A is reducible, then since Z is integral, there must exist a nontrivial (and thus linear) factor $L \mid A$ such that $V(L, B) \supseteq Z$; and thus $e = 3$ and $Z = V(L, B) \subseteq V(L)$. Conversely, suppose $Z = V(L, C)$ for some nonzero linear form L and cubic form C (with $L \nmid C$). Then $V(A, B) \supseteq V(L, C)$, i.e. A, B lie in the saturated homogeneous ideal (L, C). So for degree reasons, $L \mid A$ and $e = 3$ (or else $L \mid A, B$, which is impossible).

Thus we have shown that if $\deg Z = 3$, then Z is a $(1,3)$-CI if and only if A is reducible, in which case $e = 3$ must hold. This proves (1)–(2) when $\deg Z = 3$. □

Let us now record some results and proofs of a folklore nature.

Remark 5.2.3. Technical points (some implicit) in [Wan22] include taking care to work with saturated homogeneous ideals when necessary (as in the proof of Lemma 5.2.2); also recall that for homogeneous ideals, prime implies radical implies saturated.

Proposition 5.2.4. *For a cubic scroll* Σ *over* $K = \overline{K}$*, the following hold:*

(1) Σ *contains a nonempty open subscheme of a* $(2,2)$*-CI in* $\mathrm{Span}(\Sigma)$*.*

(2) Σ *is singular if and only if it is a cone, in which case it is a cone over a cubic scroll of dimension* $\dim(\Sigma) - 1$*.*

Proof. By Definition 5.1.25, Σ is, in the sense of [EH87], a "variety of minimal degree" in $\mathrm{Span}(\Sigma) \cong \mathbb{P}^{d+2}$, where $d := \dim(\Sigma)$. Now inspect the statement [EH87, Theorem 1] and elaboration [EH87, pp. 4–6 of §1]. We find that Σ is, in the language of [EH87], a "rational normal scroll" of the form $S(a_1, \dots, a_d)$, where $a_1, \dots, a_d$ are integers with $0 \le a_1 \le \cdots \le a_d$ and $a_1 + \cdots + a_d = 3$; that $S(a_1, \dots, a_d)$ is smooth if $a_1 \ge 1$, and a cone over $S(a_2, \dots, a_d)$ if $a_1 = 0$; and that the schemes $S(1,1,1) \subseteq \mathbb{P}^5$, $S(1,2) \subseteq \mathbb{P}^4$, and $S(3) \subseteq \mathbb{P}^3$ are (up to linear change of coordinates) defined by the homogeneous ideals $(x_3x_6 - x_4x_5, x_1x_6 - x_2x_5, x_1x_4 - x_2x_3)$, $(x_3x_5 - x_4^2, x_1x_5 - x_2x_4, x_1x_4 - x_2x_3)$, and $(x_2x_4 - x_3^2, x_1x_4 - x_2x_3, x_1x_3 - x_2^2)$, respectively.

Claim (1) follows, since $V_{\mathbb{P}^5}(x_1x_6 - x_2x_5, x_1x_4 - x_2x_3)$, $V_{\mathbb{P}^4}(x_1x_5 - x_2x_4, x_1x_4 - x_2x_3)$, and $V_{\mathbb{P}^3}(x_1x_4 - x_2x_3, x_1x_3 - x_2^2)$ are $(2,2)$-CI's that coincide with $S(1,1,1)$, $S(1,2)$, and $S(3)$, respectively, away from $V(x_1)$. It also follows that if Σ is singular, then it is a cone. On the other hand, if Σ is a cone, say over Y, then $Y \subseteq \Sigma$ is an integral $(d-1)$-fold with $\deg Y = \deg \Sigma = 3$ and $\dim(\mathrm{Span}(Y)) = \dim(\mathrm{Span}(\Sigma)) - 1 = d+1$, i.e. Y is a cubic $(d-1)$-scroll; so by [EH87, Theorem 1] and [EH87, pp. 4–6 of §1], we have $d - 1 \ge 1$, and $Y \cong S(b_1, \dots, b_{d-1})$ and $\Sigma \cong S(0, b_1, \dots, b_{d-1})$ for some integers $b_1, \dots, b_{d-1}$; and thus Σ is singular. □

Proposition 5.2.5. *Let $S \subseteq \mathbb{P}^3$ be a (3)-CI over $K = \overline{K}$. Then S contains a line.*

Proof. See [Mus17, Theorem 1.1]. □

For the rest of this section, let k denote a finite field.

Definition 5.2.6. Fix a prime $\ell \nmid |k|$. Let Y, X denote arbitrary projective k-varieties.

(1) Define $H^\bullet(Y) := H^\bullet(Y \times \overline{k}, \mathbb{Q}_\ell)$ using ℓ-adic cohomology with $\mathbb{Q}_\ell$-coefficients.

(2) For each $i \geq 0$, let $\mathcal{E}^i(Y)$ denote the multiset of (geometric) Frobenius eigenvalues on $H^i(Y)$.

(3) Let $\mathcal{E}^i_\triangle(X, \mathbb{P}) := (\mathcal{E}^i(X) \cup \mathcal{E}^i(\mathbb{P}^{\dim X}_k)) \setminus (\mathcal{E}^i(X) \cap \mathcal{E}^i(\mathbb{P}^{\dim X}_k))$ for $i \geq 0$. In other words, if $\alpha \in \overline{\mathbb{Q}}_\ell$ has multiplicities $j_1, j_2 \geq 0$ in $\mathcal{E}^i(X), \mathcal{E}^i(\mathbb{P}^{\dim X}_k)$, let it have multiplicity $|j_1 - j_2|$ in $\mathcal{E}^i_\triangle(X, \mathbb{P})$.

(4) Define $\mathcal{E}_\triangle(X, \mathbb{P}) := \sum_{i \geq 0} \mathcal{E}^i_\triangle(X, \mathbb{P})$ by "summing multiplicities" over i.

Remark 5.2.7. Deligne's purity theorem implies that each $\mathcal{E}^i(Y)$ above consists of $|k|$-Weil numbers $\alpha \in \overline{\mathbb{Q}}_\ell$ of weight $w(\alpha) \leq i$. (See e.g. [KW01, Remark I.7.2 and Theorem I.9.3(2)] for a precise textbook reference in English.)

The following statement combines [Hoo91, Katz's Appendix, assertion (2) in the proof of Theorem 1] and [GL02, first sentence of Remark 3.5]. Both are important for us, but the latter seems to appear without proof in [GL02], so we sketch some.

Theorem 5.2.8 (Katz, Skorobogatov, et al.)**.** *Fix $\ell \nmid |k|$. Fix integers $n, N \geq 0$, and a complete intersection $X \subseteq \mathbb{P}^n_k$ with $\dim X = N$ and $\operatorname{codim} X \geq 1$. Let $D := \dim(X_{\mathrm{sing}})$. If $i \in \mathbb{Z}$, then (1) if $i \geq N + D + 2$, then $\mathcal{E}^i_\triangle(X, \mathbb{P}) = \emptyset$; and (2) if $i = N + D + 1$, then $\mathcal{E}^i(\mathbb{P}^N_k) \subseteq \mathcal{E}^i(X)$.*

Remark 5.2.9. By [GL02, Proposition 3.2], one could add the following to Theorem 5.2.8: (3) if $i = N$, then $\mathcal{E}^i(\mathbb{P}^N_k) \subseteq \mathcal{E}^i(X)$; and (4) if $0 \leq i \leq N - 1$, then $\mathcal{E}^i_\triangle(X, \mathbb{P}) = \emptyset$. But we only need (1)–(2).

Proof sketch for Theorem 5.2.8. When X is a hypersurface in a smooth projective complete intersection Y/k of dimension ≥ 2, Theorem 5.2.8 follows from [Sko92, Corollary 2.2, up to Veronese embedding], Theorem 5.1.8, "Betti comparison" for Y, and the geometric irreducibility of Y. In general, one can prove Theorem 5.2.8 either inductively or directly.

Inductive proof. If $D \geq N - 1$, use [Poo17, Corollary 7.5.21]. Now suppose $D \leq N - 2$ (which implies $N \geq 1$, since $D \geq -1$). Using $N - D \geq 2$, induct on $\operatorname{codim} X$ using [Sko92, Corollary 2.2], [GL02, Lemma 1.1(ii)], and [GL02, proofs of Theorem 2.4 and Proposition 2.5, up to Veronese embeddings].

Direct proof. Suppose $N \geq 1$ and $D \leq N - 1$. Claim (1) follows from Katz [Hoo91, loc. cit.]. And if $D = -1$, then (2) follows from weak Lefschetz. Now assume $D \geq 0$, and let $i := N + D + 1$. If $2 \nmid i$, then $\mathcal{E}^i(\mathbb{P}^N_k) = \emptyset$, so (2) holds trivially. Now suppose $2 \mid i$. Let $q := |k|$. Then $\mathcal{E}^i(\mathbb{P}^N_k) = \{q^{i/2}\}$, since $i \leq 2N$. It remains to show that $q^{i/2} \in \mathcal{E}^i(X)$. Let $T := \mathbb{A}^1_k$. Following Katz, we can reduce to the case in which there exists a closed subscheme

$Z \subseteq \mathbb{P}^n_T$, flat over T, such that (i) $Z_0 = X$ and (ii) $Y := Z_1$ is a smooth complete intersection in $\mathbb{P}^n_k$ with $\dim Y = N$. In this case, [Gro72, Deligne's Exposé I, Corollaire 4.3] implies that the specialization map $H^i(Z_t) \to H^i(Z \times_T \overline{k(T)}, \mathbb{Q}_\ell)$ is an isomorphism at $t = 1$ (since $D \geq 0$), and a surjection at $t = 0$. So by G_k-equivariance, $\mathcal{E}^i(Y) \subseteq \mathcal{E}^i(X)$. But $i \geq N + 1$ (since $D \geq 0$), so $\mathcal{E}^i(Y) = \mathcal{E}^i(\mathbb{P}^N_k)$ (by (1) for Y). Thus $\{q^{i/2}\} = \mathcal{E}^i(\mathbb{P}^N_k) \subseteq \mathcal{E}^i(X)$. ☐

The following standard result is also essential in the proof of Proposition 5.1.9:

Lemma 5.2.10 (Real amplification)**.** *Fix an integer $l \geq 0$ and a tuple $\boldsymbol{\beta} \in \{z \in \mathbb{C} : |z| \geq 1\}^l$. Then $\limsup_{n\to\infty} \Re(\beta_1^{dn} + \cdots + \beta_l^{dn}) \geq l$ holds for every integer $d \geq 1$.*

Proof. We may assume $l \geq 1$ and $d = 1$. Then, Dirichlet's approximation theorem implies that $\limsup_{n\to\infty} \Re(\cdots)$ is $+\infty$ if $\|\boldsymbol{\beta}\|_\infty > 1$, and l otherwise. ☐

Proof of Proposition 5.1.18. The first part is clear: $X_{\mathrm{sing}}(\overline{k})$ is G_k-invariant, so the geometric point $x\colon \operatorname{Spec} \overline{k} \to X$ must factor through $\operatorname{Spec} k$. It remains to prove the second part. Suppose $k = \mathbb{F}_q$, and let $k_r := \mathbb{F}_{q^r}$ and $N_r(Y) := \#Y(k_r)$ for any given integer $r \geq 1$ and k-scheme Y. Clearly X is reduced, so by [DLR17, §2.3, eq. (8)], the Galkin–Shinder formulas imply (for all $r \geq 1$) the equality

$$N_r(F(X)) = \frac{N_r(X)^2 - 2(1 + |k_r|^{\dim X}) N_r(X) + N_{2r}(X)}{2|k_r|^2} + |k_r|^{-2+\dim X} N_r(X_{\mathrm{sing}}).$$

Now recall our assumption on $X_{\mathrm{sing}}(\overline{k})$. By an inspection of the three cases of [DLR17, §4.4, Proposition 4.8], it follows that there exists a smooth genus 4 curve $C/\mathbb{F}_q$ such that $N_r(F(X)) = \frac{1}{2}\left(N_r(C)^2 + O(N_r(C)) + N_{2r}(C)\right)$. Thus $N_r(F(X)) = |k_r|^2 + O(|k_r|^{3/2})$. Since $\dim X = 3$ and $N_r(X_{\mathrm{sing}}) = O(1)$, it follows that

$$N_r(X)^2 - 2(1 + |k_r|^3) N_r(X) + N_{2r}(X) = 2|k_r|^4 + O(|k_r|^{7/2}).$$

For all $d \geq 1$, let $E_d(X) := N_d(X) - |\mathbb{P}^3(k_d)|$. Then $N_r(X) - |k_r|^3 = E_r(X) + |k_r|^2 + O(|k_r|)$ and $N_{2r}(X) = E_{2r}(X) + |k_r|^6 + |k_r|^4 + O(|k_r|^2)$, so we easily obtain

$$(E_r(X) + |k_r|^2)^2 + O(E_r(X) \cdot |k_r|) + O(N_r(X)) + E_{2r}(X) = |k_r|^4 + O(|k_r|^{7/2}).$$

Now fix $\ell \neq p$, fix an embedding $\iota\colon \mathbb{Q}_\ell \hookrightarrow \mathbb{C}$, and let $\alpha_1, \ldots, \alpha_b \in \mathbb{C}$ denote the weight-4 eigenvalues on $\mathbb{C} \otimes_{\mathbb{Q}_\ell} H^4(X)$. For each $i \in [b]$, let $\tilde{\alpha}_i := |k|^{-4/2}\alpha_i \in S^1$. Then $E_d(X) = |k_d|^{4/2}(\tilde{\alpha}_1^d + \cdots + \tilde{\alpha}_b^d - 1) + O_X(|k_d|^{3/2})$ uniformly over $d \geq 1$, by Theorem 5.2.8(1). In particular, $E_d(X) \ll_X |k_d|^2$ and $N_d(X) \ll_X |k_d|^3$, so ultimately,

$$((\tilde{\alpha}_1^r + \cdots + \tilde{\alpha}_b^r - 1) + 1)^2 + \left(\tilde{\alpha}_1^{2r} + \cdots + \tilde{\alpha}_b^{2r} - 1\right) = 1 + O_X(|k_r|^{-1/2}).$$

Dirichlet's approximation theorem (applied to the *phases* $\tilde{\alpha}_1, \ldots, \tilde{\alpha}_b$) then yields $b^2 + (b-1) = 1$ (cf. Lemma 5.2.10), whence $b = 1$ (since $b \geq 0$). Theorem 5.2.8(2) then gives $\tilde{\alpha}_1 = 1$, which implies (by the trace formula) that X is $|E|$-good. ☐

5.3 Analyzing cubic threefolds as conic bundles

Fix a finite field $k := \mathbb{F}_q$. Let $X := V(C)$ denote a projective cubic hypersurface in $\mathbb{P}^4_k$. *Assume* $\dim(X_{\text{sing}}) \le 0$. (In particular, C must be absolutely irreducible, by Theorem 5.2.8(1) and [Poo17, Corollary 7.5.21].)

5.3.1 Counting using a conic bundle structure

In addition to assuming $\dim(X_{\text{sing}}) \le 0$, we now make the following assumptions.

(1) We assume $p \ge 3$ whenever necessary or convenient.

(2) We *assume* that $X_{\overline{k}}$ does *not* contain any 2-plane P.

Under the above assumptions, we will proceed to count points roughly along the lines of [DLR17, §4.3] and [Laf16] (but note that we have not assumed smoothness).

First, it is known that every projective cubic threefold over $\overline{k}$ contains a line (since every projective cubic surface does). By enlarging k (and applying a $\mathrm{GL}_5(k)$-transformation) if necessary, we may assume that X contains $L := V(x_1, x_2, x_3)$.

Next, given a 5-vector $\boldsymbol{x}$, let $\pi(\boldsymbol{x})$ denote the 3-vector (x_1, x_2, x_3). Then we may write

$$C(\boldsymbol{x}) = f + 2q_1x_4 + 2q_2x_5 + l_1x_4^2 + 2l_2x_4x_5 + l_3x_5^2,$$

for some forms f, q_i, l_j in $\pi(\boldsymbol{x})$, with f cubic, q_i quadratic, and l_j linear.

By assumption, if $\boldsymbol{x}' = (x_1, x_2, x_3) \in k^3 \setminus \{\boldsymbol{0}\}$, the plane $P_{[\boldsymbol{x}']}$ through $L, [\boldsymbol{x}']$ (i.e. through $[e_4], [e_5], [\boldsymbol{x}']$) is *not* contained in X. In other words, $(f, q_i, l_j)(\boldsymbol{x}') \neq \boldsymbol{0}$. So let $Q_{\boldsymbol{x}'}$ denote the projective conic

$$V(fy_1^2 + 2q_1y_1y_2 + 2q_2y_1y_3 + l_1y_2^2 + 2l_2y_2y_3 + l_3y_3^2) \subseteq \{[\boldsymbol{y}] \in \mathbb{P}^2\},$$

canonically embedded into X via $[\boldsymbol{y}] \mapsto [y_1\boldsymbol{x}' + y_2e_4 + y_3e_5]$, with image $Q_{[\boldsymbol{x}']} \subseteq X$, say.

Note that the (restricted) projection $[\pi]\colon X \setminus L \to \mathbb{P}^2$ fibers $X \setminus L$ into affine conics $Q_{[\boldsymbol{x}']} \setminus L$; equivalently, the plane $P_{[\boldsymbol{x}']}$ intersects X in the product of L and $Q_{[\boldsymbol{x}']}$.

Remark 5.3.1. We are essentially writing $\pi(\boldsymbol{x}) = y_1\boldsymbol{x}'$, and "factoring out" y_1 to get $Q_{\boldsymbol{x}'}$.

Definition 5.3.2. Let $M := \begin{bmatrix} f & q_1 & q_2 \\ q_1 & l_1 & l_2 \\ q_2 & l_2 & l_3 \end{bmatrix}$ and $(\delta_1, \delta_2, \delta_3) := (l_1l_3 - l_2^2, fl_3 - q_2^2, fl_1 - q_1^2)$. Let $\Gamma := V(\det M) \subseteq \{[\boldsymbol{x}'] \in \mathbb{P}^2\}$, where

$$\det M = f\delta_1 - q_1(q_1l_3 - q_2l_2) + q_2(q_1l_2 - q_2l_1) = f\delta_1 - q_1^2l_3 - q_2^2l_1 + 2q_1q_2l_2 \in k[\boldsymbol{x}'].$$

Proposition 5.3.3. *Here*

$$\begin{aligned}\#X(k) - \#\mathbb{P}^3(k) = &- q^2\#\{[\boldsymbol{x}] \in L(k) : l_1x_4^2 + 2l_2x_4x_5 + l_3x_5^2 = 0\} \\ &+ q\sum_{[\boldsymbol{x}'] \in \Gamma(k)} (-1 + \#\{\text{distinct } k\text{-lines on } Q_{[\boldsymbol{x}']}\}).\end{aligned}$$

Remark 5.3.4. As $[\boldsymbol{x}']$ varies over Γ, the lines on $Q_{[\boldsymbol{x}']}$ presumably cut out a closed subscheme Γ' of the Fano scheme of lines $F(X)$, which in turn sits in the Grassmannian $\operatorname{Gr}(2,5) \subseteq \mathbb{P}(\bigwedge^2 k^5)$. However, it is unclear (both theoretically and computationally) how suitable or well-developed the theory of Γ' is for our purposes.

Proof sketch. Here each $Q_{[\boldsymbol{x}']}$ is in fact a conic (due to the plane-free assumption on X). So we may follow the casework of [DLR17, pp. 8–9, proof of Proposition 4.6]. Specifically, write $\#X(k) - \#L(k)$ as

$$\sum_{[\boldsymbol{x}'] \in \mathbb{P}^2(k)} \#Q_{[\boldsymbol{x}']}(k) - \sum_{[\boldsymbol{x}'] \in \mathbb{P}^2(k)} \#\{[\boldsymbol{x}] \in L(k) : l_1 x_4^2 + 2l_2 x_4 x_5 + l_3 x_5^2 = 0\}.$$

The first sum Σ_1 simplifies (upon considering smooth and singular Q's separately) to

$$(q+1)(q^2+q+1) + q \sum_{[\boldsymbol{x}'] \in \Gamma(k)} (-1 + \#\{\text{distinct } k\text{-lines on } Q_{[\boldsymbol{x}']}\}),$$

while the second sum Σ_2 simplifies (upon "switching the order of $\boldsymbol{x}', \boldsymbol{x}$") to

$$(q+1)^2 + q^2 \#\{[\boldsymbol{x}] \in L(k) : l_1 x_4^2 + 2l_2 x_4 x_5 + l_3 x_5^2 = 0 \ \forall \boldsymbol{x}'\}.$$

The result follows upon noting $\#L(k) + (q+1)(q^2+q+1) - (q+1)^2 = \#\mathbb{P}^3(k)$. □

Remark 5.3.5. The proof above works for an arbitrary projective cubic threefold X/k containing L such that there exists no 2-plane $P/\overline{k}$ with $L_{\overline{k}} \subseteq P \subseteq X_{\overline{k}}$. (No hypothesis on isolated singularities is needed.) However, the formula could easily (but somewhat tediously) be adjusted to allow for such 2-planes.

Proposition 5.3.6 (Cf. [DLR17, p. 9, Proposition 4.7])**.** *Fix $[\boldsymbol{x}'] \in \Gamma(k)$. Then $Q_{[\boldsymbol{x}']}$ is*

(1) reducible over k, *i.e. the product of two* distinct k*-lines, if and only if $(\delta_1, \delta_2, \delta_3) \neq \mathbf{0}$ and $-\delta_i \in (k^\times)^2 \cup \{0\}$ for all $i \in [3]$;*

(2) non-reduced *over k, i.e. a* double k*-line, if and only if $(\delta_1, \delta_2, \delta_3) = \mathbf{0}$; and*

(3) integral over k, *i.e. the product of two* distinct conjugate *lines defined over $\overline{k}$ but not over k, otherwise.*

Proof. By assumption, $Q_{[\boldsymbol{x}']}$ is in fact a conic. Here $Q_{[\boldsymbol{x}']} \cong_k Q_{\boldsymbol{x}'} := V(\boldsymbol{y}^T M \boldsymbol{y})$, where

$$\boldsymbol{y}^T M \boldsymbol{y} = f y_1^2 + 2q_1 y_1 y_2 + 2q_2 y_1 y_3 + l_1 y_2^2 + 2l_2 y_2 y_3 + l_3 y_3^2.$$

By assumption, $\det M = 0$, from which a routine computation yields $\delta_2 \delta_3 = (q_1 q_2 - f l_2)^2 \in k^2$. Similarly, $\delta_1 \delta_2 \in k^2$ and $\delta_1 \delta_3 \in k^2$. In particular, if $-\delta_i \in (k^\times)^2$ for *some* $i \in [3]$, then $-\delta_j \in k^2$ for *all* $j \in [3]$.

Now fix $i \in [3]$. Essentially by [DLR17, first paragraph of proof of Proposition 4.7] (i.e. casework on $|(Q_{\boldsymbol{x}'} \cap V(y_i))(k)|$), we know that *if* $\delta_i \neq 0$, then $V(y_i) \not\subseteq Q_{\boldsymbol{x}'}$, and

- $Q_{\boldsymbol{x}'}$ is k-reducible if $-\delta_i \in (k^\times)^2$, while

- $Q_{\boldsymbol{x}'}$ is k-reduced *and* k-irreducible (i.e. k-integral) otherwise.

Consequently, *if* $\boldsymbol{\delta} \neq \mathbf{0}$, then we have established the desired dichotomy between (1) and (3). Finally, *suppose* $\boldsymbol{\delta} = \mathbf{0}$. Then it remains precisely to show that $Q_{\boldsymbol{x}'}$ is a double k-line, or equivalently (by Galois theory) that $(Q_{\boldsymbol{x}'})_{\overline{k}}$ is a double $\overline{k}$-line. To this end, write $(Q_{\boldsymbol{x}'})_{\overline{k}}$ as the product of two $\overline{k}$-lines L_1, L_2, and note that for each $i \in [3]$, we either have $V(y_i)_{\overline{k}} \in \{L_1, L_2\}$, or else $V(y_i)_{\overline{k}} \cap L_1 = \{p_i\} = V(y_i)_{\overline{k}} \cap L_2$ for some point $p_i \in \mathbb{P}^2(k)$. In any case, we can find points $q_i \in \mathbb{P}^2(k)$ with $q_i \in V(y_i)_{\overline{k}} \cap L_1 \cap L_2$. Here we must have $\{q_1\} \cap \{q_2\} \cap \{q_3\} = \emptyset$, since $V(y_1, y_2, y_3)_{\overline{k}} = \emptyset$. But also, $\{q_1, q_2, q_3\} \subseteq L_1 \cap L_2$. Thus $L_1 = L_2$, as desired. □

To go further, we must analyze X's singularities (or lack thereof).

Observation 5.3.7. C is singular at a given (geometric) point $\boldsymbol{x} = [x_4 e_4 + x_5 e_5] \in L(\overline{k})$ if and only if $l_1 x_4^2 + 2l_2 x_4 x_5 + l_3 x_5^2$ vanishes identically as a linear form in $\boldsymbol{x}'$.

Proof. By assumption, "$C|_L$" vanishes identically, so $C, \partial_{x_4} C, \partial_{x_5} C$ vanish on the (affine) cone of L. Next, if we write $C(\boldsymbol{x}) = O((\boldsymbol{x}')^2) + l_1 x_4^2 + 2l_2 x_4 x_5 + l_3 x_5^2$, then $\partial_{x_i} C$, for $i \in [3]$, simplifies to $(\partial_{x_i} l_1) x_4^2 + 2(\partial_{x_i} l_2) x_4 x_5 + (\partial_{x_i} l_3) x_5^2$ (a binary quadratic form in x_4, x_5) on the cone of L. But l_1, l_2, l_3 are linear, so the desired result immediately follows. □

Remark 5.3.8. In particular, $\#(X_{\mathrm{sing}} \cap L)(\overline{k}) \leq 2$, since X has isolated singularities. There is a conceptual proof of this fact: "$\nabla x_i C|_L$" is a quadratic polynomial on L for each i, so $L \not\subseteq X_{\mathrm{sing}}$ implies that there exists i such that "$\nabla x_i C|_L$" has at most 2 distinct roots.

Example 5.3.9. The data in [Wan21f, `Data for... X (in general).xlsx`] seems consistent with the propositions (and observation) above.

5.3.2 A convenient choice of a line

By enlarging k (and applying a $\mathrm{GL}_5(k)$-transformation) if necessary, we may assume that

$$|X_{\mathrm{sing}}(\overline{k}) \cap \{[e_4], [e_5]\}| = \min(|X_{\mathrm{sing}}(\overline{k})|, 2),$$

and that X contains the line $L := V(x_1, x_2, x_3)$ through $[e_4], [e_5]$. To justify the existence of such a "convenient line" L requires some proof, which we now give (by casework).

Proof when $|X_{\mathrm{sing}}(\overline{k})| \geq 2$. Say $X_{\mathrm{sing}}(\overline{k}) \supseteq \{[e_4], [e_5]\}$. Then the binary cubic form "$C|_L$" vanishes identically, because C itself is singular at $[e_4], [e_5] \in L(\overline{k})$ (forcing "$C|_L$" to vanish to order ≥ 4). □

Proof when $|X_{\mathrm{sing}}(\overline{k})| = 1$. Say $X_{\mathrm{sing}}(\overline{k}) = \{[e_4]\}$. We must show that $[e_4]$ is contained in a line on $X_{\overline{k}}$. To this end, consider an arbitrary (nontrivial) hyperplane section S of $X_{\overline{k}}$. Then S is a projective cubic surface, so it must contain a line l. If $[e_4] \in l$, then we are done, so suppose not. Now let P denote the (unique) plane through $[e_4], l$. If $P \subseteq S$, then we may simply choose any line on P through $[e_4]$. So suppose $P \not\subseteq S$. Then $P \cap S$ is a plane cubic containing $[e_4], l$. But, crucially, the ternary cubic form "$C|_{P \cap S}$" must be singular at $[e_4]$ (since C itself is singular at $[e_4]$). Since $[e_4] \notin l$, we conclude that $[e_4]$ is a singular point of a *conic* Q, which can only occur if Q is a product of two lines passing through $[e_4]$. Either line then suffices. □

Proof when $X_{\rm sing}(\overline{k}) = \emptyset$. We must show that $X_{\overline{k}}$ contains a line. It suffices to choose any (nontrivial) hyperplane section $S \subseteq X_{\overline{k}}$, and then any line $l \subseteq S$. (By Bertini's theorem, we can choose S to be smooth if desired, but this is unnecessary.) $\square$

Henceforth, we fix a "convenient line" L as above, though in principle other lines could also be used.

5.3.3 The smooth case

See [DLR17, par. 4 of §4.3]. (Here $\Gamma' \to \Gamma$ is a fairly "nice" double cover, already analyzed by [BSD67].)

5.3.4 The case of exactly one singular point contained in L

By assumption, either $l_1 = 0$ or $l_3 = 0$. Assume the former ($l_1 = 0$). Then either $l_2 = 0$ and $l_3 \neq 0$, or else l_2, l_3 are linearly independent. (Otherwise, $X_{\rm sing}$ would contain a point in $L \setminus \{[e_4]\}$.) In particular,

$$\#\{[\boldsymbol{x}] \in L(k) : l_1 x_4^2 + 2l_2 x_4 x_5 + l_3 x_5^2 = 0\} = \#\{[\boldsymbol{x}] \in L(k) : x_5 = 0\} = 1.$$

Here $(\delta_1, \delta_2, \delta_3) = (-l_2^2, f l_3 - q_2^2, -q_1^2)$ and

$$\det M = f\delta_1 - q_1^2 l_3 + 2q_1 q_2 l_2 = -f l_2^2 + 2q_2 l_2 q_1 - l_3 q_1^2.$$

Here $C = f + 2q_1 x_4 + 2q_2 x_5 + 2l_2 x_4 x_5 + l_3 x_5^2$.

Proposition 5.3.10. *Here* $\det M$ *is a* nonzero *ternary quintic form in* $k[\boldsymbol{x}']$.

Proof when $l_2 \neq 0$. Noting the linearity of C in x_4, we write

$$C = f + 2q_2 x_5 + l_3 x_5^2 + 2(q_1 + l_2 x_5) x_4.$$

By assumption, C is (absolutely) irreducible, so $\gcd(f + 2q_2 x_5 + l_3 x_5^2, q_1 + l_2 x_5) = 1$ in $\overline{k}[\boldsymbol{x}', x_5]$. Consequently, Gauss' lemma implies that $\gcd(f + 2q_2 x_5 + l_3 x_5^2, q_1 + l_2 x_5) = 1$ in $\overline{k}(\boldsymbol{x}')[x_5]$. But in $\overline{k}(\boldsymbol{x}')[x_5]$, the remainder of $f + 2q_2 x_5 + l_3 x_5^2$ modulo $l_2 x_5 + q_1$ is precisely $-l_2^{-2} \det M$, so $\det M \neq 0$, as desired. $\square$

Proof when $l_2 = 0$. Here $l_3 \neq 0$, and $\det M = -l_3 q_1^2$. Also, $C = f + 2q_1 x_4 + 2q_2 x_5 + l_3 x_5^2$. Suppose for contradiction that $q_1 = 0$. Then X is the (projective) cone over a projective cubic surface S. Now fix a line l on $S_{\overline{k}} = X_{\overline{k}} \cap \{x_4 = 0\}$. Then l extends to a plane on $X_{\overline{k}}$, contradicting our plane-free assumption on X. Thus in fact $q_1 \neq 0$, so $\det M \neq 0$. $\square$

Corollary 5.3.11. *Here*

$$\#X(k) - \#\mathbb{P}^3(k) \leq -q^2 + (q^2 + O(q^{3/2}))\#\{\text{irreducible components of } \Gamma_{\overline{k}}\}.$$

Proof sketch. Similar to the analogous corollary when $|X_{\rm sing}(\overline{k})| \geq 2$ (treated below). $\square$

Proposition 5.3.12. *If $[\boldsymbol{x}'] \in \Gamma(k)$, then the number of distinct k-lines on $Q_{[\boldsymbol{x}']}$ is*

(1) exactly $1 + (\frac{-\delta_2(\boldsymbol{x}')}{k})$, defined using the Legendre symbol over k, if $l_2(\boldsymbol{x}') = q_1(\boldsymbol{x}') = 0$; and

(2) exactly 2, otherwise.

Proof. Apply Proposition 5.3.6. $\square$

5.3.5 The case of two or more isolated singularities in L

By assumption, $l_1 = l_3 = 0$. Hence $l_2 \neq 0$ (or else X_{sing} would contain L entirely). In particular,

$$\#\{[\boldsymbol{x}] \in L(k) : l_1 x_4^2 + 2l_2 x_4 x_5 + l_3 x_5^2 = 0\} = \#\{[\boldsymbol{x}] \in L(k) : x_4 x_5 = 0\} = 2.$$

Here $(\delta_1, \delta_2, \delta_3) = (-l_2^2, -q_2^2, -q_1^2)$ and $\det M = f\delta_1 + 2q_1q_2l_2 = (-fl_2 + 2q_1q_2)l_2$.

Proposition 5.3.13. *Here* $\det M$ *is a* nonzero *ternary quintic form in* $k[\boldsymbol{x}']$.

Proof. Here $C = f + 2q_1x_4 + 2q_2x_5 + 2l_2x_4x_5$. But by assumption, C is (absolutely) irreducible, so $f \cdot 2l_2 \neq 2q_1 \cdot 2q_2$. (Otherwise, there would exist forms a, b, c, d with $(f, 2l_2) = (ab, cd)$ and $(2q_1, 2q_2) = (ac, bd)$, and we would have $C = (a + dx_5)(b + cx_4)$, with both factors necessarily nonzero of degree ≥ 1; contradiction.)

Earlier, we noted that $l_2 \neq 0$. So in fact, $\det M = (-fl_2 + 2q_1q_2)l_2 \neq 0$. □

Corollary 5.3.14. *Here*

$$\#X(k) - \#\mathbb{P}^3(k) \leq -2q^2 + (q^2 + O(q^{3/2}))\#\{\text{irreducible components of } \Gamma_{\overline{k}}\}.$$

Proof sketch. Apply the trivial bound $-1 + \#\{\text{distinct } k\text{-lines on } Q_{[\boldsymbol{x}']}\} \leq 1$. Then note that

$$|\Gamma(k)| = O(1) + (q + O(q^{1/2}))\#\{\text{irreducible components of } \Gamma_{\overline{k}} \text{ defined over } k\}$$

by the Lang–Weil bound, for instance. (Note that if $\mathcal{C}/\overline{k}$ is an embedded projective curve *not* defined over k, then there exists a conjugate $\mathcal{C}'/\overline{k}$ distinct from $\mathcal{C}$, so that "$\mathcal{C}(k)$" is contained in the *finite* set $(\mathcal{C} \cap \mathcal{C}')(\overline{k})$ of size $O(1)$.) □

Proposition 5.3.15. *If* $[\boldsymbol{x}'] \in \Gamma(k)$, *then the number of distinct* k*-lines on* $Q_{[\boldsymbol{x}']}$ *is exactly* 1 *if* $(l_2, q_1, q_2)(\boldsymbol{x}') = \mathbf{0}$, *and exactly* 2 *otherwise.*

Proof. Apply Proposition 5.3.6. □

Remark 5.3.16 (Original approach to one direction of Theorem 5.1.26). Suppose $\Gamma_{\overline{k}}$ has ≥ 3 irreducible components. Then $fl_2 - 2q_1q_2$ has ≥ 2 prime factors (counted without multiplicity) distinct from (l_2).

Case 1: $A \mid fl_2 - 2q_1q_2$ *for some linear form* $A \in \overline{k}[x_1, x_2, x_3]$ *with* $(A) \neq (l_2)$. Then the reduction of $C = f + 2q_1x_4 + 2q_2x_5 + 2l_2x_4x_5$ modulo A is reducible, and therefore has a linear factor B mod A such that C mod (A, B) vanishes identically. So $X_{\overline{k}} = V_{\mathbb{P}^4}(C)_{\overline{k}}$ contains the 2-plane $A = B = 0$ in in $\mathbb{P}^4_{\overline{k}}$. Thus the condition 5.1.26(3) holds; so does condition 5.1.26(4) (cf. [Wan22, Proposition 5.9(3)⇒(1)]).

Case 2: $fl_2 - 2q_1q_2 = AB$ *for some irreducible quadratic forms* $A, B \in \overline{k}[x_1, x_2, x_3]$. Then $C \cdot l_2 = AB + 2(q_1 + l_2x_5)(q_2 + l_2x_4)$. Let U denote the open subscheme $\{l_2 \neq 0\} \subseteq \mathbb{P}^4_{\overline{k}}$. Then $X_{\overline{k}}$ contains the scheme-theoretic intersection $V_U(A, q_1 + l_2x_5)$, where $A, q_1 + l_2x_5$ are quadratic forms "missing" the variable x_4 (and therefore sharing a singularity in $\mathbb{P}^4_{\overline{k}}$). Furthermore, $l_2 \nmid A$ (since A is irreducible), so there exists $(x_1, x_2, x_3, x_5) \in \overline{k}^4 \setminus \{\mathbf{0}\}$ with $A = 0 \neq l_2$ and $x_5 = -q_1/l_2$; thus $V_U(A, q_1 + l_2x_5)$ is nonempty. So the condition 5.1.26(4) holds.

Chapter 6

Isolating special integral solutions

6.1 Introduction

As in §4.1, let $m := 6$ and $F := x_1^3 + \cdots + x_6^3$. Fix $w \in C_c^\infty(\mathbb{R}^m)$ with (F, w) smooth in the sense of Definition 1.4.3. As we mentioned in §4.1, the analysis of the delta method for $N_{F,w}(X)$ (see (3.2)–(3.3)) naturally breaks up into two parts: one over the locus $F^\vee(\boldsymbol{c}) \neq 0$ (where "reciprocal" Hasse–Weil L-functions $1/L(s, V_{\boldsymbol{c}})$ naturally arise), and one over the locus $F^\vee(\boldsymbol{c}) = 0$ (where the $\mathbb{Q}$-varieties $V_{\boldsymbol{c}}$ are singular, and thus need to be treated separately).

Since $m \geq 5$, it is known that the $\boldsymbol{c} = \boldsymbol{0}$ terms isolate the singular series in HLH for (F, w) (see Definition 1.4.6 and §3.4). In [Wan21d], we interpret the sum

$$Y^{-2} \sum_{n \geq 1} \sum_{\boldsymbol{c} \in \mathbb{Z}^m} n^{-m} S_{\boldsymbol{c}}(n) I_{\boldsymbol{c}}(n) \cdot \mathbf{1}_{\boldsymbol{c} \neq \boldsymbol{0}} \cdot \mathbf{1}_{V_{\boldsymbol{c}} \text{ is singular}}$$

in terms of special $\mathbb{Q}$-subvarieties (specifically, linear spaces) on V, which allows one to cleanly reformulate HLH for (F, w) in a useful way (for Chapter 8). The main goal of the present chapter is to summarize the work done in [Wan21d]. Throughout Chapter 6, we let $\sum'_{\boldsymbol{c} \in *} \cdots$ denote $\sum_{\boldsymbol{c} \in * \text{ with } F^\vee(\boldsymbol{c}) = 0} \cdots$, unless specified otherwise.

Theorem 6.1.1. *For some absolute constant $\delta > 0$, we have (uniformly over $X \geq 1$)*

$$Y^{-2} \sideset{}{'}\sum_{\boldsymbol{c} \in \mathbb{Z}^m} \mathbf{1}_{\boldsymbol{c} \neq \boldsymbol{0}} \cdot \sum_{n \geq 1} n^{-m} S_{\boldsymbol{c}}(n) I_{\boldsymbol{c}}(n) = O_{F,w}(X^{m/2 - \delta}) + \sum_{L \in C(\mathrm{SSV})} \sum_{\boldsymbol{x} \in L \cap \mathbb{Z}^m} w(\boldsymbol{x}/X).$$

Remark 6.1.2. The set $C(\mathrm{SSV})$ is known to be finite for general reasons (recalled in §6.3 below). For diagonal F as in Theorem 6.1.1, one can in fact compute the set $C(\mathrm{SSV})$ explicitly: see Observation 6.3.9 (essentially classical). In any case, on the Diophantine (right-hand) side of Theorem 6.1.1, the *sum* over L is roughly equivalent to a *union*, since $\#C(\mathrm{SSV}) < \infty$ and the pairwise intersections fit in the error term.

Remark 6.1.3. The *h-invariant* introduced by [DL62, DL64] provides an equivalent way to think about linear subvarieties of cubic varieties. (See e.g. [Die17, Lemma 1.1] for a modern record of a more general equivalence.) In fact, the proof of the important Lemma 6.4.1 below (see [Wan21d, §6]) essentially relies on (a convenient choice of) "h-decompositions of F" corresponding to the L's in Theorem 6.1.1.

Here is the main result here (in the present chapter) needed for Chapter 8:

Corollary 6.1.4. *For some absolute $\delta > 0$, we have (uniformly over $X \geq 1$)*

$$Y^{-2} \sideset{}{'}\sum_{\boldsymbol{c} \in \mathbb{Z}^m} \sum_{n \geq 1} n^{-m} S_{\boldsymbol{c}}(n) I_{\boldsymbol{c}}(n) = \sigma_{\infty,F,w} \mathfrak{S}_F X^{m-3} + O_{F,w}(X^{m/2-\delta}) + \sum_{L \in C(\mathrm{SSV})} \sigma_{\infty,L^\perp,w} X^{m/2}.$$

Proof. Combine Theorem 6.1.1 with the routine $\boldsymbol{c} = \boldsymbol{0}$ computation in §3.4 (isolating the singular series). In the exponents, note that $m/2 = m - 3 = 3$. □

Remark 6.1.5. Theorem 6.1.1 and Corollary 6.1.4 are *completely unconditional* results. These results let us *reformulate* HLH for (F, w) as a statement purely about *cancellation* in the sum $\sum_{\boldsymbol{c} \in \mathbb{Z}^m} \mathbf{1}_{F^\vee(\boldsymbol{c}) \neq 0} \cdot \sum_{n \geq 1} n^{-m} S_{\boldsymbol{c}}(n) I_{\boldsymbol{c}}(n)$; see Chapter 8. With any luck, a similar reformulation might be possible much more generally. But at least for cubic forms in 4 variables, subtleties in the Manin–Peyre constant would likely demand a more sophisticated "delta method recipe" beyond "restriction to $F^\vee = 0$".

Let us sketch the proof of the theorem. (In this sketch, we restrict L to $C(\mathrm{SSV})$.)

We *start* generally, observing that $F^\vee|_{L^\perp} = 0$ for all L (see the first part of Proposition 6.3.2). Conversely, at least for diagonal F, most $\boldsymbol{c}$'s on the left-hand side of Theorem 6.1.1 are in fact *trivial*, in the sense of the following definition:

Definition 6.1.6. Call a solution $\boldsymbol{c} \in \mathbb{Z}^m$ to $F^\vee(\boldsymbol{c}) = 0$ *trivial* if $\boldsymbol{c} \in L^\perp$ for some L.

Actually, we cannot analyze all trivial $\boldsymbol{c}$'s uniformly, but only the "least degenerate" ones. Under (plausibly mild) hypotheses satisfied by diagonal F, the second part of Proposition 6.3.2 establishes a vanishing baseline for the jets $j^\bullet F^\vee$ over $\bigcup_L L^\perp$—which inspires the following definition:

Definition 6.1.7. Call $F^\vee$ *unsurprising* if uniformly over $C \geq 1$, the equation $F^\vee(\boldsymbol{c}) = 0$ has at most $O_\epsilon(C^{m/2-1+\epsilon})$ integral solutions $\boldsymbol{c} \in [-C, C]^m$ that are either nontrivial, or else trivial with $j^{2^{m/2-1}} F^\vee(\boldsymbol{c}) = \boldsymbol{0}$.

Indeed, in [Wan21d, §5], we prove (using the diagonality of F) that $F^\vee$ is unsurprising—in which case "almost all solutions to $F^\vee = 0$ are trivial with nonzero $2^{m/2-1}$-jet" (qualitatively speaking).

Remark 6.1.8. In particular (for the Gauss map $\gamma \colon V \to V^\vee$ introduced in §6.2), $\gamma(\mathbb{Q}) \colon V(\mathbb{Q}) \to V^\vee(\mathbb{Q})$ can be far from surjective, even though $\gamma \colon V \to V^\vee$ itself is birational and roughly log-height–preserving (up to a constant factor).

Remark 6.1.9. A weaker bound of the form $O(C^{m/2-\delta})$ in Definition 6.1.7 would likely suffice for our main purposes.

For the "least degenerate" trivial $\boldsymbol{c}$'s, Lemma 6.4.1 isolates an explicit positive bias

$$\widetilde{S}_{\boldsymbol{c}}(p^l) = [A_{p^l}(\boldsymbol{c}) + O(p^{-l/2})] \cdot (1 - p^{-1}) \cdot p^{l/2}$$

for most primes p, with $A_{p^l}(\boldsymbol{c}) \ll 1$ essentially a combinatorial factor measuring the p-adic "extent of speciality" of $\boldsymbol{c}$; on average *vertically*, $\mathbb{E}[A_{p^l}(\boldsymbol{c})] \approx 1$. In the dominant ranges (i.e. for n large), the resulting reduction in arithmetic complexity of $S_{\boldsymbol{c}}(n)$ lets us dramatically simplify each sum of the form $\sum_{\boldsymbol{c}\in\Lambda^\perp} S_{\boldsymbol{c}}(n) I_{\boldsymbol{c}}(n)$ by "undoing" Poisson summation to $I_{\boldsymbol{c}}(n)$ over various individual residue classes $\boldsymbol{c} \equiv \boldsymbol{b} \bmod n_0\Lambda^\perp$ with $n_0 \ll X^{1/2}$ dividing n.

Ultimately, this process produces corresponding sums over $\boldsymbol{x} \in \Lambda$ (as desired for Theorem 6.1.1). However, when $m = 6$, we must be careful to separate $\boldsymbol{c} = \boldsymbol{0}$ from $\boldsymbol{c} \neq \boldsymbol{0}$; Lemma 3.4.1 (decay of the singular series over large moduli) provides the necessary input, when contrasted with Lemma 3.3.5 giving decay of $I_{\boldsymbol{c}}(n)$ over small moduli for $\boldsymbol{c} \neq \boldsymbol{0}$.

For a full cross-referenced outline of the proof of Theorem 6.1.1, see [Wan21d, §5].

Remark 6.1.10. We do not need *horizontal* cancellation over n: at least morally, the terms $S_{\boldsymbol{c}}(n), I_{\boldsymbol{c}}(n)$ are positive for trivial $\boldsymbol{c}$'s, while nontrivial $\boldsymbol{c}$'s are relatively sparse. This *morally* explains why we can prove Theorem 6.1.1 unconditionally. In fact, the deepest result we use on L-functions (when $m = 6$) is the (purely local) Weil bound for hyperelliptic curves of genus ≤ 2.

Remark 6.1.11. The full proof of Theorem 6.1.1 requires an *error analysis* to "reduce consideration" to biases. Fortunately, several convenient features make our error analysis here (reducing consideration to biases over $F^\vee(\boldsymbol{c}) = 0$) "half an inch" easier, or clearer, than that in Chapter 8 (reducing to L-functions over $F^\vee(\boldsymbol{c}) \neq 0$). Specifically, small moduli n and bad primes p here cause very little pain compared to those in Chapter 8.

See Remark 6.4.3 for a discussion of what is missing for non-diagonal F.

6.2 Algebraic geometry background

Let $F, V, \ldots$ be as in §3.1. For general F, we need some classical results on the gradient map ∇F, its image, and its ramification. For diagonal F, a more explicit analysis is possible (see §6.3.2 below).

6.2.1 The dual hypersurface and the discriminant form

Since V is smooth, the *polar map* $[\nabla F]\colon \mathbb{P}^{m-1} \to \mathbb{P}^{m-1}$ is *regular*, i.e. defined everywhere. (It would be better to write $\mathbb{P}^{m-1} \to (\mathbb{P}^{m-1})^\vee$, but no confusion should arise.) The map $[\nabla F]$ is finite of degree 2^{m-1} [Dol12, p. 29] between irreducible equidimensional projective varieties, hence surjective. Since $\mathbb{P}^{m-1}$ is smooth, $[\nabla F]$ must then be flat (by "miracle flatness").

Upon restricting $[\nabla F]$ to V, we get the finite surjective *Gauss map* $\gamma\colon V \to V^\vee$, where $V^\vee \subseteq (\mathbb{P}^{m-1}_{\mathbb{Q}})^\vee$ denotes the *dual variety* of V, i.e. the closure of the union of hyperplanes tangent to the *smooth locus* V_{sm} of V. (For us, V is a hypersurface, and $V_{\mathrm{sm}} = V$. So $V^\vee = \overline{[\nabla F]V_{\mathrm{sm}}} = \overline{[\nabla F]V} = [\nabla F]V$. Hence γ is indeed well-defined and surjective; and γ is finite because it is a quasi-finite map between projective varieties.)

Here $V/\mathbb{Q}$ is irreducible over $\mathbb{C}$, so $V^\vee = \gamma(V)$ must be too. Since γ is finite, $V^\vee$ must therefore be a geometrically integral hypersurface, namely the zero locus of $F^\vee$ from Proposition-Definition 3.2.3. (At least for diagonal F, one can explicitly compute $F^\vee$; see (6.1) in §6.3.2 below.) The notation $F^\vee$ is thus convenient for us, but it is likely *not* standard.

It is known that $(V^\vee)^\vee = V$ [Dol12, p. 30, Reflexivity Theorem]. Furthermore, the definition of $V^\vee$ (together with the fact that $V, V^\vee$ are hypersurfaces with V irreducible) implies the divisibility $F(\boldsymbol{x}) \mid F^\vee(\nabla F(\boldsymbol{x}))$, and reflexivity (together with the fact that $V^\vee, V$ are hypersurfaces with $V^\vee$ irreducible) similarly implies the divisibility $F^\vee(\boldsymbol{c}) \mid F(\nabla F^\vee(\boldsymbol{c}))$. These (symmetric!) divisibilities capture much of the basic duality theory for V.

The apparent symmetry between $V, V^\vee$ thus far is deceptive, however. What complicates matters is that $V^\vee$, or equivalently $F^\vee$, must be *singular* if $\deg F \geq 3$. (Otherwise, $(V^\vee)^\vee$ would be a hypersurface of degree larger than $\deg F$, contradicting reflexivity.) Hence the polar map $[\nabla F^\vee]\colon \mathbb{P}^{m-1} \dashrightarrow \mathbb{P}^{m-1}$ is only a *rational* map, defined away from $\operatorname{Sing}(V^\vee)$, the *set of singular points* of $V^\vee$. (Here $\operatorname{Sing}(V^\vee)$ is a proper closed subset of $V^\vee$.)

Nonetheless, given *smooth* points $[\boldsymbol{x}] \in V$ and $[\boldsymbol{c}] \in V^\vee$, the biduality theorem says that $[\nabla F(\boldsymbol{x})] = [\boldsymbol{c}] \iff [\nabla F^\vee(\boldsymbol{c})] = [\boldsymbol{x}]$. (See e.g. [Dol12, p. 30, sentence after Reflexivity Theorem], which however does not explicitly refer to the biduality theorem by name.) For us, $V_{\mathrm{sm}} = V$, so biduality implies that $[\nabla F], [\nabla F^\vee]$ restrict to *inverse morphisms* between $V \setminus [\nabla F]^{-1}(\operatorname{Sing}(V^\vee))$ and $V^\vee \setminus \operatorname{Sing}(V^\vee)$.

Remark 6.2.1. The map $\gamma\colon V \to V^\vee$ is finite surjective (and birational, i.e. of degree 1), but not necessarily *flat* (or equivalently, *locally free*). In fact, the finite $\mathcal{O}_{V^\vee}$-algebra $\gamma_*\mathcal{O}_V$ is isomorphic to $\mathcal{O}_{V^\vee}$ *generically* over $V^\vee$, but *not necessarily everywhere*—for instance, the geometric fiber $V \times_\gamma \overline{k(p)} = \operatorname{Spec}((\gamma_*\mathcal{O}_V)_p \otimes_{\mathcal{O}_{V^\vee,p}} \overline{k(p)})$, and in fact the analogous set-theoretic geometric fiber, may have size ≥ 2 at some point $p \in \operatorname{Sing}(V^\vee)$.

Question 6.2.2. What is known about the fiber of γ over a *singular* point $[\boldsymbol{c}] \in V^\vee$?

Presumably the singular structure of $V_{\boldsymbol{c}}$ should play a role. Perhaps works of Aluffi and Cukierman (such as [AC93, Alu95]) can help to give a precise statement.

6.2.2 Ramification and the Hessian

$[\nabla F]$ is a finite surjective morphism of smooth $\mathbb{Q}$-varieties, so its ramification theory is well-behaved. By [Dol12, p. 29, Proposition 1.2.1],

(1) the *ramification divisor* of $[\nabla F]\colon \mathbb{P}^{m-1} \to \mathbb{P}^{m-1}$ is the Hessian hypersurface $\operatorname{hess}(V) \subseteq \mathbb{P}^{m-1}$; and

(2) its image, the *branch divisor*, is $\operatorname{St}(V)^\vee$, the dual of the Steinerian hypersurface $\operatorname{St}(V)$ (with $\operatorname{St}(V)$ defined scheme-theoretically as in [Dol12, p. 19, §1.1.6]).

In particular, both $\operatorname{hess}(V), \operatorname{St}(V)^\vee$ are (possibly reducible or non-reduced) projective hypersurfaces in $\mathbb{P}^{m-1}$. (It is worth noting that the precise geometry of $\operatorname{hess}(V), \operatorname{St}(V)^\vee$ is not so important in [Wan21d], but it may become more important in the future if one wants to study general V. Most important for us now is that the branch divisor exists, and that it can be computed explicitly when F is diagonal.)

When studying the Hasse principle (or even weak approximation), one can certainly localize before counting points. For instance, the following proposition shows that on an *arbitrary* smooth cubic V, as in [Hoo88] or [Hoo14] for instance, it is always *non-vacuous* to count points on V with nonvanishing Hessian determinant.

Proposition 6.2.3 ([Hoo88, Lemma 1 and its proof]). *If V is a smooth cubic, then $V \not\subseteq \operatorname{hess}(V)$. Furthermore, $V(\mathbb{R})$ is Zariski dense in V,*[1] *so $V(\mathbb{R}) \not\subseteq (\operatorname{hess} V)(\mathbb{R})$.*

Remark 6.2.4 (Cf. [Hoo88, remarks in the paragraph before Lemma 1]). The intersection $V \cap \operatorname{hess}(V)$ consists of *inflection points* if $m = 3$, and of *parabolic points* if $m \geq 4$ [Dol12, p. 17, Theorem 1.1.20]. But it does not seem easy to find a general reference proving the existence of non-inflection or non-parabolic points on V (according as $m = 3$ or $m \geq 4$).

The following stronger technical question comes up in Proposition 6.3.2, although we happen to be able to sidestep it there.

Question 6.2.5. Is it always true (for smooth cubic V) that $V^\vee \not\subseteq \operatorname{St}(V)^\vee$?

Remark 6.2.6. If $V \subseteq \operatorname{hess}(V)$, then applying $[\nabla F]$ would imply $V^\vee \subseteq \operatorname{St}(V)^\vee$. Thus an affirmative answer to the question would give another proof of Proposition 6.2.3.

6.3 Maximal linear subvarieties under duality

Let $F, V, \ldots$ be as in §3.1. If $2 \mid m$ but $m \neq 6$, extend the definition of $C(\mathrm{SSV})$ from Definition 1.4.6 in the obvious way (involving $m/2$-dimensional vector spaces $L/\mathbb{Q}$). The reader only interested in diagonal F can skim forwards to §6.3.2, which explicitly analyzes $C(\mathrm{SSV})$ through the lens of $F^\vee$ (in the diagonal case).

6.3.1 A preliminary general analysis

In general, classical duality theory comes into play, leading to Proposition 6.3.2 below. For the definitions and basic properties of the *polar* and *Gauss* maps $[\nabla F], \gamma$ associated to V, see the first few paragraphs of §6.2.

Consider the affine cone $C(V)$ (i.e. $F(\boldsymbol{x}) = 0$ sitting in $\mathbb{A}^m_{\mathbb{Q}}$). Suppose $m \geq 3$ and $L \subseteq C(V)$ is a vector space over $\mathbb{Q}$ (of arbitrary dimension). Then differentiating along L implies $L \perp \nabla F(\boldsymbol{x})$ for all $\boldsymbol{x} \in L$. In other words, the restriction $\gamma|_{\mathbb{P}L}$ maps into $\mathbb{P}L^\perp$. But $[\nabla F] \colon \mathbb{P}^{m-1} \to \mathbb{P}^{m-1}$ is regular and finite, so $\gamma|_{\mathbb{P}L}$ is itself regular and finite (since $\mathbb{P}L$ is closed in $\mathbb{P}^{m-1}$). Thus $\dim(L) \leq \dim(L^\perp)$, i.e. $\dim(L) \leq \lfloor m/2 \rfloor$.

Remark 6.3.1. Although $\mathbb{P}L^\perp \subseteq (\mathbb{P}^{m-1})^\vee$ is the dual variety of $\mathbb{P}L \subseteq \mathbb{P}^{m-1}$, we prefer to write $L^\perp$ instead of $L^\vee$, to avoid confusion with the dual vector space $\operatorname{Hom}(L, \mathbb{Q})$.

Since $\deg F \geq 3$, [Deb03, Lemma 3] (or Starr [BHB06, Appendix], of which I learned from [Dao10]) proves more: if m is *even*, then there are at most *finitely many* L's of dimension $m/2$. We would like to understand these "maximal" linear spaces L in terms of the delta method. Proposition 6.3.2 suggests one plausible *starting* route: duality (i.e. detecting $L^\perp$ through $F^\vee$).

Proposition 6.3.2. *Suppose $2 \mid m \geq 4$, and fix $L \in C(\mathrm{SSV})$. Then $\gamma|_{\mathbb{P}L}$ is a finite flat surjective morphism $\mathbb{P}L \to \mathbb{P}L^\perp$ of degree $2^{m/2-1} \geq 2$, and $\mathbb{P}L^\perp \subseteq \operatorname{Sing}(V^\vee)$. Furthermore,* if $\mathbb{P}L^\perp \not\subseteq [\nabla F](\operatorname{hess}(V))$ *set-theoretically, then the affine jet* $j^{2^{m/2-1}-1} F^\vee$ vanishes *over* $L^\perp$.

[1] One can prove density using the general Proposition 1.3.1, or the real [Law19].

Remark 6.3.3. For diagonal F, we provide an explicit proof of Proposition 6.3.2 in §6.3.2. For general $\mathbb{P}^{m-1}_{\mathbb{Q}}$-smooth F, we instead rely on some classical algebraic geometry—duality theory and the ramification behavior of $[\nabla F]$—detailed in §6.2.

We now begin the proof of Proposition 6.3.2. We know (from earlier) that $\gamma|_{\mathbb{P}L}$ is a finite map from $\mathbb{P}L$ to $\mathbb{P}L^\perp$. Yet $\dim(L) = \dim(L^\perp)$. So $\gamma|_{\mathbb{P}L}$ is *surjective*, since $\mathbb{P}L^\perp$ is irreducible. But γ maps V into $V^\vee$ by definition, so $\mathbb{P}L^\perp \subseteq \operatorname{im}\gamma|_{\mathbb{P}L} \subseteq \operatorname{im}\gamma \subseteq V^\vee$.

Proof of first part. Since $\gamma|_{\mathbb{P}L}$ is finite surjective and $\mathbb{P}L, \mathbb{P}L^\perp$ are smooth, "miracle flatness" implies flatness of $\gamma|_{\mathbb{P}L}$. Also, $\gamma|_{\mathbb{P}L}$ has degree $2^{m/2-1}$ (cf. [Dol12, top of p. 29]), since it is a morphism given by quadratic polynomials, between projective spaces of dimension $m/2 - 1$. In particular, $\gamma|_{\mathbb{P}L} \colon \mathbb{P}L \to \mathbb{P}L^\perp$ is *nowhere birational*, so the biduality theorem implies $\mathbb{P}L^\perp \subseteq \operatorname{Sing}(V^\vee)$. □

The second part of Proposition 6.3.2 is inspired by the factorization of $F^\vee$ over $\overline{\mathbb{Q}}[\boldsymbol{c}^{1/2}]$ when F is diagonal (see (6.1) in §6.3.2 below). However, giving a rigorous "factorization" of $F^\vee$ seems to require a bit of setup, since the map $[\nabla F]$ presumably need not be Galois in general.

First, *assume* $\mathbb{P}L^\perp \not\subseteq [\nabla F](\operatorname{hess}(V))$. Now consider the hypersurface complements $S := \mathbb{P}^{m-1} \setminus [\nabla F](\operatorname{hess}(V))$ and $X := [\nabla F]^{-1}S \subseteq \mathbb{P}^{m-1}$. Then $S \cap \mathbb{P}L^\perp$ is a *nonempty* open subset of $\mathbb{P}L^\perp$, yet $[\nabla F]|_X \colon X \to S$ is finite *étale* of degree 2^{m-1}. Write $\phi := [\nabla F]|_X$. By Grothendieck's Galois theory, there exists a finite étale Galois cover $\pi \colon X' \to X$ with X' connected and $\phi \circ \pi \colon X' \to S$ (finite étale) Galois. Let $G := \operatorname{Aut}_S(X')$ and $H := \operatorname{Aut}_X(X')$.

Constructing a product "divisible" by $F^\vee$

Over every geometric point $[\boldsymbol{c}] \in (S \cap V^\vee)(\overline{\mathbb{Q}})$, there exists a (geometric) point $[\boldsymbol{x}] \in (X \cap V)_{[\boldsymbol{c}]}$, i.e. $[\boldsymbol{x}] \in X_{[\boldsymbol{c}]}$ with $F(\boldsymbol{x}) = 0$. Although G may not act on X itself, it acts transitively on X'—so after fixing a (geometric) point $p \in X'_{[\boldsymbol{x}]}$, we can characterize the fiber $X_{[\boldsymbol{c}]}$ as the set $\{\pi(gp) : g \in H \backslash G\} \subseteq X(\overline{\mathbb{Q}})$.

Now view $F(\boldsymbol{x})$ as a section of $\mathcal{O}_X(3)$, and pull it back to $\pi^*F \in \Gamma(X', \pi^*\mathcal{O}_X(3))$. Then the product $\alpha := \prod_{g \in H\backslash G} g^*(\pi^*F)$ defines a *G-invariant section*[2] of the *G-equivariant line bundle* $\mathcal{L} := \bigotimes_{g \in H\backslash G} g^*(\pi^*\mathcal{O}_X(3))$ on X', with

$$\alpha|_{(\phi\pi)^{-1}(S \cap V^\vee)} = 0 \quad (\text{since } X \cap V \subseteq \phi^{-1}(S \cap V^\vee), \text{ and } F|_{X \cap V} = 0).$$

By faithfully flat (Galois) descent,[3] there exist line bundles $\mathcal{F}, \mathcal{D}$ on X, S with $\mathcal{L} \cong \pi^*\mathcal{F}$ and $\mathcal{F} \cong \phi^*\mathcal{D}$,[4] and sections β, δ on X, S vanishing along $\phi^{-1}(S \cap V^\vee), S \cap V^\vee$, respectively, with $\alpha = \pi^*\beta$ and $\beta = \phi^*\delta$.

But S, X are hypersurface complements in $\mathbb{P}^{m-1}$, so $\operatorname{Pic}(\mathbb{P}^{m-1}) \to \operatorname{Pic}(S)$ and $\operatorname{Pic}(\mathbb{P}^{m-1}) \to \operatorname{Pic}(X)$ are surjective and we may identify $\mathcal{F}, \mathcal{D}$ with suitable powers of $\mathcal{O}_X(1), \mathcal{O}_S(1)$, respectively. Then up to a choice of nonzero constant factors, we may view β, δ as homogeneous rational functions (i.e. ratios of homogeneous m-variable polynomials) with $F^\vee(\boldsymbol{c}) \mid \delta$ and $F^\vee(\nabla F(\boldsymbol{x})) = \phi^*F^\vee \mid \beta$. Here we interpret divisibility of two sections on a scheme to mean their ratio is a *global section* of the obvious "tensor-quotient" line bundle.

[2] A *G-invariant section* $\alpha \in \Gamma(X', \mathcal{L})$ is equivalent in data to a *G-equivariant morphism* $\alpha \colon \mathcal{O}_{X'} \to \mathcal{L}$.

[3] in the form of an equivalence of categories

[4] It suffices to find $\mathcal{D}$ with $\mathcal{L} \cong (\phi\pi)^*\mathcal{D}$, and then define $\mathcal{F} := \phi^*\mathcal{D}$.

“Factoring” $F^\vee$

By duality theory, $F(\boldsymbol{x}) \mid F^\vee(\nabla F(\boldsymbol{x}))$. However, $F(\boldsymbol{x})^2 \nmid \beta$ on X, since $(\pi^*F)^2 \nmid \alpha$ on X'. Indeed, given a geometric point $[\boldsymbol{c}]$ of $(S \cap V^\vee) \setminus \operatorname{Sing}(V^\vee)$, biduality furnishes a *unique* point $[\boldsymbol{x}] \in X_{[\boldsymbol{c}]}$ with $F(\boldsymbol{x}) = 0$. So if $p \in X'_{[\boldsymbol{x}]}$ and $g \in G$, then the section $g^*\pi^*F$ evaluates to 0 at p *if and only if* $g \in H$. Thus $\pi^*F \nmid \prod_{[1]\neq g\in H\backslash G}(g^*\pi^*F)$ on X'. It follows that $(\pi^*F)^2 \nmid \alpha$, whence $F(\boldsymbol{x})^2 \nmid \beta$; whence $F^\vee(\boldsymbol{c})^2 \nmid \delta$.

However, $F \mid \phi^*F^\vee$, and $\phi\pi g = \phi\pi$ for all $g \in G$, so by inspection, $\alpha \mid (\pi^*\phi^*F^\vee)^{\deg\phi}$, i.e. $\delta \mid (F^\vee)^{\deg\phi}$. By absolute irreducibility of $F^\vee$, we conclude that in fact $\delta \mid F^\vee$ and $\alpha \mid (\phi\pi)^*F^\vee$. (Divisibility also holds in the other direction, but we will not need this.)

Remark 6.3.4. Our proof of $\delta \mid F^\vee$ requires $(S \cap V^\vee) \setminus \operatorname{Sing}(V^\vee) \neq \emptyset$, i.e. that $S \cap V^\vee$ and $V^\vee \setminus \operatorname{Sing}(V^\vee)$ are nonempty (open) subsets of $V^\vee$; since $\mathbb{P}L^\perp \subseteq V^\vee$, the former nonemptiness follows conveniently from our assumption $S \cap \mathbb{P}L^\perp \neq \emptyset$ (but see Question 6.2.5), while the latter follows (unconditionally) from “generic smoothness” in characteristic 0 (since $V^\vee$ is reduced).

Differentiating the product

Using $S \cap \mathbb{P}L^\perp \neq \emptyset$ one last time (more seriously than before), we will now complete the proof of the second part of Proposition 6.3.2.

Completion of proof. By assumption, $S \cap \mathbb{P}L^\perp \neq \emptyset$. Yet $\phi|_{X\cap\mathbb{P}L} = \gamma|_{X\cap\mathbb{P}L} \colon X \cap \mathbb{P}L \to S \cap \mathbb{P}L^\perp$ is finite étale of degree $2^{m/2-1}$. Say $[\boldsymbol{c}] \in (S \cap \mathbb{P}L^\perp)(\overline{\mathbb{Q}})$ is a geometric point, and fix $p \in X'_{[\boldsymbol{c}]}$. Then there exist at least $2^{m/2-1}$ cosets $g \in H\backslash G$ with $\pi g p \in (X \cap \mathbb{P}L)(\overline{\mathbb{Q}}) \subseteq V(\overline{\mathbb{Q}})$. In a G-invariant affine open neighborhood of (the image of) p in X',[5] the Leibniz rule—applied after locally trivializing $g^*\pi^*\mathcal{O}_X(3)$ for all $g \in G$—thus implies that $j^r_p\alpha(p) = \mathbf{0}$ for $r := 2^{m/2-1} - 1$, where $j^r \colon \mathcal{L} \to J^r\mathcal{L}$ denotes the rth-order jet map “along” $\mathcal{L}$ (from $\mathcal{L}$ to its rth jet bundle $J^r\mathcal{L}$).

Since $\alpha \mid (\phi\pi)^*F^\vee$, Leibniz then implies $j^r_p(\phi\pi)^*F^\vee(p) = \mathbf{0}$ “along” the pullback line bundle $(\phi\pi)^*\mathcal{O}_S(\deg F^\vee)$. But $\phi\pi \colon X' \to S$ is étale at $p \in X'$, so $j^r_{[\boldsymbol{c}]}F^\vee([\boldsymbol{c}]) = \mathbf{0}$ “along” $\mathcal{O}_S(\deg F^\vee)$ itself, over all points $[\boldsymbol{c}] \in (S \cap \mathbb{P}L^\perp)(\overline{\mathbb{Q}})$. Finally, $S \cap \mathbb{P}L^\perp$ is dense in $\mathbb{P}L^\perp$, so the vanishing of the rth-order jet section $j^rF^\vee$ extends to all points $[\boldsymbol{c}] \in \mathbb{P}L^\perp$, as desired. $\square$

Remark 6.3.5. In the friendly setting above, our étale morphisms (such as $\phi\pi \colon X' \to S$), after base change to an algebraically closed field, always induce isomorphisms on completed local rings. So e.g. at regular points we can do calculus purely in terms of formal power series (in general by the Cohen structure theorem, but for us S is already given as a piece of $\mathbb{P}^{m-1}$).

Question 6.3.6. For smooth V, is $\mathbb{P}L^\perp \subseteq [\nabla F](\operatorname{hess}(V))$ possible (for $L \in C(\mathrm{SSV})$)? If so, then in such situations, does the conclusion of Proposition 6.3.2 still hold?

[5] Such a neighborhood exists by [Mus11, paragraph after Corollary A.3], since X' is certainly quasi-projective.

6.3.2 The diagonal case

Say F is diagonal, and write $F = F_1x_1^3 + \cdots + F_mx_m^3$. Then we can explicitly verify all the theory above. Here $[\nabla F]\colon [\boldsymbol{x}] \mapsto [3F_1x_1^2, \dots, 3F_mx_m^2]$ is Galois with abelian Galois group $\mu_2^m/\mu_2 \cong (\mathbb{Z}/2)^{m-1}$, and $\operatorname{hess}(V)$ is cut out by $(6F_1x_1)\cdots(6F_mx_m) = 0$.

Describing $C(\mathrm{SSV})$

Since F is diagonal, the number of $m/2$-dimensional vector spaces $L \subseteq C(V)_{\mathbb{C}}$ over $\mathbb{C}$ is $(\deg F)^{m/2} = 3^{m/2}$ times $C_{m/2-1} := (m-1)!!$, the number of *pairings* of $[m]$. (See e.g. [BHB06, Starr's Appendix, top of p. 302] for a more general statement on Fermat hypersurfaces of degree $d \geq 3$ in $s \in \{4, 6, 8, \dots\}$ variables.) So we make a combinatorial definition:

Definition 6.3.7. Let $\mathcal{J} = (\mathcal{J}(k))_{k\in\mathcal{K}}$ denote an *ordered set partition* of $[m]$: a list of pairwise disjoint nonempty sets $\mathcal{J}(k) \subseteq [m]$ covering $[m]$, indexed by a set $\mathcal{K} \in \{[1], [2], [3], \dots\}$. Call $\mathcal{J}, \mathcal{J}'$ *equivalent* if they define the same unordered partition of $[m]$ (i.e. if $\mathcal{K} = \mathcal{K}'$ and $\{\mathcal{J}(k) : k \in \mathcal{K}\} = \{\mathcal{J}'(k) : k \in \mathcal{K}'\}$).

Call $\mathcal{J}$ a *pairing* if $|\mathcal{J}(k)| = 2$ for all $k \in \mathcal{K}$. Call $\mathcal{J}$ *permissible* if for all $k \in \mathcal{K}$ and $i, j \in \mathcal{J}(k)$, we have $F_j/F_i \in (\mathbb{Q}^\times)^3$. For a permissible $\mathcal{J}$, let $\mathcal{R}_{\mathcal{J}} := \{\boldsymbol{c} \in \mathbb{Z}^m :$ if $k \in \mathcal{K}$ and $i, j \in \mathcal{J}(k)$, then $c_i/F_i^{1/3} = c_j/F_j^{1/3}\}$—and given $\boldsymbol{c} \in \mathcal{R}_{\mathcal{J}}$, define $c\colon \mathcal{K} \to \mathbb{R}$ so that for all $k \in \mathcal{K}$ and $i \in \mathcal{J}(k)$, we have $c_i/F_i^{1/3} = c(k)$.

Knowing the number of $L \subseteq C(V)_{\mathbb{C}}$, we can then give an exhaustive construction: each equivalence class of pairings $\mathcal{J}$ yields $3^{m/2}$ distinct $L/\mathbb{C}$, obtained by setting $F_ix_i^3 + F_jx_j^3 = 0$ for each part $\mathcal{J}(k) = \{i, j\}$. Over $\mathbb{Q}$, we must set $x_i + (F_j/F_i)^{1/3}x_j = 0$—which is valid when $F_i \equiv F_j \bmod (\mathbb{Q}^\times)^3$. It follows that the only $m/2$-dimensional L's on $C(V)_{\mathbb{C}}, C(V)$ (over $\mathbb{C}, \mathbb{Q}$, respectively) are the "obvious" ones.

Remark 6.3.8. Since we do not know of a reference *proving* the aforementioned statement of Starr, we should mention that given d, s, the statement easily follows from Gaussian elimination, symmetry, and the fact that over $\mathbb{C}$, the only *linear* automorphisms of the "halved" (i.e. $s/2$-variable degree-d) Fermat hypersurface are the "obvious" ones (see e.g. [Shi88] or [Kon02, proofs of Proposition 3.1 and Example 1]).

Therefore we obtain the following (essentially classical) result:

Observation 6.3.9. There is a canonical bijection, between $C(\mathrm{SSV})$ and the set of equivalence classes of *permissible* pairings $\mathcal{J}$, characterized by $L \cap \mathbb{Z}^m = \mathcal{R}_{\mathcal{J}}^{\perp}$ (an equality of sublattices of $\mathbb{Z}^m$).

Analyzing the discriminant

For convenience, fix square roots $F_i^{1/2} \in \overline{\mathbb{Q}}^\times$. Up to scaling (which matters in the non-archimedean analysis underlying Lemma 6.4.1, but not here), the discriminant form $F^\vee(\boldsymbol{c})$ factors in $\overline{\mathbb{Q}}[\boldsymbol{c}^{1/2}]$ as

$$\prod_{\boldsymbol{\epsilon}} (\epsilon_1 F_1^{-1/2} c_1^{3/2} + \epsilon_2 F_2^{-1/2} c_2^{3/2} + \cdots + \epsilon_m F_m^{-1/2} c_m^{3/2}) \in \mathbb{Q}[\boldsymbol{c}], \tag{6.1}$$

with the product taken over $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_m)$ with $\epsilon_1 = 1$ and $\epsilon_2, \ldots, \epsilon_m = \pm 1$. (This formula is classical; see [Wan21c, §1.2, proof of Proposition-Definition 1.8 for diagonal F] or [HB98, eq. (4.2)].)

Now fix a tuple $\boldsymbol{c} \neq \mathbf{0}$ with $F^\vee(\boldsymbol{c}) = 0$, and fix square roots $c_i^{1/2} \in \overline{\mathbb{Q}}$. Then using formal power series calculus over $c_i \neq 0$ (by Remark 6.3.5, adapted to $\mathbb{A}^1_{\overline{\mathbb{Q}}} \to \mathbb{A}^1_{\overline{\mathbb{Q}}}$, $t \mapsto t^2$ away from the origin), we will prove the following result, which precisely characterizes the order of vanishing of $F^\vee$ at $\boldsymbol{c}$:

Proposition 6.3.10. *Fix $r \geq 0$. Then the affine jet $j^r F^\vee$ vanishes at a given point $\boldsymbol{c} \neq \mathbf{0}$ if and only if there exist at least $r+1$ distinct $\boldsymbol{\epsilon}$ with $(\cdots) = 0$.*

Remark 6.3.11. A short computation yields the equality

$$\#\{\text{such distinct } \boldsymbol{\epsilon}\text{'s}\} = \sum_{[\boldsymbol{x}] \in \gamma(\overline{\mathbb{Q}})^{-1}([\boldsymbol{c}])} 2^{\#\{i \in [m] : x_i = 0\}},$$

where $\gamma(\overline{\mathbb{Q}})^{-1}([\boldsymbol{c}]) := \{[\boldsymbol{x}] \in V(\overline{\mathbb{Q}}) : [\nabla F(\boldsymbol{x})] = [\boldsymbol{c}]\} = \{\text{singular } \overline{\mathbb{Q}}\text{-points of } V_{\boldsymbol{c}}\}$. (Here x_i corresponds to $\epsilon_i F_i^{-1/2} c_i^{1/2}$, with some ambiguity or "multiplicity" in ϵ_i when $x_i = 0$.) The previous display provides a geometric interpretation of the number of $\boldsymbol{\epsilon}$'s in Proposition 6.3.10; thus we can formulate the proposition more geometrically, without reference to $\boldsymbol{\epsilon}$'s. Does this *geometric formulation* generalize somehow to arbitrary $\mathbb{P}^{m-1}_{\mathbb{Q}}$-smooth F?

Proof. Induct on $r \geq 0$. The base case $r = 0$ follows directly from the factorization of $F^\vee$. Now suppose $r \geq 1$, and assume the result for $r - 1$.

First, we prove the forwards implication for r. Here it suffices to work with "pure" derivatives $\partial_{c_i}^{\leq r}$, for just a single index i with $c_i \neq 0$. For example, if $c_1 \neq 0$, and there exist *exactly* r distinct $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_r$ with $(\cdots) = 0$, then $r \leq 2^{m-1}$, and the product rule implies

$$\partial_{c_1}^r F^\vee(\boldsymbol{c}) \propto_{3,r,F} (c_1^{1/2})^r \prod_{\boldsymbol{\epsilon} \neq \boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_r} (\cdots) \neq 0.$$

But if $j^r F^\vee(\boldsymbol{c}) = 0$, then by the inductive hypothesis, there must exist *at least* r distinct $\boldsymbol{\epsilon}$'s with $(\cdots) = 0$, and thus at least $r+1$. This proves the the forwards implication for r.

It remains to prove the backwards implication, i.e. that if there exist *at least* $r+1$ distinct $\boldsymbol{\epsilon}$'s with $(\cdots) = 0$, then $j^r F^\vee(\boldsymbol{c}) = \mathbf{0}$. We must take extra care if $c_1 \cdots c_m = 0$. Say $c_i = 0$ for $i \in I$. Then the following "formal analytic functions"—indexed by certain triples $(\boldsymbol{a}, \boldsymbol{b}, E)$—span a $\overline{\mathbb{Q}}$–vector space *closed under $\boldsymbol{c}$-differentiation*:

$$\prod_{i \in I} c_i^{a_i} \prod_{i \notin I} c_i^{b_i/2} \prod_{\boldsymbol{\epsilon} \in E} (\epsilon_1 F_1^{-1/2} c_1^{3/2} + \cdots + \epsilon_m F_m^{-1/2} c_m^{3/2}) \in \overline{\mathbb{Q}}[c_i]_{i \in I}[c_i^{1/2}, c_i^{-1/2}]_{i \notin I},$$

for $(\boldsymbol{a}, \boldsymbol{b}) \in \mathbb{Z}_{\geq 0}^I \times \mathbb{Z}^{[m] \setminus I}$ and $E \subseteq \{\boldsymbol{\epsilon} \in \{\pm 1\}^m : \epsilon_1 = 1\}$ with $E \bmod \pm 1$ invariant under flipping ϵ_i for $i \in I$.

Specifically, differentiating in c_i leads to terms with $a_i \mapsto a_i - 1$ or $(a_i, |E|) \mapsto (a_i + 2, |E| - 2)$ if $i \in I$, and to terms with $b_i \mapsto b_i - 2$ or $(b_i, |E|) \mapsto (b_i + 1, |E| - 1)$ if $i \notin I$. In each case, applying ∂_{c_i} decreases $\min_{\boldsymbol{a}, \boldsymbol{b}, E}(|\boldsymbol{a}| + |E|)$ by at most 1.

Now fix $\boldsymbol{r} \in \mathbb{Z}_{\geq 0}^m$ with $|\boldsymbol{r}| \leq r$. Then $\partial_{\boldsymbol{c}}^{\boldsymbol{r}} F^\vee$ is a $\overline{\mathbb{Q}}$-linear combination of functions indexed by triples $(\boldsymbol{a}, \boldsymbol{b}, E)$ with $|\boldsymbol{a}| + |E| \geq 2^{m-1} - r$ (and thus $|E| \geq 2^{m-1} - r$ or $|\boldsymbol{a}| \geq 1$). Each such function must vanish at our original given point $\boldsymbol{c}$, so $\partial_{\boldsymbol{c}}^{\boldsymbol{r}} F^\vee = 0$, as desired. $\square$

Evaluating $F^\vee$ on $L^\perp$ for special L's

Fix $L \in C(\mathrm{SSV})$. Proposition 6.3.10 has the following corollary:

Corollary 6.3.12. *For L as above, we have $(j^{2^{m/2-1}-1}F^\vee)|_{L^\perp} = \mathbf{0}$.*

Proof. By Observation 6.3.9, L corresponds to some permissible pairing $\mathcal{J}$. For each part $\mathcal{J}(k) = \{i, j\}$, there are exactly two choices of signs $(\epsilon_i, \epsilon_j) \in \{\pm 1\}^2$—or only one choice if $1 \in \mathcal{J}(k)$—such that $\epsilon_i F_i^{-1/2} c_i^{3/2} + \epsilon_j F_j^{-1/2} c_j^{3/2}$ vanishes over *all* $\boldsymbol{c} \in L^\perp \cap \mathbb{Z}^m = \mathcal{R}_{\mathcal{J}}$ lying in a given orthant of $\mathbb{R}^m$. So given $\boldsymbol{c} \in L^\perp \setminus \{\mathbf{0}\}$, we can apply Proposition 6.3.10 "backwards" with $r := 2^{m/2-1} - 1$. (Of course, $L^\perp \setminus \{\mathbf{0}\}$ is dense in $L^\perp$, so the vanishing then extends to all of $L^\perp$.) □

Thus we have explicitly verified the conclusion of Proposition 6.3.2. The next result shows that in fact, $F^\vee$ generally does not vanish to higher order along $L^\perp$.

Observation 6.3.13. Given $L, \mathcal{J}$ as above, fix $\boldsymbol{c} \in L^\perp \cap \mathbb{Z}^m = \mathcal{R}_{\mathcal{J}}$. For each $k \in \mathcal{K}$, choose a square root $c(k)^{3/2} := F_i^{-1/2} c_i^{3/2}$ in $\overline{\mathbb{Q}}$ (say). Then $j^{2^{m/2-1}}F^\vee(\boldsymbol{c}) = \mathbf{0}$ if and only if there exist $l \geq 1$ distinct indices $k_1, \dots, k_l \in \mathcal{K}$ such that $c(k_1)^{3/2} \pm \cdots \pm c(k_l)^{3/2} = 0$ for some choice of signs. Consequently, if $j^{2^{m/2-1}}F^\vee(\boldsymbol{c}) = \mathbf{0}$, then $c(k_1)^3 c(k_2)^3 \in (\mathbb{Q}^\times)^2 \cup \{0\}$ for some distinct $k_1, k_2 \in \mathcal{K}$.

Proof. For the equivalence, apply Proposition 6.3.10 "forwards" with $r := 2^{m/2-1}$ (and then "simplify" the resulting conclusion using the fact that $\mathcal{J}$ is a *pairing*). To obtain the final conclusion, note that the condition $c(k_1)^{3/2} \pm \cdots \pm c(k_l)^{3/2} = 0$ *implies* the following, provided l is *minimal* among all possible l's (as we may certainly assume):

(1) If $l = 1$, then $c(k)^3 = 0$ for some $k \in \mathcal{K}$.

(2) If $l = 2$ is *minimal*, then $c(k_1)^3 = c(k_2)^3 \in \mathbb{Q}^\times$ for some distinct $k_1, k_2 \in \mathcal{K}$.

(3) If $l \geq 3$ is *minimal*, then $c(k_t)^3 \in \mathbb{Q}^\times$ for all $t \in [l]$, and by multi-quadratic field theory in characteristic 0, the square classes $c(k_1)^3, \dots, c(k_l)^3 \bmod (\mathbb{Q}^\times)^2$ must all coincide. (More precisely, given indices $i_t \in \mathcal{J}(k_t)$ for $t \in [l]$, we must have $c(k_t)^3 = F_{i_t}^2 x_{i_t}^6 d^3 \in d \cdot (\mathbb{Q}^\times)^2$ for some $d, x_{i_t} \in \mathbb{Q}^\times$ such that $F_{i_1} x_{i_1}^3 + \cdots + F_{i_l} x_{i_l}^3 = 0$. If $\mathcal{J}(k_1), \dots, \mathcal{J}(k_l)$ cover $[m]$, as must be the case if $m = 6$, then this would imply that $[\boldsymbol{c}] \in \mathbb{P}L^\perp$ actually lies in the image of $\gamma(\mathbb{Q})$—unlike most points of $\mathbb{P}L^\perp$.)

In each case, the claimed multiplicative relationship exists for some distinct $k_1, k_2 \in \mathcal{K}$. This completes the proof. □

Remark 6.3.14. §6.3.2 is written in characteristic 0, but since $F^\vee \in \mathbb{Z}[\boldsymbol{c}]$, the results, with their *algebraic* proofs, carry over to arbitrary fields of characteristic $p \nmid (6^m)! F_1 \cdots F_m$. Over $\mathbb{F}_p$, such extensions of Proposition 6.3.10 (in its geometric formulation), Corollary 6.3.12, and (the equivalence part of) Observation 6.3.13 prove useful in the proof of the important Lemma 6.4.1 stated below.

(Though unimportant for us over $\mathbb{F}_p$, the other results of §6.3.2 also carry over. For instance, regarding the field theory behind the last part of Observation 6.3.13: if K is a field of characteristic $p \nmid 2$, and $d_1, \dots, d_l \in K^\times$ are pairwise incongruent modulo $(K^\times)^2$, then $\sqrt{d_1}, \dots, \sqrt{d_l} \in K(\sqrt{d_1}, \dots, \sqrt{d_l})$ are linearly independent over K.)

6.4 Some key ingredients

The following lemma is close in spirit to one direction of Corollary 5.1.31 (see also Question 5.1.1). The proof in [Wan21d, §6] begins with a change of coordinates somewhat related to van der Corput or Weyl differencing, or (probably) to certain blow-ups along $m/2$-planes; see [Wan21d, Remark 6.5].

Lemma 6.4.1 ([Wan21d, Lemma 5.5]). *Assume F is diagonal, with $m \in \{4, 6\}$. Fix $L \in C(\mathrm{SSV})$. Suppose $\boldsymbol{c} \in \Lambda^{\perp}$ is trivial, and $p \nmid (j^{2^{m/2-1}} F^{\vee})(\boldsymbol{c})$ is a prime. Then*

$$\widetilde{S}_{\boldsymbol{c}}(p) = \phi(p)p^{-1/2} + O(1).$$

Also, if we take $\mathcal{J}$ corresponding to L via Observation 6.3.9, then in the notation of Definition 6.3.7, $c(k)^3 \in \mathbb{Q}$ is a well-defined p-adic unit for all k, and

$$\widetilde{S}_{\boldsymbol{c}}(p^l) = \phi(p^l)p^{-l/2} \cdot \prod_{k \in [m/2-1]} \left[1 + \chi\left(c(k)^3 c(k+1)^3\right)\right] \ll \phi(p^l)p^{-l/2}$$

for all integers $l \geq 2$, if we let $\chi(r) := (\frac{r}{p})$. Both implied constants depend only on m.

To state the next ingredient, let $\mathcal{S} \subseteq \{\boldsymbol{c} \in \mathbb{Z}^m : F^{\vee}(\boldsymbol{c}) = 0\}$ be a *homogeneous* (i.e. invariant under scaling, so $\boldsymbol{c} \in \mathcal{S}$ implies $\mathbb{Z} \cdot \boldsymbol{c} \subseteq \mathcal{S}$) subset of $\boldsymbol{c}$'s with $F^{\vee}(\boldsymbol{c}) = 0$. At several technical points in the proof of Theorem 6.1.1 (see [Wan21d, §5]), the following lemma lets us cleanly discard various contributions from *sparse* homogeneous sets $\mathcal{S}$—when restricted to $\boldsymbol{c} \neq \boldsymbol{0}$, at least. The proof (based on [HB98, pp. 688–689]) is rather awkward, due to a lack of a deeper algebro-geometric theory for $S_{\boldsymbol{c}}$ when $V_{\boldsymbol{c}}$ is singular.

Lemma 6.4.2 ([Wan21d, Lemma 4.1]). *Assume F is diagonal. If $\mathcal{S} \cap [-C, C]^m$ has size $O(C^{m/2-\delta+\epsilon})$ for all $C \gg 1$, then*

$$Y^{-2} \sum_{\boldsymbol{c} \in \mathcal{S} \setminus \{\boldsymbol{0}\}} \sum_{n \geq 1} |I_{\boldsymbol{c}}(n)| \cdot n^{1-m/2} \cdot \max_{n_\star \mid n} n_\star^{-1/2} |\widetilde{S}_{\boldsymbol{c}}(n_\star)| \ll_\epsilon X^{(m-\delta)/2+\epsilon}$$

provided $4 \leq m \leq 6$ *and* $\delta \leq \min((m+2)/4, (m-1)/2)$.

Remark 6.4.3. We can "axiomatize" our proof of Theorem 6.1.1, in the hope of extending Theorem 6.1.1 to general non-diagonal F (though a full diagnosis of the relevant issues would seem to require algebraic geometry beyond the author's current expertise). Assume

(1) $m \geq 4$ is even, F has nonzero discriminant, and $F^{\vee}$ is unsurprising (in the sense of Definition 6.1.7);

(2) (F, w) is clean in the sense of Definition 1.4.3;

(3) a version of Lemma 6.4.2, with the same hypotheses but with a modified conclusion

$$\frac{X^m}{Y^2} \sum_{\boldsymbol{c} \in \mathcal{S} \setminus \{\boldsymbol{0}\}} \sum_{n \geq 1} \left(1 + \frac{X\|\boldsymbol{c}\|}{n}\right)^{1-m/2} \nu_1\left(\frac{\|\boldsymbol{c}\|}{X^{1/2}}\right) \cdot n^{(1-m)/2} |\widetilde{S}_{\boldsymbol{c}}(n)| \ll_\epsilon X^{(m-\delta)/2+\epsilon},$$

remains true, where $\nu_1(\star)$ is a fixed function decaying as $O_A(\max(1, \star)^{-A})$; and

(4) in Lemma 6.4.1, the formula for $S_{\boldsymbol{c}}(p)$ and upper bound for $S_{\boldsymbol{c}}(p^{\geq 2})$ remain true, provided $p \gg_{3,m,F} 1$ exceeds some threshold (only allowed to depend on $3, m, F$).

Then the conclusion of Theorem 6.1.1 still holds for F, as does that of Corollary 6.1.4 if $m \geq 6$. (Depending on the precise "Definition 6.1.7" in (1), a weaker version of (3) may suffice.)

Remark 6.4.4. In Remark 6.4.3, we expect that (4) should be relatively routine to prove (if true), but (1) and (3) may well require substantial new ideas.

6.5 Discussion of Manin-type conjectures

Let $F, V, \ldots$ be as in §3.1. The Hardy–Littlewood "randomness" (singular series) prediction for $F = 0$ may fail, even when $m = 6$ (see Example 1.2.1).

Remark 6.5.1. The "randomness" prediction is expected to hold for $m \geq 7$, since V is *smooth*. (At least when $m \geq 8$ and (F, w) is clean, [Hoo15] provides a conditional affirmative proof, unconditional for $m \geq 9$.) For *singular* cubics, however, failure can occur even when $m \geq 7$; see [BW19] for an interesting example when $m = 8$.

But every "randomness failure" should have a good *excuse*!

A plausible version of Manin's conjecture (in a smoothed form) for $C(\mathcal{V}) \subseteq \mathbb{A}^m_{\mathbb{Z}}$ says that away from a certain *special structured locus*—namely the empty set if $m \geq 7$, and the union of all (*linear*) vector spaces $L \subseteq C(V)$ over $\mathbb{Q}$ of dimension $\lfloor m/2 \rfloor$ if $m \leq 6$—one should have

$$N_{F,w}(X) - \sum_{\boldsymbol{x} \in \mathbb{Z}^m} w(\boldsymbol{x}/X) \cdot \mathbf{1}_{\boldsymbol{x} \in \bigcup_L L} = (c + o_{F,w;X\to\infty}(1)) \cdot X^{m-3} (\log X)^{r-1+\mathbf{1}_{m=4}}$$

for a certain precise *predicted* constant $c = c_{C(\mathrm{MP}),w} \in \mathbb{R}$ and integer rank $r \geq 1$.

Remark 6.5.2. If $m \geq 5$, then $r = 1$ always—so the $\log X$ disappears—while $c_{C(\mathrm{MP}),w}$, the "coned" (or "unsieved") Manin–Peyre constant, always equals the Hardy–Littlewood constant $c_{\mathrm{HL},w}$. If $m = 4$, then *typically*, but not always, we have $(r, c_{C(\mathrm{MP}),w}) = (1, c_{\mathrm{HL},w})$—if one interprets $c_{\mathrm{HL},w}$ generously as in [Jah14, Chapter II, Remarks 7.5–7.7].

In these cases, "Manin" differs from "Hardy–Littlewood" only in the *special* part. (But *sometimes* when $m = 4$, Brauer–Manin obstructions or other phenomena lead to *further* differences; e.g. loosely speaking, it is possible for "$c_{\mathrm{HL},w} \neq 0 = c_{C(\mathrm{MP}),w}$" or related behavior to occur.)

For a *general* overview of "Manin", see [Bom09, §2 and references within] or [Bro09, Jah14]. In general, "Manin" is imprecise regarding the *special* part (though nowadays there are precise, but sometimes complicated, proposals available; see e.g. [LST19]). However, at least when $m = 6$ and $F = x_1^3 + \cdots + x_6^3$, the *specific* conjecture recorded above was—in essence—stated first, and precisely, by Hooley [Hoo86a, Conjecture 2]. (See also [VW95, Appendix].)

Finally, we raise some further questions (about "more complicated" varieties) that one also might (at first) study in the "$\boldsymbol{c}$-restricted" sense of Theorem 6.1.1.

Example 6.5.3. It would be interesting to extend our analysis to $m = 5$, even just for diagonal F. See [Bom09, §3] for a discussion of the potentially infinite family of lines on a cubic threefold $V \subseteq \mathbb{P}^4$. Can one see these lines via $V^\vee$ (cf. Proposition 6.3.2)?

Now consider the following example (which I learned from a talk of Wooley) of a situation in which (one expects that) *nonlinear* special subvarieties arise.

Example 6.5.4 ([Woo19]). Over boxes $[-X, X]^6$ as $X \to \infty$, one expects the 6-variable quartic $x_1^4 + x_2^4 + x_3^4 = x_4^4 + x_5^4 + x_6^4$ to have not only a "purely probabilistic" source of $\asymp X^2$ points (as $X \to \infty$), but also at least two relevant "special" sources of points:

(1) the "trivial" or "diagonal-type" linear locus

$$x_1 \pm x_4 = x_2 \pm x_5 = x_3 \pm x_6 = 0 \quad \text{(and the obvious permutations)}$$

contributing $\asymp X^3$ points (as $X \to \infty$), as well as

(2) a secondary *quadratic* locus

$$x_1 \pm x_2 \pm x_3 = x_4 \pm x_5 \pm x_6 = (x_1^2 + x_2^2 + x_3^2) \pm (x_4^2 + x_5^2 + x_6^2) = 0$$

contributing $\asymp X^2 \log X$ points (as $X \to \infty$).

(The underlying identity behind (2) is $a^4 + b^4 + (a+b)^4 = 2(a^2 + ab + b^2)^2$.)

Question 6.5.5. Can one detect (2) naturally via the delta method; or if not, why? Where might the quadratic aspect naturally arise?

Remark 6.5.6. Question 5.1.1 may or may not be relevant to this question.

Chapter 7

Discriminating pointwise estimates

In Chapter 8, we would like to "remove the ϵ from §4.1" (and go beyond it), at least conditionally. In the present chapter, we will summarize the estimates from [Wan21a] designed in response to Remarks 4.1.1 and 4.1.3 towards Chapter 8. These estimates may well be related to Igusa local zeta functions, model theory, o-minimality, real algebraic geometry, or resolution of singularities (in mixed characteristic). But a consideration of such perspectives will have to wait for now.

Let $F, V, w, \dots$ be as in §3.1. At least for some pairs (F, w), we will present new unconditional pointwise bounds on the oscillatory integrals $I_{\boldsymbol{c}}(n)$ and exponential sums $S_{\boldsymbol{c}}(n)$ (see Lemmas 7.1.4 and 7.2.3 below, respectively), and new conditional average bounds on $S_{\boldsymbol{c}}(n)$ (see Conjecture 7.3.7 (B3) below, and the surrounding discussion).

7.1 Pointwise integral estimates

Definition 7.1.1. Call $\nu_d \colon \mathbb{R}^d \to \mathbb{R}$ a *decay weight* if $\nu_d \in \mathcal{S}(\mathbb{R}^d)$ (i.e. if ν_d is *Schwartz*). For such ν_d, write $\nu_d \geq 0$ to mean $\operatorname{im} \nu_d \subseteq \mathbb{R}_{\geq 0}$, and $\nu_d > 0$ to mean $\operatorname{im} \nu_d \subseteq \mathbb{R}_{>0}$.

Remark 7.1.2. We prefer the somewhat informal name "decay weight" because while the regularity assumptions on ν are very convenient in [Wan21a] (see e.g. [Wan21a, Remark 8.4, and the Hölder argument in §9]), they are not morally essential.

Definition 7.1.3. Define $J_{\boldsymbol{c},X}(n)$ as in [Wan21a, Definition 3.43]; in our setting, $J_{\boldsymbol{c},X}(n) = I_{\boldsymbol{c}}(n)$, but it will be helpful to keep the X-dependence explicit.

Given $\boldsymbol{c} \in \mathbb{R}^m \setminus \{\boldsymbol{0}\}$, let $\tilde{\boldsymbol{c}} := \boldsymbol{c}/\lVert\boldsymbol{c}\rVert$, so that $|F^\vee(\tilde{\boldsymbol{c}})|^{-1} \in [\Theta(1), \infty]$ measures the "degeneracy" of $(V_{\mathbb{R}})_{\boldsymbol{c}}$. In the following statement, we only need V to be smooth, not diagonal; in fact, we only need V to be smooth along $V(\mathbb{R})$, not $V(\mathbb{C})$.

Lemma 7.1.4 ([Wan21a, Lemma 4.5])**.** *Suppose* (F, w) *is* clean *in the sense of Definition 1.4.3. Then for some decay weight* $\nu_m > 0$*,* fixed *in terms of* F, w*, the following statements hold, uniformly over* $(X, \boldsymbol{c}, n) \in \mathbb{R}_{>0} \times \mathbb{R}^m \times \mathbb{R}_{>0}$*:*

[J1] Modulus cutoff*:* $J_{\boldsymbol{c},X}(n) = 0$ *holds unless* $n \ll_{F,w} Y$*.*

[J2] Integral bound*: If we interpret $\boldsymbol{c}/|F^\vee(\tilde{\boldsymbol{c}})|^{-1}$ as $\mathbf{0}$ if $\boldsymbol{c} = \mathbf{0}$, then we* always *have*

$$|J_{\boldsymbol{c},X}(n)| \ll_{F,w} \min\left[1, \left(\frac{X\|\boldsymbol{c}\|}{n}\right)^{1-m/2}\right] \cdot \nu_m\left(\frac{\boldsymbol{c}}{X^{1/2}}\right)^3 \cdot \nu_m\left(\frac{X\boldsymbol{c}}{n|F^\vee(\tilde{\boldsymbol{c}})|^{-1}}\right).$$

[J3] Homogeneous *dimensional analysis: For each integer $j \geq 0$, the same bound holds for $|\partial_{\log n}^j J_{\boldsymbol{c},X}(n)|$, up to an $O_{F,w,j}(1)$ factor loss. (Here $\partial_{\log n} := n \cdot \partial_n$.)*

[J4] Vertical *variation: In [J3], we can replace $J_{\boldsymbol{c},X}(n)$ with $\partial_{\log \boldsymbol{c}}^{\boldsymbol{\alpha}} J_{\boldsymbol{c},X}(n)$ (where we define $\partial_{\log \boldsymbol{c}}^{\boldsymbol{\alpha}} := \prod_{i\in[m]} \partial_{\log c_i}^{\alpha_i}$ for $\boldsymbol{\alpha} \in \mathbb{Z}_{\geq 0}^m$), up to an* additional *multiplicative loss of*

$$O_{F,w,\boldsymbol{\alpha}}(1) \cdot \prod_{i\in[m]} \left(1 + \frac{X|c_i|}{n}\right)^{\alpha_i}.$$

Remark 7.1.5. [HB96, HB98] proved most of Lemma 7.1.4 *except for* [$J2$]–[$J4$]'s "decay over $n \ll \|X\boldsymbol{c}/|F^\vee(\tilde{\boldsymbol{c}})|^{-1}\|$". (Hooley's "pre-delta method" works [Hoo86b, Hoo97] have a similar lacuna.)

Remark 7.1.6. Without the cleanliness assumption, similar but messier bounds in [$J2$]–[$J4$] might still hold, at least for diagonal F (cf. Lemma 3.3.5), and plausibly—at least in a weaker sense—even for general F (using the the o-minimal geometric framework of [BGZZK21]). It would be interesting, and likely useful or enlightening, to find such bounds for smooth pairs (F, w), say, even for diagonal F. However, in order to focus on the most essential issues in this thesis, we have decided to keep the (qualitatively harmless) cleanliness assumption on (F, w).

The requirement $\nu_m \in \mathcal{S}(\mathbb{R}^m)$ in Lemma 7.1.4 (via Definition 7.1.1) may seem unnaturally restrictive at first glance. But in fact, the following (surely folklore-type) result shows that when *proving* Lemma 7.1.4, it suffices (for [$J2$]–[$J4$]) to obtain "decay bounds" *without regard to regularity.*

Proposition 7.1.7. *Given an* arbitrary *function $f\colon \mathbb{R}^d \to \mathbb{R}_{\geq 0}$ such that $f(\boldsymbol{a}) \ll_b (1+\|\boldsymbol{a}\|)^{-b}$ (for all $b \in \mathbb{Z}_{\geq 0}$), there must exist a decay weight $\nu_d > 0$ with $f \leq \nu_d$.*

Proof. See [Wan21a, Proposition 4.7]. □

Given Proposition 7.1.7, let us end by summarizing some of the main ideas behind Lemma 7.1.4; see [Wan21a, Appendix B] for details.

Partial proof sketch for Lemma 7.1.4. [$J1$]: See Proposition 3.1.6.

[$J3$]: This is based on a recursive structure proven by an integration by parts argument of [HB96] (using the homogeneity of $F = 0$).

[$J2$]: This involves a lot of stationary phase; the assumption that (F, w) is clean makes the analysis much cleaner (allowing conceptual arguments based on gradient descent, for instance). Let $r := n/Y$ and $\boldsymbol{v} := X\boldsymbol{c}/n$; then it suffices to bound, suitably uniformly over Schwartz functions q, ϕ, the $(q \circ r, \phi)_{r,\boldsymbol{v}}$*-integral*

$$\int_{u\in\mathbb{R}} du\, q(ru) \int_{\boldsymbol{x}\in\mathbb{R}^m} d\boldsymbol{x}\, \phi(\boldsymbol{x})\, \underbrace{e(uF(\boldsymbol{x}) - \boldsymbol{v}\cdot\boldsymbol{x})}_{e(\psi(\boldsymbol{x}))},$$

where $\operatorname{Supp}\phi \subseteq \operatorname{Supp} w$. If $\nabla\psi(\boldsymbol{x}) = u\nabla F(\boldsymbol{x}) - \boldsymbol{v} \gg |u| + \|\boldsymbol{v}\| \gg 1$, then the principle of *non-stationary phase* suffices. In many other ranges, the trivial bound suffices. The critical range occurs when $|u| \asymp_{F,w} \|\boldsymbol{v}\| \gg 1$, with $\boldsymbol{x} \in \operatorname{Supp} w$ restricted to $\|\nabla\psi(\boldsymbol{x})\| \le \delta|u|$. In this setting, fix $u, \boldsymbol{v}$. Locally over $\boldsymbol{x}$, there exists a unique $\boldsymbol{s} \approx \boldsymbol{x}$ with $\nabla\psi(\boldsymbol{s}) = \boldsymbol{0}$. In fact $\boldsymbol{s}$ is *Morse*, i.e. if $\boldsymbol{y} = \boldsymbol{x} - \boldsymbol{s}$ then $\psi(\boldsymbol{x}) - \psi(\boldsymbol{s}) \approx uQ_{\boldsymbol{s}}(\boldsymbol{y}) \in \operatorname{Sym}^2(\mathbb{R}^m)^\vee$ looks non-degenerate. Then, by stationary phase expansion,

$$\begin{aligned}\int_{\boldsymbol{x}\in\mathbb{R}^m} d\boldsymbol{x}\ \phi(\boldsymbol{x})e(\psi(\boldsymbol{x})) &\approx e(\psi(\boldsymbol{s})) \cdot \int_{\boldsymbol{y}\in\mathbb{R}^m} d\boldsymbol{y}\ \phi(\boldsymbol{s}+\boldsymbol{y})e(uQ(\boldsymbol{y})) \\ &\approx e(\psi(\boldsymbol{s})) \cdot \frac{i^{\operatorname{sgn}(uQ)/2}\phi(\boldsymbol{s})}{|2u|^{m/2}|\det Q|^{1/2}},\end{aligned}$$

roughly speaking. (One heuristic for the growth rate $|u|^{-m/2}$ as $|u| \to \infty$ is to compare the integral over $\boldsymbol{y}$ with the discrete Gauss-like sum $|u|^{-m}\sum_{\mathbb{Z}^m \ni u\boldsymbol{y} \ll u} e(Q(u\boldsymbol{y})/u)$, which one might expect to exhibit square-root cancellation over $\mathbb{Z}^m \ni u\boldsymbol{y} \ll u$.) Hooley and Heath-Brown integrated $|u|^{-m/2}$ absolutely over $|u| \asymp \|\boldsymbol{v}\|$. But there is really a factor of $e(\psi(\boldsymbol{s}))$ in front, so by a calculation using non-stationary phase over $|u| \asymp \|\boldsymbol{v}\|$ (as $\boldsymbol{s}$ varies with u), decay occurs if $|F^\vee(\tilde{\boldsymbol{v}})\boldsymbol{v}| \gg 1$ is large. $\square$

7.2 Pointwise sum estimates

Assume F is diagonal and $m \ge 4$. Write $F = F_1x_1^3 + \cdots + F_mx_m^3$, with $F_1, \ldots, F_m \in \mathbb{Z} \setminus \{0\}$. For technical convenience and discussion, we make the following definition:

Definition 7.2.1. For each nonempty set $U \subseteq [m]$, choose $F_U^\vee \in \mathbb{Q}^\times \cdot \operatorname{rad}(F^\vee|_{c_j=0,\, j\notin U})$ with $F_U^\vee \in \mathbb{Z}[c_i]_{i\in U}$.

Example 7.2.2. If $F = x_1^3 + \cdots + x_m^3$, then $F_{\{1,2\}}^\vee \propto \prod(c_1^{3/2} \pm c_2^{3/2}) = c_1^3 - c_2^3$, while $F_{\{1\}}^\vee \propto c_1$. (Cf. the factorization (6.1) in §6.3.2.)

Recall the general pointwise bound Proposition 3.3.3 on $\widetilde{S}_{\boldsymbol{c}}(n)$. By taking $F^\vee(\boldsymbol{c})$ into account, we can sometimes do better. For convenience, let $\boldsymbol{F} := (F_1, \ldots, F_m)$ and $\operatorname{lcm}(\boldsymbol{F}) := \operatorname{lcm}(F_1, \ldots, F_m)$. The following is essentially [Wan21a, Lemma 4.11]:

Lemma 7.2.3. *Assume* $\operatorname{lcm}(\boldsymbol{F})$ *is cube-free.*[1] *Up to scaling* $F^\vee, F_U^\vee$ *once and for all, the following* vanishing and boundedness criteria *hold uniformly over* p*:*

(1) If $\boldsymbol{c} \in \mathbb{Z}_p^m$ *and* $v_p(F^\vee(\boldsymbol{c})) \le 1$*, then* $|\widetilde{S}_{\boldsymbol{c}}(p)| \ll_F 1$*.*

(2) If $\boldsymbol{c} \in \mathbb{Z}_p^m$ *and* $v_p(F^\vee(\boldsymbol{c})) \le 1$*, then* $|\widetilde{S}_{\boldsymbol{c}}(p^2)| \ll_F 1$*.*

(3) If $\boldsymbol{c} \in \mathbb{Z}_p^m$*, then* $S_{\boldsymbol{c}}(p^l) = 0$ *for all integers* $l \ge 2 + v_p(F^\vee(\boldsymbol{c}))$*.*

(4) If $U \subseteq [m]$ *and* $|U| \ge 2$*, then (2)–(3) hold for* $\boldsymbol{c} \in \mathbb{Z}_p^U \times \{0\}^{[m]\setminus U}$, even if *we "replace"* $F^\vee$ *with* $F_U^\vee \in \mathbb{Z}[\boldsymbol{c}|_U]$*.*

[1] This assumption can surely be removed with more work, but it is extremely mild anyways.

(5) *If $i \in [m]$ and $U := \{i\}$, then (2)–(3) hold for $\boldsymbol{c} \in \mathbb{Z}_p^U \times \{0\}^{[m]\setminus U}$,* even if *we "replace" $v_p(F^\vee(\boldsymbol{c}))$ with $\frac{3}{2}v_p(F_U^\vee(c_i)) - \frac{1}{2} \cdot \mathbf{1}_{l=3}$. (Here, when "modifying" (2), we let $l := 2$, so that $\mathbf{1}_{l=3} = 0$. The $\mathbf{1}_{l=3}$ only plays a role when "modifying" (3).)*

Proof sketch. For (1), use Theorem 5.1.14(1) and [PS20, Theorem 1.1].

The proofs of (2)–(5), especially (3)–(5), are much more technical; the main ingredient is Theorem 7.2.5 below, proven in [Wan21a, Appendix D]. Essentially one does a lot of Hensel lifting, and needs to design a suitable "step-by-step ladder" into increasingly deep "approximate singularities" of $(V_{\mathbb{Q}_p})_{\boldsymbol{c}}$. Unfortunately, despite the simplicity of the idea, the full argument involves many clumsy preparations (designed to align ourselves in an "efficient direction" at each "ladder step"). □

Remark 7.2.4. If F were quadratic (rather than cubic), with $m \geq 4$ *even*, then Lemma 7.2.3(1) would be false. This is because in general, if $p \neq 2$, then every rank-r quadratic form $Q/\mathbb{F}_p$ in $s \geq 3$ variables has $\#Q(\mathbb{F}_p) = \#\mathbb{P}^{s-2}(\mathbb{F}_p) \pm p^{(s-r)/2}\mathbf{1}_{2|r} \cdot p^{(s-2)/2}$; cf. Proposition 5.1.21.

On a first reading of the following statement, it may help to focus on the special but "general" case $v_p(\operatorname{lcm}(6\boldsymbol{F})) = v_p(\boldsymbol{c}) = 0$.

Theorem 7.2.5 ([Wan21a, Theorem D.1])**.** *Fix $o \in \{0,1\}$. Fix $\boldsymbol{c} \in \mathbb{Z}_p^m$ and $l \in \mathbb{Z}$. Assume the following:*

(1) *$v_p(\operatorname{lcm}(\boldsymbol{F})) \leq 2$;*

(2) *$l \geq 2 + 10v_p(\operatorname{lcm}(3\boldsymbol{F})) + 8v_p(4) + \frac{3}{2}v_p(\boldsymbol{c})$; and*

(3) *$l - 2 - v_p(48) + o \geq v_\wp(F(\boldsymbol{z}))$ for all tuples $(\lambda, K, \boldsymbol{z}, v_\wp)$ consisting of a unit $\lambda \in \mathbb{Z}_p^\times$, a finite extension $K/\mathbb{Q}_p$, a solution $\boldsymbol{z} \in K^m$ to $\nabla F(\boldsymbol{z}) = \lambda\boldsymbol{c}$, and a valuation $v_\wp \colon K^\times \to \mathbb{Q}$ normalized by $v_\wp(p) = v_p(p) = 1$.*

Then $\mathbf{1}_{o=0} \cdot S_{\boldsymbol{c}}(p^l) = 0$. Also, $\mathbf{1}_{(o,l)=(1,2)} \cdot |S_{\boldsymbol{c}}(p^l)| \leq 2^{m-1}p^{l(1+m)/2}$.

Remark 7.2.6. Theorem 7.2.5 simplifies greatly when F is $\mathbb{P}^{m-1}_{\mathbb{F}_p}$-smooth, i.e. $p \nmid 3\operatorname{lcm}(\boldsymbol{F})$. *If* the simplified statement extends directly (without change) to general $\mathbb{P}^{m-1}_{\mathbb{F}_p}$-smooth cubic forms $F/\mathbb{Z}_p$, a different proof would be needed (maybe less explicit, and more geometric or model-theoretic). It might then take even more work to find a reasonable statement (and proof) for general $\mathbb{P}^{m-1}_{\mathbb{Q}_p}$-smooth cubic forms $F/\mathbb{Z}_p$.

Remark 7.2.7. If one strengthened the assumption "$l \geq v_\wp(F(\boldsymbol{z})) + O(1)$" to "$l \geq 2v_\wp(F(\boldsymbol{z})) + O(1)$", then Theorem 7.2.5 would likely be (i) much easier to prove: it might suffice to use [Wan21a, Proposition D.21], with $d = \frac{1}{2}l - O(1)$; and (ii) still satisfactory (for our purposes), when complemented with suitable results for $l \leq O(1)$; but (iii) messier, and less enlightening.

Remark 7.2.8. One could likely adapt the proof of Theorem 7.2.5 (see [Wan21a, Appendix D]) to obtain a fairly efficient algorithm to compute $(S_{\boldsymbol{c}}(p^l))_{l\geq 2}$, at least when $p \nmid 6\operatorname{lcm}(\boldsymbol{F})$.[2] Such work might also eventually help to improve, or at least clarify, existing bounds on $S_{\boldsymbol{c}}(p^l)$ (e.g. those from [Hoo86b, Hoo97, HB98] that "resort to diagonality"), either for diagonal F or in general.

[2]It would be nice to numerically verify Theorem 7.2.5 (or at least some of its proof ingredients).

Remark 7.2.9. A similar result should hold without the assumption $v_p(\mathrm{lcm}(\boldsymbol{F})) \le 2$; see [Wan21a, Remark D.12] for a possible first step towards such an extension (under the present approach). In the same vein, there might be a variant of Theorem 7.2.5 suited for weak-approximation questions. But for practical reasons, we have restricted ourselves to Theorem 7.2.5.

The assumption $v_p(\mathrm{lcm}(\boldsymbol{F})) \le 2$ essentially lets us "cleanly anchor ourselves" to certain "minimally degenerate" indices. The relevant "clean initial combinatorics" is captured by Observation 7.2.11 below.

Definition 7.2.10. Given $\boldsymbol{c} \in \mathbb{Z}_p^m$, let $R_{\min} := \arg\min_{i\in[m]} v_p(c_i^3/F_i)$, and if $m \in R_{\min}$, then let $U_{\min} := R_{\min} \setminus \{m\}$.

Observation 7.2.11. Fix $\boldsymbol{c} \in \mathbb{Z}_p^m$. After permuting $[m]$ if necessary, *assume* $v_p(c_1^3/F_1) \ge \cdots \ge v_p(c_m^3/F_m)$. Now fix $i, j \in [m]$ with $i \ge j$. Then $v_p(\mathrm{lcm}(\boldsymbol{F})) \le 2$ implies

(1) $v_p(c_i) \ge v_p(c_j)$;

(2) $U_{\min} = \{k \in [m-1] : (v_p(c_k), v_p(F_k)) = (v_p(c_m), v_p(F_m))\}$; and

(3) if $v_p(c_i) - v_p(c_j) \in \{0\} \cup [2, \infty]$ or $2 \mid v_p(c_i/F_i) - v_p(c_j/F_j)$, then $v_p(c_i/F_i) \ge v_p(c_j/F_j)$.

Proof. Suppose we write the integers $2 + v_p(c_i^3/F_i) = 3v_p(c_i) + [2 - v_p(F_i)]$ in base 3, noting that $2 - v_p(F_i) \in \{0, 1, 2\}$. Then we find that "$v_p(c_i^3/F_i) \ge v_p(c_j^3/F_j)$" is equivalent to "$(v_p(c_i), -v_p(F_i)) \ge (v_p(c_j), -v_p(F_j))$ under the lexicographic ordering of $\mathbb{Z}^2$".

Now fix i, j with $i \ge j$. The lexicographic ordering immediately implies (1)–(2), as well as the case "$v_p(c_i) = v_p(c_j)$" of (3). Now assume $v_p(c_i) \ne v_p(c_j)$. Then $v_p(c_i/F_i) - v_p(c_j/F_j) \ge v_p(c_i) - v_p(c_j) - 2 \ge -1 + \mathbf{1}_{v_p(c_i) \ge v_p(c_j)+2}$, which yields the remaining cases of (3). $\square$

7.3 Conditional average sum estimates

Let us state some quantitative forms of the Square-free Sieve Conjecture (cf. [Mil04, p. 956] and [Gra98, Poo03, Bha14])—restricted to a certain range (specified in terms of a real exponent $b \ge 0$)—regarding "unlikely divisors" of certain polynomial outputs. Function field analogs of these statements (for arbitrary $b \ge 0$) are, at least up to literature gaps, known unconditionally; see [Poo03, Lemma 7.1] for a general qualitative statement for multivariate polynomials, and [Lan15, Propositions 3.2 and 3.3] for a quantification for univariate polynomials (which can surely be extended to multivariate polynomials by adapting [Bha14, Theorem 3.3] to the function field setting).

Conjecture 7.3.1 ($\mathrm{SFSC}_{p,b}$)**.** *There exists an absolute $\delta > 0$ such that*

$$\#\{\boldsymbol{c} \in \mathbb{Z}^m \cap [-Z, Z]^m : \exists \text{ a prime } p \in \mathbb{Z} \cap [P, 2P] \text{ with } p^2 \mid F^\vee(\boldsymbol{c})\} \ll_{F^\vee} Z^m P^{-\delta}$$

holds uniformly over $Z \ge 1$ and $P \le Z^{b/2}$.

Conjecture 7.3.2 ($\mathrm{SFSC}_{q,b}$)**.** *There exists an absolute $\delta > 0$ such that*

$$\#\{\boldsymbol{c} \in \mathbb{Z}^m \cap [-Z, Z]^m : \exists \text{ a square-full } q \in \mathbb{Z} \cap [Q, 2Q] \text{ with } q \mid F^\vee(\boldsymbol{c})\} \ll_{F^\vee} Z^m Q^{-\delta}$$

holds uniformly over $Z \ge 1$ and $Q \le Z^b$.

Remark 7.3.3. Note that ($\mathrm{SFSC}_{p,b}$) is stable under scaling $F^\vee$. Also, if ($\mathrm{SFSC}_{p,b}$) holds for $P \le Z^{b/2}$, then it extends (up to modified implied constant) to any *fixed* range of the form $P \ll Z^{b/2}$. Similar remarks hold for ($\mathrm{SFSC}_{q,b}$).

Remark 7.3.4. It can (probably) be shown that ($\mathrm{SFSC}_{p,b}$) implies ($\mathrm{SFSC}_{q,b}$), up to changing δ. The point is that (in the difficult case $b > 1$) a square-full integer $q \in [Z, Z^b]$ is divisible by a square $q' \in [Z^{2/3}, Z^b]$, which in turn either has a prime-squared factor $p^2 \in [Z^{1/3}, Z^b]$, or a square factor $d^2 \in [Z^{1/3}, Z^{2/3}]$ (with d composite).

In view of Lemma 7.2.3(4)–(5), it is also natural to consider the following augmented versions of ($\mathrm{SFSC}_{p,b}$) and ($\mathrm{SFSC}_{q,b}$).

Conjecture 7.3.5 ($\mathrm{SFSC}_{p,b}$+)**.** *If $U \subseteq [m]$ is nonempty, then the analog of ($\mathrm{SFSC}_{p,b}$) holds for $F_U^\vee$ over $(c_i)_{i\in U} \in \mathbb{Z}^U \cap [-Z, Z]^U$. In particular, ($\mathrm{SFSC}_{p,b}$) holds.*

Conjecture 7.3.6 ($\mathrm{SFSC}_{q,b}$+)**.** *If $U \subseteq [m]$ is nonempty, then the analog of ($\mathrm{SFSC}_{q,b}$) holds for $F_U^\vee$ over $(c_i)_{i\in U} \in \mathbb{Z}^U \cap [-Z, Z]^U$. In particular, ($\mathrm{SFSC}_{q,b}$) holds.*

Proposition 3.3.3, Lemma 7.2.3, and ($\mathrm{SFSC}_{q,6}$+) together imply the following statement, at least when F is diagonal and $\operatorname{lcm}(\boldsymbol{F})$ is cube-free:

Conjecture 7.3.7 (B3)**.** *Let $F/\mathbb{Z}$ be a $\mathbb{P}^{m-1}_{\mathbb{Q}}$-smooth cubic form in $m \ge 4$ variables. Restrict to $F^\vee(\boldsymbol{c}) \ne 0$. Then there exists an absolute $\delta > 0$ such that uniformly over $Z > 0$ and $N \le Z^3$, we have*

$$\sum_{\boldsymbol{c}\in\mathbb{Z}^m}{}' \mathbf{1}_{\|\boldsymbol{c}\|\le Z} \left(\sum_{n_{\boldsymbol{c}}\in[N,2N]} \frac{|\widetilde{S}_{\boldsymbol{c}}(n_{\boldsymbol{c}})|}{N^{1/2}} \cdot \mathbf{1}_{n_{\boldsymbol{c}} | F^\vee(\boldsymbol{c})^\infty} \right)^2 \ll_F Z^m N^{-\delta}.$$

For the (conditional) proof, see [Wan21a, Lemma 7.43]; in certain key ranges, the N-power saving in (B3) comes from Lemma 7.2.3 and ($\mathrm{SFSC}_{q,6}$+). For a discussion of our reliance on ($\mathrm{SFSC}_{q,6}$+), see e.g. Remark 7.3.13 below.

Remark 7.3.8. "B" refers to the "badness" of $n_{\boldsymbol{c}}$ and $S_{\boldsymbol{c}}(n_{\boldsymbol{c}})$; cf. §4.1's [B2'], as discussed in Remark 4.1.3. For each fixed $\boldsymbol{c}$, the sum over $n_{\boldsymbol{c}}$ is quite sparse, so there is not much difference between $\sum$ and max, but for small N the statement with $\sum$ rather than max is slightly stronger.

Remark 7.3.9. If F were *quadratic*, then (B3) would be false in general. One way to see this is to take N to be a small power of Z, restrict $n_{\boldsymbol{c}}$ to be prime, and apply Remark 7.2.4.

Remark 7.3.10. For us, the second moment—as stated above—is most convenient.[3] However, the first moment—or anything higher—could still be useful. Also note that while a bound of $Z^{m+\epsilon}N^\epsilon$ (in place of $Z^m N^{-\delta}$ in (B3)) follows from Proposition 3.3.3 for diagonal F (cf. [B2'] in §4.1), no comparable result is known for general F (a significant source of difficulty in work such as [Hoo14]); see the discussion in [Wan21c, Appendix A] for more details.

[3]Note that once we have a power saving for the second moment, we also have—due to the trivial bound on $S_{\boldsymbol{c}}$—a power saving for all moments with exponent in $[2, 2 + \Omega(\delta)]$.

Remark 7.3.11. It would be good to understand the restriction $N \leq Z^3$ at a deeper level; is it really needed? Our conditional *proof* of (B3) morally only needs the restriction over the locus $c_1 \cdots c_m = 0$. (However, in the proof, the restriction $N \leq Z^3$ does help at least at a technical level, by letting us get away with the limited range of ($\mathrm{SFSC}_{q,6}$+).) The full *truth* seems unclear.

Remark 7.3.12. In principle, in our *applications* of (B3), it would probably be OK to "drop" the absolute value from $|\widetilde{S}_{\boldsymbol{c}}(n_{\boldsymbol{c}})|$—provided we assume the corresponding statement holds for all sub-intervals of $[N, 2N]$, and not just $[N, 2N]$ itself.

Now we discuss some possible approaches to weakening the hypothesis ($\mathrm{SFSC}_{q,6}$+).

Remark 7.3.13. I expect that if instead of Lemma 7.2.3(1), one uses Theorem 5.1.14(2), then a quantitative forms of Ekedahl's geometric sieve (cf. [Bha14, Theorem 3.3]) would allow one to replace the ingredient ($\mathrm{SFSC}_{q,6}$+) with ($\mathrm{SFSC}_{q,3}$+) (and thus with ($\mathrm{SFSC}_{p,3}$+)). Furthermore, if one restricts (B3) to $F^\vee(\boldsymbol{c}) \cdot c_1 \cdots c_m \neq 0$ and not just $F^\vee(\boldsymbol{c}) \neq 0$, then I believe ($\mathrm{SFSC}_{q,6}$) (and thus ($\mathrm{SFSC}_{q,3}$) or ($\mathrm{SFSC}_{p,3}$)) would suffice; in fact, this further-restricted version of (B3), say (B3G), would suffice for the main goals of Chapter 8.

Remark 7.3.14. Going beyond Remark 7.3.13, say one uses Theorem 7.2.5 (or any improvement thereof) in place of Lemma 7.2.3(2)–(3). Then ultimately, one might be left with problems such as Problem 7.3.15 below (to give a toy example).

Problem 7.3.15. Show that uniformly over $Z \geq 1$ and $Q \leq Z^3$, the count

$$\#\{\boldsymbol{c} \ll Z : \exists\, p \asymp Q^{1/2} \text{ and } (\boldsymbol{x}, \lambda) \in \mathbb{Z}_p^m \times \mathbb{Z}_p^\times \text{ with } \nabla F(\boldsymbol{x}) = \lambda \boldsymbol{c} \text{ and } p^2 \mid F(\boldsymbol{x})\}$$

is $\ll_F Z^m Q^{-\delta}$. (Here $\boldsymbol{c}$ denotes a tuple in $\mathbb{Z}^m$, and p denotes a prime.)

Remark 7.3.16. While a direct use of Theorem 7.2.5 (with $(o, l) = (1, 2)$) to control large values of $\widetilde{S}_{\boldsymbol{c}}(p^2)$ only "guarantees" $\boldsymbol{x} \in \overline{\mathbb{Q}}_p^m$, a closer inspection of the proof of Theorem 7.2.5 shows, under mild hypotheses, that $S_{\boldsymbol{c}}(p^2) = 0$ unless there exists $(\boldsymbol{x}, \lambda) \in \mathbb{Z}_p^m \times \mathbb{Z}_p^\times$ with $\nabla F(\boldsymbol{x}) \equiv \lambda \boldsymbol{c} \bmod p$—a condition not so far from the toy problem's condition $\nabla F(\boldsymbol{x}) = \lambda \boldsymbol{c}$. (With more care when $p \mid c_1 \cdots c_m$, we could state a more precise toy problem—but at least when $Q \gg Z^2$ is sufficiently large, this would not make a difference.)

Remark 7.3.17. The toy problem above vaguely resembles [Kow21], but with an additional "multi-quadratic cover" aspect due to ∇F. Optimistically, one might try to combine [Kow21] with techniques for "Weyl sums for square roots" developed by [DFI12] and others.

Remark 7.3.18. As another approach to simplifying the work or input required, one might try combining Jutila's "approximate circle method" with [DFI93, HB96], in the hope of "reducing" the delta method to "simple" moduli n of our choice (e.g. square-free n coprime to $6F_1 \cdots F_m$). This may or may not be possible; the inherent "error bound" in Jutila's circle method may or may not be too large for our purposes in [Wan21d, Wan21a].

Chapter 8

Using the Ratios Conjectures

8.1 Introduction

Let $F, V, w, \ldots$ be as in §3.1. Assume $m = 6$ and $F = x_1^3 + \cdots + x_6^3$, as in §4.1. Under certain standard number-theoretic hypotheses, the paper [Wan21a]—to be summarized in the present chapter—"removes the ϵ from §4.1" (and goes beyond it, proving HLH when (F, w) is clean); this has statistical implications for sums of three cubes, by Observation 2.1.3 and Theorem 2.1.8. The following hypotheses suffice:

(1) some Langlands-type conjectures (applied to certain Galois representations "coming from geometry"), plus GRH (for certain automorphic L-functions);

(2) some RMT-type predictions (based on [CFZ08, §5.1, (5.6)] and [SST16])—namely the L-function Ratios Conjectures over the *multi-parameter* geometric family $\boldsymbol{c} \mapsto L(s, V_{\boldsymbol{c}})$ (with $L(s, V_{\boldsymbol{c}})$ defined as in Example 3.2.10 and §4.1);

(3) a quantitative form of the Square-free Sieve Conjecture (see §7.3); and

(4) the constancy of the local Hasse–Weil Euler factors $L_p(s, V_{\boldsymbol{c}})$, when $\boldsymbol{c}$ is restricted to a p-adic residue class of modulus $p \cdot \gcd(F^\vee(\boldsymbol{c})^{O(1)}, p^\infty)$—in the spirit of Krasner's lemma (cf. the general but possibly "ineffective" results of [Kis99]).

Remark 8.1.1. Under the current state of knowledge, we expect (1)–(2) to be the most serious hypotheses; (4) to be provable, and thus removable, by a suitable expert; and (3) to be somewhere in between.

In §8.2, we will state more precise hypotheses, especially regarding (1)–(2). For now, we make six (inessential) remarks on hypotheses (1)–(4).

Remark 8.1.2. Very roughly, to improve on §4.1, the paper [Wan21a] "factors" certain key sums into pieces addressed by (2), from the world of "RMT-based heuristics", or by (3), from the world of "unlikely divisors" (see especially Conjecture 7.3.7 (B3) and the surrounding discussion).

Remark 8.1.3. It might be worthwhile to find a suitable "elementary" replacement for the Ratios Conjectures whose formulation would not require analytic continuation or GRH. But [CFZ08, §5.1, (5.6)] itself is stated directly in terms of L and $1/L$.

Remark 8.1.4. In (1)—unlike in [Hoo86b, Hoo97, HB98]'s "Hypothesis HW"—we need to know not just about $L(s, V_{\boldsymbol{c}})$ as $\boldsymbol{c}$ varies, but also about $L(s, V)$, and about $L(s, V_{\boldsymbol{c}}, \bigwedge^2)$ and its poles as $\boldsymbol{c}$ varies. For us, such polar information comes from Langlands, via standard ideas of a representation-theoretic flavor.

Remark 8.1.5. In (2), we take the family of $V_{\boldsymbol{c}}$'s over $\{\boldsymbol{c} \in \mathbb{Z}^6 : F^\vee(\boldsymbol{c}) \neq 0\}$, indexed by very clean level sets $\|\boldsymbol{c}\| = *$ for the "ϵ-removal" result, but later indexed with an additional "adelic perturbation" for the HLH result.

Remark 8.1.6. In work towards "geometric" RMT conjectures, [Mil04] used a similar—but more qualitative—version of (3), for sieve-theoretic purposes *different from ours*. That said, it might be interesting to interpret [DFI93, HB96]'s delta method in a sieve-theoretic framework (cf. the fact that 0 is the only integer divisible by arbitrarily large primes; also cf. the "philosophy of sieves and assemblers" suggested by [MM21]).

Remark 8.1.7. (4) is a technical statement that we use first to prove the existence of certain local averages (such as those required for the "recipe for the Ratios Conjectures"), and second to pacify certain "small bad error moduli" for HLH.

8.2 Precise statements of our main hypotheses

(Recall that we are assuming $F = x_1^3 + \cdots + x_6^3$. But in general, for F with nonzero discriminant, almost all of what we say in §8.2 will apply if $m = 6$; most of it will apply if $m \geq 3$ and $2 \mid m$; and some of it will apply if $m \geq 3$ and $2 \nmid m$.)

Having already stated the necessary forms of the Square-free Sieve Conjecture (SFSC) in §7.3, it remains to discuss our other hypotheses, which involve certain "geometric" L-functions. This is technical, but the following three comments may help.

(1) Geometric L-functions capture "the horizontal variation of local behavior" (i.e. variation in p), in a way susceptible to the familiar formalism of representation theory.

(2) At least at "good" primes p, one can compute all local data in question *cleanly and concretely* in terms of certain Frobenius eigenvalues. (See e.g. Definition 3.2.7 for a friendly definition of the specific L-functions $L(s, V_{\boldsymbol{c}})$ at "good" primes. For the connection to Definition 8.2.1 below, see Observation 8.2.14 near p. 80.)

(3) At least morally, the L-functions relevant to us are no more complicated than the L-functions $L(s, A)$ of *abelian varieties* A, and various tensor-square L-functions thereof. Therefore, familiarity with fairly well-known Galois representations (e.g. ℓ-adic Tate modules $V_\ell A := T_\ell A \otimes \mathbb{Q}$, and various tensor squares thereof) basically suffices.

Definition 8.2.1 (Cf. [Tay04])**.** Fix, once and for all, a prime ℓ_0 and an inclusion $\iota\colon \mathbb{Q}_{\ell_0} \hookrightarrow \mathbb{C}$. Let $\rho\colon G_{\mathbb{Q}} \to \mathrm{GL}(M)$ denote an ℓ_0*-adic* representation of $G_{\mathbb{Q}}$.

Suppose ρ *arises from geometry*, in the "arbitrary subquotient" sense of [Tay04, pp. 79–80, the two paragraphs before Conjecture 1.1 (Fontaine-Mazur)]. Then ρ is *de Rham* at ℓ_0, and *pure* of some weight $w \in \mathbb{Z}$, so [Tay04, §2, pp. 85–86] defines (among other invariants of ρ)

(1) a global *conductor* $q(\rho)$ and *root number* $\epsilon(\rho) := \epsilon(\iota\rho)$, and

(2) an *algebraically normalized* local factor $L_{v,\mathrm{Taylor}}(s,\iota\rho)$ at each place $v \le \infty$.

Now define the *analytically normalized* local factors $L_v(s,\rho) := L_{v,\mathrm{Taylor}}(s+w/2,\iota\rho)$. Globally, set $L(s,\rho) := \prod_{p<\infty} L_p(s,\rho)$ and $\Lambda := L_\infty L$. And for each prime p, let $\{\tilde{\alpha}_{\rho,j}(p)\}_j := \{p^{-w/2}\alpha_{\rho,j}(p)\}_j$ denote the multiset of *normalized* eigenvalues of geometric Frobenius on the ($\le \dim(M)$-dimensional) representation of $W_{\mathbb{Q}_p}/I_{\mathbb{Q}_p}$ defining $L_p(s,\rho)$. Finally, let $\tilde{\lambda}_\rho(n) := [n^s]L(s,\rho)$ denote the nth coefficient of $L(s,\rho)$ for $n \ge 1$.

For a smooth projective variety Y over a field k, let $H^d(Y) := H^d(Y_{\overline{k}}, \mathbb{Q}_{\ell_0})$; for an *embedded* smooth projective variety $Z \subseteq \mathbb{P}^N_k$, let $H^d_{\mathrm{diff}}(Z) := H^d(Z)/H^d(\mathbb{P}^N_k)$. When $M := H^d_{\mathrm{diff}}(Z)$ for a *smooth projective complete intersection* $Z \subseteq \mathbb{P}^{r+d}_{\mathbb{Q}}$ of dimension $d \ge 0$ and codimension $r \ge 1$, we *abbreviate* $q(M), L(s,M), \ldots$ as $q(Z), L(s,Z), \ldots$. (Here if $2 \nmid d$, then in fact $H^d_{\mathrm{diff}}(Z) = H^d(Z)$; but when $2 \mid d$, the "diff" is important.)

Remark 8.2.2. Here each $L_p(s,\rho) = \prod_j (1 - \tilde{\alpha}_{\rho,j}(p)p^{-s})^{-1}$ has some "degree" $\#\{j\} \le \dim M$. *Purity* means "$\#\{j\} = \dim M$ and $|\tilde{\alpha}_{\rho,j}(p)| = 1\ \forall j$" holds for all p outside a *finite* set.

Remark 8.2.3. Say $Z/\mathbb{Q}$ is a smooth projective complete intersection of dimension $d \ge 0$. Then only the *middle cohomology* $H^d(Z)$—or rather, only the "primitive quotient" $H^d_{\mathrm{diff}}(Z)$ thereof—is interesting. This justifies the definition $L(s,Z) := L(s,M)$ above.

Now say $d \ge 1$. Then $H^d(Z) = H^d(\mathbb{P}^{1+d}) \oplus \ker(L\colon H^d(Z) \to H^{d+2}(Z))$, where L denotes the Lefschetz operator. So $H^d_{\mathrm{diff}}(Z) \cong \ker(L)$. But this is almost always false for $d = 0$. So even though $d \ge 1$ always for us, we prefer to define $H^d_{\mathrm{diff}}(Z)$ as in Definition 8.2.1.

Definition 8.2.4. Fix $\boldsymbol{c} \in \mathbb{Z}^m$ with $F^\vee(\boldsymbol{c}) \ne 0$. Then $V_{\boldsymbol{c}}$ is a smooth projective complete intersection in $\mathbb{P}^{m-1}_{\mathbb{Q}}$ of dimension $m_* := m-3$. So in particular, Definition 8.2.1 *fully defines* $L(s,V_{\boldsymbol{c}})$. For $n \ge 1$, now let $\tilde{\lambda}_{\boldsymbol{c}}(n) := [n^s]L(s,V_{\boldsymbol{c}})$ and $\mu_{\boldsymbol{c}}(n) := [n^s]L(s,V_{\boldsymbol{c}})^{-1}$. And for all p,j, let $\tilde{\alpha}_{\boldsymbol{c},j}(p) := \tilde{\alpha}_{M_{\boldsymbol{c}},j}(p)$, where $M_{\boldsymbol{c}} := H^{m_*}_{\mathrm{diff}}(V_{\boldsymbol{c}})$.

Remark 8.2.5. A close reading of [Tay04, §§1–2] (keeping in mind that certain geometric facts are *truly of local origin*, even though [Tay04] often "starts" globally) shows that given n, one can define $\tilde{\lambda}_{\boldsymbol{c}}(n), \mu_{\boldsymbol{c}}(n)$ on all of $\{\boldsymbol{c} \in \prod_{p\mid n} \mathbb{Z}_p^m : F^\vee(\boldsymbol{c}) \ne 0\}$. We could probably avoid using extended definitions like these, but they are convenient in some local calculations.

After (SFSC), our second technical hypothesis, Conjecture 8.2.6, is an "effective Krasner-type lemma" (an "ineffective" version likely already being known by [Kis99]).

Conjecture 8.2.6 (EKL). *Let* $S := \{\boldsymbol{c} \in \widehat{\mathbb{Z}}^m : F^\vee(\boldsymbol{c}) \ne 0\}$. *Then there exists a* nonzero *homogeneous polynomial* $H \in \mathbb{Z}[\boldsymbol{c}]$ *such that* $\tilde{\lambda}_{\boldsymbol{c}}(n)$ *is* constant *on the fibers of the map* $\mathbb{Z}_{\ge 1} \times S \to \mathbb{Z}_{\ge 1} \times \mathbb{Z}_{\ge 0} \times 2^{\widehat{\mathbb{Z}}}$ *given by*

$$(n,\boldsymbol{c}) \mapsto (n, r, \boldsymbol{c} + \mathrm{rad}(n)r\widehat{\mathbb{Z}}), \text{ where } r := (n^\infty, H(\boldsymbol{c})) \in \mathbb{Z}_{\ge 0}.$$

Remark 8.2.7. The point is that $L_p(s,V_{\boldsymbol{c}})$ should morally be computable from something like a "local minimal model" of $(V_{\boldsymbol{c}})_{\mathbb{Q}_p}$. At least for elliptic curves in Weierstrass form, $v_p(\Delta)$ controls minimality, and L_p is determined by the coefficients modulo p of a minimal model.

Question 8.2.8. Does (EKL) hold with $H = \Theta_F(1) \cdot F^\vee$ (for a suitable constant $\Theta_F(1)$)?

Now we move on to our seemingly more fundamental hypotheses—those involving (for the most part) *global aspects of L-functions.* As suggested in Remark 3.2.12, we will phrase these hypotheses in terms of automorphic representations π of $\mathrm{GL}_\bullet$ over $\mathbb{Q}$. Doing so is technical, but the following remark (partly based on the survey [FPRS19]) may help.

Remark 8.2.9. We will only work with *cuspidal* π's (on $\mathrm{GL}_\bullet$ over $\mathbb{Q}$), or more generally, *isobaric* π's. These π's have well-defined L-functions $L(s,\pi)$, and good formal properties (due to Rankin, Selberg, Langlands, Godement, Jacquet, Shalika, and others):

(1) If π is cuspidal, then $L(s,\pi)$ is *primitive* in the sense of [FPRS19, Lemma 2.4], and has certain standard analytic properties [FPRS19, Theorem 3.1]. Furthermore, the following statements hold:

  (a) Given a cuspidal π, let $\omega\colon \widehat{\mathbb{Z}}^\times \times \mathbb{R}_{>0} \to \mathbb{C}^\times$ denote the (continuous) central character of π. Then the critical line of $L(s,\pi)$ is $\Re(s) = 1/2$ if and only if ω is *unitary* (in which case one might informally say "π is of weight 0").

  (b) In the setting of (a), suppose ω is unitary. Say π is on GL_d, let $q(\pi)$ be the conductor of π, and write $L_p(s,\pi) = \prod_{1\le j\le d}(1-\alpha_{\pi,j}(p)p^{-s})^{-1}$ for each $p \nmid q(\pi)$. Then the number $\prod_{1\le j\le d}\alpha_{\pi,j}(p) \in \mathbb{C}$ is *algebraic for all* $p \nmid q(\pi)$ if and only if ω is a Dirichlet character (i.e. ω is of *finite order*; i.e. $\omega|_{\mathbb{R}_{>0}} = 1$; i.e. π is *balanced* in the sense of [FPRS19, §2.2]).

(2) For each isobaric π, there is a *unique* multiset $\{\pi_1,\dots,\pi_r\}$, consisting of *cuspidals*, such that $L(s,\pi) = L(s,\pi_1)\cdots L(s,\pi_r)$. We call the π_i's *cuspidal constituents* of π.

(3) *Strong multiplicity one* (a strengthening of (2)'s "uniqueness statement"): If π,π' are isobaric, and $L_p(s,\pi) = L_p(s,\pi')$ for all but finitely many primes p, then $L(s,\pi) = L(s,\pi')$—or equivalently, $\{\pi_1,\dots,\pi_r\} = \{\pi'_1,\dots,\pi'_{r'}\}$, in the notation of (2).

(4) Fix an isobaric π, say corresponding to $\{\pi_1,\dots,\pi_r\}$ in the notation of (2). Suppose each π_i has *unitary* central character. Then $L(s,\pi)$ has *real coefficients* if and only if π is *self-dual* (i.e. $L(s,\pi) = L(s,\pi^\vee)$; i.e. $\{\pi_1,\dots,\pi_r\} = \{\pi_1^\vee,\dots,\pi_r^\vee\}$).

(Confusingly, isobaric π's are always irreducible as automorphic representations. So in (1) it is indeed better to use the adjective "primitive" than "irreducible".)

For our automorphic purposes, the following convenient definition (though not as precise as the "algebraicity" and "arithmeticity" definitions of Clozel and Buzzard–Gee) will suffice:

Definition 8.2.10. Call an isobaric π *nice* if each cuspidal constituent of π has *unitary, finite-order* central character.

Using Definition 8.2.10, we can now state Conjecture 8.2.11, which cleanly extends Hypothesis HW (or rather (1) in Remark 3.2.11) from $L(s,V_{\boldsymbol{c}})$ to a host of related L-functions.

Conjecture 8.2.11 (HW2). *Fix $\boldsymbol{c} \in \mathbb{Z}^m$ with $F^\vee(\boldsymbol{c}) \neq 0$. Let $M_{\boldsymbol{c}} := H^{m_*}_{\mathrm{diff}}(V_{\boldsymbol{c}})$ and $M_V := H^{1+m_*}_{\mathrm{diff}}(V)$. Now let M denote one of the representations*

$$M_{\boldsymbol{c}}, M_V, M_{\boldsymbol{c}} \wedge M_{\boldsymbol{c}}, \mathrm{Sym}^2 M_{\boldsymbol{c}}, M_{\boldsymbol{c}} \otimes M_{\boldsymbol{c}}, M_{\boldsymbol{c}} \otimes M_{\boldsymbol{c}}^\vee.$$

Let S denote $\mathrm{disc}(F)$ *or* $F^\vee(\boldsymbol{c})$*, according as $M = M_V$ or $M \neq M_V$. Then*

(a) *M itself* arises from geometry *(pure of weight m_*, $1 + m_*$, or $2m_*$ according as M is $M_{\boldsymbol{c}}$, M_V, or neither), with* $\dim M = O_m(1)$ *and* $q(M) \mid \operatorname{rad}(S)^{O_m(1)} \mid S^{O_m(1)}$*;*

(b) *the gamma factor $L_\infty(s, M)$ lies in a* finite set *depending only on m;*

(c) *there exists a* nice isobaric π_M *on* $\mathrm{GL}_{\dim M}$ *over* $\mathbb{Q}$ *such that (i) $L_v(s, M) = L_v(s, \pi_M)$ at all places $v \leq \infty$, and (ii) $(q(M), \epsilon(M)) = (q(\pi_M), \epsilon(\pi_M))$;*

(d) *each cuspidal constituent of π_M satisfies the* Generalized Ramanujan Conjectures (GRC)*—whence $|\tilde{\alpha}_{M,j}(p)| \leq 1$ for all p, j, and $L_\infty(s, M)$ is holomorphic on $\{\Re(s) > 0\}$; and*

(e) *$L(s, \pi_M)$ satisfies the* Grand Riemann Hypothesis (GRH)*.*

Remark 8.2.12. Here "HW" stands for "Hasse–Weil" (cf. "Hypothesis HW"). The "2" refers to the "second-order" nature in which $L(s, V), L(s, V_{\boldsymbol{c}}, \bigwedge^2)$ will appear later. Here and later, we let $L(s, V_{\boldsymbol{c}}, \bigwedge^2) := L(s, M_{\boldsymbol{c}} \wedge M_{\boldsymbol{c}})$, and similarly define $L(s, V_{\boldsymbol{c}}, \mathrm{Sym}^2), L(s, V_{\boldsymbol{c}}, \bigotimes^2), \ldots$.

Remark 8.2.13. Some of the "conjectures" above (excluding GRH) may be known in some generality or specificity; see §8.5.1 (which begins near p. 91) for some discussion.

Observation 8.2.14. In the notation of (HW2), fix a prime $p \nmid \ell_0 S$. Then the multiset $\{\alpha_{M,j}(\rho)\}_j$ *coincides* with the multiset of eigenvalues of geometric Frobenius on "the obvious $G_{\mathbb{F}_p}$-representation $\mathcal{M}_{\mathbb{F}_p} \in \{H^{m_*}_{\mathrm{diff}}((\mathcal{V}_{\boldsymbol{c}})_{\mathbb{F}_p}), \ldots\}$ corresponding to M". (For proof, use smooth proper base change—and some linear algebra on $\bigwedge^2, \mathrm{Sym}^2, - \otimes -$.)

In particular, this explains the connection between the (full, sophisticated) Definition 8.2.4 and the (partial, simpler) Definition 3.2.7 for $L(s, V_{\boldsymbol{c}})$.

Observation 8.2.15. Assume (HW2), and fix $\boldsymbol{c}$. Then $L(s, V_{\boldsymbol{c}}, \bigotimes^2)$ is (isobaric) automorphic, so by strong multiplicity one (Remark 8.2.9(3)), *the Rankin–Selberg L-function $L(s, \pi_{M_{\boldsymbol{c}}}, \bigotimes^2)$ is automorphic, and equal to $L(s, V_{\boldsymbol{c}}, \bigotimes^2)$*. Similar comments apply to $L(s, \pi_{M_{\boldsymbol{c}}}, \bigwedge^2), \ldots$.

In particular, $L(s, \pi_{M_{\boldsymbol{c}}} \times \pi^\vee_{M_{\boldsymbol{c}}})$—which has a pole at $s = 1$ for general reasons—must be (isobaric) automorphic, and must satisfy GRH. So *RH for $\zeta(s)$ must hold*, under (HW2).

For convenience, we now refine (HW2) to a conjecture more specific to our family of $M_{\boldsymbol{c}}$'s. From here to the end of §8.2, it will be important that $2 \mid m$.

Conjecture 8.2.16 (HWSp). *Conjecture* (HW2) holds, and *each π_M in (HW2) is* self-dual. Furthermore*, for each $\boldsymbol{c}$ in (HW2), there exists a* nice isobaric $\phi_{\boldsymbol{c},2}$ *over* $\mathbb{Q}$ *such that*

(a) *$L_v(s, V_{\boldsymbol{c}}, \bigwedge^2) = \zeta_v(s) L_v(s, \phi_{\boldsymbol{c},2})$ holds at all places $v \leq \infty$, and*

(b) *the conductors and ϵ-factors match accordingly.*

Remark 8.2.17. Fix $\boldsymbol{c}$, and say $M_{\boldsymbol{c}}$ is *irreducible*. (Under Langlands-type conjectures and GRH, one can prove that $M_{\boldsymbol{c}}$ is *typically* irreducible. See also Proposition A.2.1 for some heuristic evidence for a stronger statement.) Then Schur's lemma and Poincaré duality (and the fact that $m_* = m - 3$ is *odd*) suggest (via Langlands-type conjectures) that the putative $\pi_{M_{\boldsymbol{c}}}$ in (HW2) should be *cuspidal self-dual symplectic* as defined on [SST16, p. 533], and in particular that $L(1, \pi_{M_{\boldsymbol{c}}}, \bigwedge^2) = \infty$. Hence the "Sp" in "HWSp".

It is then no surprise[1] that under (HW2), say, the *homogeneity* type of the family $\boldsymbol{c} \mapsto L(s, V_{\boldsymbol{c}})$ is "symplectic" (in the sense of [SST16]); see §8.5.2 (which begins near p. 95) for details. Under [SST16, Universality Conjecture], one then expects the *symmetry* type to be "orthogonal"—a point that we will clarify in due time (even though the RMT-based conjectures we use will come mostly from [CFZ08] rather than [SST16]).

See §8.5.1 for a more thorough discussion of (HW2) and (HWSp), and where they come from. But most directly relevant to us are the *analytic estimates* (in Remark 8.5.3) implied by (HW2) and (HWSp), leading via [IK04, Perron's formula (5.111)] to *pointwise* "square-root cancellation up to ϵ" bounds (such as (3) in Remark 3.2.11). Such bounds "barely" fail us. Fortunately, there exist more precise *mean-value* predictions of RMT type. For some (but not all) of our purposes here and elsewhere, we can work with *smooth weights*; to this end, recall Definition 7.1.1.

Let $\pi_{\boldsymbol{c}} := \pi_{M_{\boldsymbol{c}}}$. In §8.3.3 (which begins near p. 88), we will explain—assuming (HW2) and (EKL)—how the "recipe" (or "heuristic") of [CFZ08], when interpreted in the general framework of [SST16, pp. 534–535, Geometric Families and Remark (i)], leads to the following clean two-part "Ratios Conjecture" (and to the slightly messier "Ratios Conjecture" (RA1) stated later below as Conjecture 8.2.28):

Conjecture 8.2.18 (Cf. [CFZ08, §5.1, (5.6)])**.** *Restrict to $F^\vee(\boldsymbol{c}) \neq 0$, let $s := \sigma + it$, and* assume (HW2)*. Fix a* decay weight $\nu_m \geq 0$*. Then the following hold:*

(R1) Fix $\sigma > 1/2$. Then uniformly over $Z \geq 1$ and $|t| \leq Z^\hbar$, we have

$$\begin{aligned}&\sideset{}{'}\sum_{\boldsymbol{c}\in\mathbb{Z}^m} \nu_m\left(\frac{\boldsymbol{c}}{Z}\right) \cdot \frac{1}{L(s,\pi_{\boldsymbol{c}})} \\ &= O_{\nu_m,\sigma}((Z^m)^{1-\delta}) + \sideset{}{'}\sum_{\boldsymbol{c}\in\mathbb{Z}^m} \nu_m\left(\frac{\boldsymbol{c}}{Z}\right) \cdot \zeta(2s)L(s+1/2,V)A_F(s),\end{aligned}$$

for some absolute $\hbar, \delta > 0$ independent of σ.

(R2) Fix $\sigma_1, \sigma_2 > 1/2$. Then uniformly over $Z \geq 1$ and $\|\boldsymbol{t}\| \leq Z^\hbar$, we have

$$\begin{aligned}&\sideset{}{'}\sum_{\boldsymbol{c}\in\mathbb{Z}^m} \nu_m\left(\frac{\boldsymbol{c}}{Z}\right) \cdot \frac{1}{L(s_1,\pi_{\boldsymbol{c}})L(s_2,\pi_{\boldsymbol{c}})} \\ &= O_{\nu_m,\boldsymbol{\sigma}}((Z^m)^{1-\delta}) + \sideset{}{'}\sum_{\boldsymbol{c}\in\mathbb{Z}^m} \nu_m\left(\frac{\boldsymbol{c}}{Z}\right) \cdot A_{F,2}(\boldsymbol{s})\zeta(s_1+s_2)\prod_{j\in[2]}\zeta(2s_j)L(s_j+1/2,V),\end{aligned}$$

for some absolute $\hbar, \delta > 0$ independent of $\boldsymbol{\sigma}$.

Here $A_F(s), A_{F,2}(\boldsymbol{s})$ are certain Euler products (defined in §8.3.3, in terms of F), absolutely convergent *on the regions $\Re(s), \Re(\boldsymbol{s}) \geq 1/2 - \delta'$, respectively, for some absolute $\delta' > 0$.*

Before proceeding, we make four remarks on Conjecture 8.2.18.

Remark 8.2.19. The letter "R" signifies "Random Matrix Theory" or "Ratios Conjecture" (or "recipe" thereof).

[1] since we are working over a number field (rather than a global function field)

Remark 8.2.20. Following [CS07, (2.11a)–(2.11b)], one could require $\sigma, \boldsymbol{\sigma} \le 1/2 + \delta'$, and restrict the "absolute convergence" statement to $\Re(s), \Re(\boldsymbol{s}) \in [1/2 - \delta', 1/2 + \delta']$, to be safe (see [CS07, Remark 2.3]). Our applications would certainly permit that. But since our "ratios" only involve $1/L$ and not L, the present formulation of Conjecture 8.2.18 should be OK.

Remark 8.2.21. We restrict t to $[-Z^{\hbar}, Z^{\hbar}]$ to (comfortably) respect the constraint [CS07, (2.11c)] on "vertical shifts". We could alternatively allow *arbitrary* $t \in \mathbb{R}$, but only after including a sufficiently large factor of the form $(1+|t|)^{O(1)}$ in the error terms of (R1)–(R2).

(Here we are being careful, but it might actually be possible to relax [CS07, (2.11c)] to allow $t \in [-Z^{O(1)}, Z^{O(1)}]$; cf. [BC20, p. 4, the sentence before Conjecture 2].)

Remark 8.2.22. [CFZ08, §5.1, (5.6)] has a "square-root error term"—too much to ask for in general.[2] However, all theoretical and numerical data so far suggest *power-saving error terms*—and an *arbitrary* power saving is all we need (and all we assume).

Actually, to prove $N_F(X) \ll X^3$, we do not need (R2) itself, but only the weaker statement (R2') below. However, we need to keep uniformity in mind when integrating against certain *varying weights* $f(s)$. So we first make a definition:

Definition 8.2.23. Let $f\colon \{\Re(s) \in (1/5, 6/5)\} \to \mathbb{C}$ be Schwartz on vertical strips. Given $Z, \nu_1, \hbar$, suppose $|f(s)| \le \nu_1(t)$ for all $\sigma \in (1/5, 6/5)$ and $|t| \ge Z^{\hbar}$. Also let $M_I(f) := \sup_{\sigma \in I} \|f(\sigma + it)\|_{L^1(t \in \mathbb{R})}$ and suppose $M_{(1/5,6/5)}(f) < \infty$. Then we say f is $(Z, \nu_1, \hbar)$-*good.*

Conjecture 8.2.24 (R2'). Assume *(HW2). Then there exist absolute constants* $\hbar, \delta > 0$ *such that if* $\nu_1, \nu_m \ge 0$ *are decay weights and* $f(s)$ *is an* $(Z, \nu_1, \hbar)$-good holomorphic *function, then uniformly over any* given *range of the form* $1 \ll N \ll Z^3$*, we have*

$$\sideset{}{'}\sum_{\boldsymbol{c} \in \mathbb{Z}^m} \nu_m\left(\frac{\boldsymbol{c}}{Z}\right) \left| \int_{(\sigma)} ds\, \frac{\zeta(2s)^{-1} L(s+1/2, V)^{-1}}{L(s, \pi_{\boldsymbol{c}})} \cdot f(s) N^s \right|^2 \ll_{\nu_1, \nu_m} M^{(2)}(f) \cdot Z^m N$$

for all $\sigma > 1/2$*, where the implied constant on the right-hand side depends only on* ν_1, ν_m*. Here* $M^{(2)}(f) := [1 + M_{\{1/2\}}(f^2) + M_{[1/2-\delta, 1/2+\delta]}(s^3 f(s))]^2$.

Remark 8.2.25. In [Wan21a, §7.4], we prove (R2') *assuming* (HW2) and (R2).

Remark 8.2.26. Sometimes, "sharp-up-to-constant upper bounds" like (R2') can be "easier" than true "leading-order asymptotics" like (R1)–(R2). See e.g. [Har13].

On the other hand, (R1) itself *does not suffice* to fully establish HLH (in our setting). Rather, we need a slight "adelic perturbation" (R$\mathbf{A}_{\mathbb{Z}}$1) of (R1). To state the Ratios Conjecture (R$\mathbf{A}_{\mathbb{Z}}$1), we first make a convenient definition:

Definition 8.2.27. Let $r_- := \min(r, 0)$ and $r_+ := \max(r, 0)$ for every $r \in \mathbb{R}$. Then given $\boldsymbol{Z} \in \mathbb{R}^m$, we let $\mathcal{B}(\boldsymbol{Z}) := [(Z_1)_-, (Z_1)_+] \times \cdots \times [(Z_m)_-, (Z_m)_+]$.

[2] See [DW21, Theorem 1.2] for an unconditional example in the setting of [CFK$^+$05]. See also [CFLS22, first two sentences of the paragraph before Theorem 1.4, as $\sigma \to 0^+$] for a presumably unconditional example in the setting of one-level densities.

Conjecture 8.2.28 (RA$_{\mathbb{Z}}$1, or RA1)**.** *Fix $\sigma > 1/2$. Then for some absolute $\hbar, \delta > 0$ independent of σ, we have*

$$\sum_{\boldsymbol{c} \equiv \boldsymbol{a} \bmod n_0}' \mathbf{1}_{\boldsymbol{c} \in \mathcal{B}(\boldsymbol{Z})} \cdot \frac{1}{L(s, \pi_{\boldsymbol{c}})} = O_\sigma(|\mathcal{F}|^{1-\delta}) + \sum_{\boldsymbol{c} \equiv \boldsymbol{a} \bmod n_0}' \mathbf{1}_{\boldsymbol{c} \in \mathcal{B}(\boldsymbol{Z})} \cdot \zeta(2s) L(s+1/2, V) A_F^{\boldsymbol{a}, n_0}(s),$$

uniformly over $|t|, n_0 \leq Z^\hbar$ and $(\boldsymbol{a}, \boldsymbol{Z}) \in \mathbb{Z}^m \times \mathbb{R}^m$ with $|Z_1|, \dots, |Z_m| \in [Z^{1-\hbar}, Z]$. Here $A_F^{\boldsymbol{a}, n_0}(s)$ is a certain Euler product (defined in §8.3.3, in terms of $F, \boldsymbol{a}, n_0$), absolutely convergent on $\{\Re(s) \geq 1/2 - \delta'\}$ for some absolute $\delta' > 0$.

Remark 8.2.29. Note that the "prediction" (i.e. the "right-hand side") is sensitive to both the archimedean data "$\boldsymbol{Z}$" and the non-archimedean data "$\boldsymbol{a} \bmod n_0$". Furthermore, in (RA1), we forbid $\boldsymbol{Z}$ from being too lopsided, and we also forbid n_0 from being too large. For more details on the philosophy behind (RA1), see Remark 8.5.13 (near p. 96).

8.3 Some critical calculations

8.3.1 A second-order approximation

Recall, from Chapter 3 and §4.1, the Dirichlet series $\Phi(\boldsymbol{c}, s), L(s, V_{\boldsymbol{c}})$, and the "first-order approximation" $\Phi = \Psi_1 \Psi_2$, with $\Psi_1 := 1/L$ and $\Psi_2 := \Phi L$. As suggested in Remark 4.1.3, the "first-order error" Ψ_2 is a "source of ϵ" in §4.1, coming from both *good* primes $p \nmid F^\vee(\boldsymbol{c})$ and *bad* primes $p \mid F^\vee(\boldsymbol{c})$.

From the perspective of the good primes—which define "standard" data—we would like a more precise "standard" approximation of Φ. We thus further approximate $\Phi / L(s, V_{\boldsymbol{c}})^{-1}$ by $L(1/2 + s, V)^{-1} L(2s, V_{\boldsymbol{c}}, \bigwedge^2)^{-1}$, in view of the "second-order phenomena"

$$\widetilde{S}_{\boldsymbol{c}}(p) - \mu_{\boldsymbol{c}}(p) = \widetilde{S}_{\boldsymbol{c}}(p) - \widetilde{E}_{\boldsymbol{c}}(p) = -p^{-1/2} \widetilde{E}_F(p) = -p^{-1/2} \tilde{\lambda}_V(p) \quad \text{for } p \nmid F^\vee(\boldsymbol{c})$$

(valid for *even* integers $m \geq 4$) and

$$\widetilde{S}_{\boldsymbol{c}}(p^2) - \mu_{\boldsymbol{c}}(p^2) = 0 - \sum_{i<j} \tilde{\alpha}_{\boldsymbol{c},i}(p) \tilde{\alpha}_{\boldsymbol{c},j}(p) = -\tilde{\lambda}_{V_{\boldsymbol{c}}, \bigwedge^2}(p) \quad \text{for } p \nmid F^\vee(\boldsymbol{c}).$$

Remark 8.3.1. If m were *odd*, we would instead have $\widetilde{S}_{\boldsymbol{c}}(p) - \tilde{\lambda}_{\boldsymbol{c}}(p) = p^{-1/2} \tilde{\lambda}_V(p)$ for $p \nmid F^\vee(\boldsymbol{c})$, and also instead care about the identity $\widetilde{S}_{\boldsymbol{c}}(p^2) - \tilde{\lambda}_{\boldsymbol{c}}(p^2) = -\tilde{\lambda}_{V_{\boldsymbol{c}}, \mathrm{Sym}^2}(p)$ for $p \nmid F^\vee(\boldsymbol{c})$.

For technical reasons, we separately study the "standard Hasse–Weil part" of Φ and the "bad part" of Φ, based on the following definition:

Definition 8.3.2. Given $\boldsymbol{c} \in \mathbb{Z}^m$ with $F^\vee(\boldsymbol{c}) \neq 0$, let $\Phi^{\mathrm{HW}}(\boldsymbol{c}, s) := \prod_{p \nmid F^\vee(\boldsymbol{c})} \Phi_p(\boldsymbol{c}, s)$ and $\Phi^{\mathrm{B}}(\boldsymbol{c}, s) := \prod_{p \mid F^\vee(\boldsymbol{c})} \Phi_p(\boldsymbol{c}, s)$. If $m \geq 3$ is even, define

$$\begin{aligned}
\Phi_1(\boldsymbol{c}, s) &:= L(s, V_{\boldsymbol{c}})^{-1} L(1/2 + s, V)^{-1} \zeta(2s)^{-1} \\
\Phi_2(\boldsymbol{c}, s) &:= \zeta(2s) / L(2s, V_{\boldsymbol{c}}, \textstyle\bigwedge^2) \\
\Phi_3 = \Phi_3^{\mathrm{HW}}(\boldsymbol{c}, s) &:= \Phi^{\mathrm{HW}} \Phi_1^{-1} \Phi_2^{-1} = \Phi^{\mathrm{HW}}(\boldsymbol{c}, s) L(s, V_{\boldsymbol{c}}) L(1/2 + s, V) L(2s, V_{\boldsymbol{c}}, \textstyle\bigwedge^2).
\end{aligned}$$

For each $j \in [3]$, let $a_{\boldsymbol{c},j}(n)$ be the n^s coefficient of $\Phi_j(\boldsymbol{c}, s)$.

Example 8.3.3. $a_{\boldsymbol{c},2}(n) = 0$ if n is not a square.

Remark 8.3.4. Here $\Phi_1(\boldsymbol{c}, s)$ is precisely the "mollified" version of $1/L(s, V_{\boldsymbol{c}})$ considered in Conjecture 8.2.24 (R2')! And $\Phi_2(\boldsymbol{c}, s) = L(2s, \phi_{\boldsymbol{c},2})^{-1}$ under Conjecture 8.2.16 (HWSp).

Remark 8.3.5. For each $j \in [3]$, the definition of $a_{\boldsymbol{c},j}(n)$ can be extended to tuples $\boldsymbol{c} \in \prod_{p|n} \mathbb{Z}_p^m$ with $F^\vee(\boldsymbol{c}) \neq 0$. In local calculations, we sometimes abuse notation accordingly.

Proposition 8.3.6. *Fix $\boldsymbol{c} \in \mathbb{Z}^m$ with $F^\vee(\boldsymbol{c}) \neq 0$, and suppose $m \in \{4, 6\}$.* Assume *Conjecture 8.2.11 (HW2). Then* Φ_3 converges absolutely *over* $\Re(s) > 1/3$*;* locally*, we have, uniformly over $\boldsymbol{c}, p, l$, that*

$$a_{\boldsymbol{c},3}(p) \cdot \mathbf{1}_{p \nmid F^\vee(\boldsymbol{c})} = 0 \quad \text{and} \quad a_{\boldsymbol{c},3}(p^2) \cdot \mathbf{1}_{p \nmid F^\vee(\boldsymbol{c})} \ll p^{-1/2} \quad \text{and} \quad a_{\boldsymbol{c},3}(p^l) \ll_\epsilon p^{l\epsilon}.$$

Also, under *(HWSp),* Φ_2 *is* holomorphic *on the region* $\Re(s) > 1/4$*.*

Proof. Our definition of Φ^{HW}, along with GRC, ensures that each local factor $\Phi_{j,p}(\boldsymbol{c}, s)$ is "well-behaved over $\Re(s) > 0$" in a standard sense—even at "bad" primes $p \mid F^\vee(\boldsymbol{c})$. Specifically, we have $|a_{\boldsymbol{c},j}(p^l)| \ll_\epsilon p^{l\epsilon}$ uniformly over tuples $\boldsymbol{c}$ (with $F^\vee(\boldsymbol{c}) \neq 0$), primes p, and integers $l \geq 1$.

In particular, "bad local factors" do *not* affect holomorphy, or absolute convergence, over any subset of the half-plane $\Re(s) > 0$. Now fix $\boldsymbol{c}$ and a prime $p \nmid F^\vee(\boldsymbol{c})$, so that $\Phi_p(\boldsymbol{c}, s) = 1 + \widetilde{S}_{\boldsymbol{c}}(p)p^{-s}$, where

$$\widetilde{S}_{\boldsymbol{c}}(p) = \widetilde{E}_{\boldsymbol{c}}(p) - p^{-1/2}\widetilde{E}_F(p) = -\tilde{\lambda}_{\boldsymbol{c}}(p) - p^{-1/2}\tilde{\lambda}_V(p).$$

The product $\Phi_p(\boldsymbol{c}, s)L_p(s, V_{\boldsymbol{c}})$ simplifies to

$$\begin{aligned}
&(1 - \tilde{\lambda}_{\boldsymbol{c}}(p)p^{-s} - \tilde{\lambda}_V(p)p^{-1/2-s})(1 + \tilde{\lambda}_{\boldsymbol{c}}(p)p^{-s} + \tilde{\lambda}_{\boldsymbol{c}}(p^2)p^{-2s} + O(p^{-3s})) \\
&= 1 - \tilde{\lambda}_V(p)p^{-1/2-s} + [\tilde{\lambda}_{\boldsymbol{c}}(p^2) - \tilde{\lambda}_{\boldsymbol{c}}(p)^2]p^{-2s} + O(p^{-1/2-2s}) + O(p^{-3s}).
\end{aligned}$$

To get further cancellation, we multiply by

$$L_p(1/2 + s, V) = 1 + \tilde{\lambda}_V(p)p^{-1/2-s} + O(p^{-1-2s}),$$

and also—motivated by the identities "$(\tilde{\lambda}_{\boldsymbol{c}}(p^2), \tilde{\lambda}_{\boldsymbol{c}}(p)^2) = (\tilde{\lambda}_{V_{\boldsymbol{c}},\mathrm{Sym}^2}(p), \tilde{\lambda}_{V_{\boldsymbol{c}},\bigotimes^2}(p))$ if $p \nmid F^\vee(\boldsymbol{c})$" and "$\bigotimes^2 = \mathrm{Sym}^2 \oplus \bigwedge^2$ in general"—by

$$L_p(2s, V_{\boldsymbol{c}}, \textstyle\bigwedge^2) = 1 + \tilde{\lambda}_{V_{\boldsymbol{c}},\bigwedge^2}(p)p^{-2s} + O(p^{-4s}),$$

to get

$$\Phi_p(\boldsymbol{c}, s)L_p(s, V_{\boldsymbol{c}})L_p(1/2 + s, V)L_p(2s, V_{\boldsymbol{c}}, \textstyle\bigwedge^2) = 1 + O(p^{-1/2-2s}) + O(p^{-3s}).$$

By definition, the left-hand side is precisely the "second-order error" $\Phi_{3,p}$, as desired. □

Remark 8.3.7. From the perspective of Φ_1, we consider Φ_2, Φ_3 "error factors" and need to separately consider large and small "error moduli" (with only the small moduli participating in the RMT-type predictions we use).

Remark 8.3.8. If m were *odd*, then $\Phi^{\mathrm{HW}}(\boldsymbol{c}, s)$ would be

$$\approx L(s, V_{\boldsymbol{c}})L(1/2+s, V)/L(2s, V_{\boldsymbol{c}}, \mathrm{Sym}^2).$$

Corollary 8.3.9 (Φ3E)**.** *Assume the setting and hypotheses of Proposition 8.3.6. Let $\mathcal{S} := \{\boldsymbol{c} \in \mathbb{Z}^m : F^\vee(\boldsymbol{c}) \neq 0\}$. Fix $A \in \mathbb{R}_{>0}$. Then uniformly over $Z, N > 0$, we have*

$$\sum_{\boldsymbol{c}\in\mathcal{S}} \mathbf{1}_{\|\boldsymbol{c}\|\le Z} \left(\sum_{n\in[N,2N]} |a_{\boldsymbol{c},3}(n)|\right)^A \ll_{A,\epsilon} Z^m N^{(1/3+\epsilon)A}.$$

Proof. If we define $0^0 := 1$, then the result trivially holds when $A = 0$. By Hölder over $\boldsymbol{c} \in \mathcal{S}$, it thus suffices to prove the (extended) result for $A \in \mathbb{Z}_{\ge 0}$. The case $A = 0$ is trivial (as noted already), so from now on, assume $A \in \mathbb{Z}_{\ge 1}$.

Given $\boldsymbol{c}, n$, let $n_{\boldsymbol{c}} := (n, F^\vee(\boldsymbol{c})^\infty)$ and $n^{\boldsymbol{c}} := n/n_{\boldsymbol{c}}$. Proposition 8.3.6 implies—uniformly over $(\boldsymbol{c}, n) \in \mathcal{S} \times \mathbb{Z}_{\ge 1}$—that $|a_{\boldsymbol{c},3}(n)| \ll_\epsilon n^\epsilon \cdot \mathbf{1}_{n^{\boldsymbol{c}}=\mathrm{sq}(n^{\boldsymbol{c}})} \cdot (n^{\boldsymbol{c}}/\operatorname{cub}(n^{\boldsymbol{c}}))^{-1/2}$.

If we fix $\boldsymbol{c} \in \mathcal{S}$, then it follows that

$$\sum_{n\in[N,2N]} |a_{\boldsymbol{c},3}(n)| \ll_\epsilon N^\epsilon \sideset{}{'}\sum_{n_{\boldsymbol{c}}\le 2N} \sideset{}{'}\sum_{n^{\boldsymbol{c}}\asymp N/n_{\boldsymbol{c}}} \mathbf{1}_{n^{\boldsymbol{c}}=\mathrm{sq}(n^{\boldsymbol{c}})} \cdot (\operatorname{cub}(n^{\boldsymbol{c}})/n^{\boldsymbol{c}})^{1/2}.$$

(On the right-hand side, we think of $n_{\boldsymbol{c}}, n^{\boldsymbol{c}}$ as separate variables, subject to the constraints $n_{\boldsymbol{c}} \mid F^\vee(\boldsymbol{c})^\infty$ and $n^{\boldsymbol{c}} \perp F^\vee(\boldsymbol{c})$.) But in general, the sum of $(\operatorname{cub}(a)/a)^{1/2}$ over *square-full* $a \asymp A$ is at most the sum over *cube-full* $d \ll A$ of the quantity

$$(d/A)^{1/2} \cdot \#\{\text{square-full } a \asymp A : d \mid a \text{ and } \sqrt{a/d} \in \mathbb{Z}\} \ll (d/A)^{1/2} \cdot (A/d)^{1/2} = 1.$$

Furthermore, $\#\{\text{cube-full } d \ll A\} \asymp A^{1/3}$, so ultimately we conclude that

$$\sum_{n\in[N,2N]} |a_{\boldsymbol{c},3}(n)| \ll_\epsilon N^\epsilon \sideset{}{'}\sum_{n_{\boldsymbol{c}}\le 2N} (N/n_{\boldsymbol{c}})^{1/3} \le N^{1/3+\epsilon} \sideset{}{'}\sum_{n_{\boldsymbol{c}}\le 2N} 1.$$

In general, if $G \in \mathbb{Z}_{\neq 0}$, then $\#\{u \le 2N : u \mid G^\infty\} \ll_\epsilon (2N \cdot G)^\epsilon$ [HB98, antepenultimate display of p. 683]. If we take $G := F^\vee(\boldsymbol{c}) \ll_F \|\boldsymbol{c}\|^{O_m(1)}$, then Corollary 8.3.9 immediately follows, *provided* $2N > Z^\eta$ holds with $\eta := 1/(2A)$, say.

Now suppose $2N \le Z^\eta$, expand $(\sum'_{n_{\boldsymbol{c}}\le 2N} 1)^A$ as a sum over $u_1, \dots, u_A \le 2N$, and switch the order of $\boldsymbol{c}, \boldsymbol{u}$. Then for each $\boldsymbol{u}$, we have $\operatorname{rad}(u_1 \cdots u_A) \le u_1 \cdots u_A \le Z^{A\eta} \le Z$. By Lang–Weil and the Chinese remainder theorem (and the well-known bound $O(1)^{\omega(\operatorname{rad}(-))} \ll_\varepsilon \operatorname{rad}(-)^\varepsilon$), we may therefore reduce Corollary 8.3.9 to the statement that for all $\varepsilon > 0$, we have

$$\sum_{u_1,\dots,u_A\le 2N} \operatorname{rad}(u_1\cdots u_A)^{\varepsilon-1} \ll_{A,\varepsilon} (2N)^{O(A^2\varepsilon)}.$$

To prove this last statement, observe that if $u_1, \dots, u_A \le 2N$ and $\operatorname{rad}(u_1 \cdots u_A) = r$, then $u_1, \dots, u_A \mid r^\infty$ and $r \le (2N)^A$; but $|\{u \le 2N : u \mid r^\infty\}^A| \ll_{A,\varepsilon} (2N \cdot r)^{A\varepsilon}$ and $\sum_{r\le(2N)^A} (2N \cdot r)^{A\varepsilon} r^{\varepsilon-1} \ll_{A,\varepsilon} (2N)^{O(A^2\varepsilon)}$, since $A \ge 1$. $\square$

Remark 8.3.10. Such care to keep Corollary 8.3.9 (Φ3E) "Z^ϵ-free" is only important in the proof of Theorem 8.4.1(a), not in the proof of Theorem 8.4.1(b).

8.3.2 Computing local averages

The local behavior of $\boldsymbol{c} \mapsto \tilde{\lambda}_{\boldsymbol{c}}(n)$ or $\boldsymbol{c} \mapsto \mu_{\boldsymbol{c}}(n)$ on average plays a basic role in our understanding of families of L-functions. For reference, write $L_p(s) = \prod_i (1 - \tilde{\alpha}_i(p)p^{-s})^{-1}$ and $1/L_p(s) = \prod_i (1 - \tilde{\alpha}_i(p)p^{-s})$, so e.g.

$$\mu_{\boldsymbol{c}}(p) = -\sum_i \tilde{\alpha}_{\boldsymbol{c},i}(p) = -\tilde{\lambda}_{\boldsymbol{c}}(p)$$

and

$$\mu_{\boldsymbol{c}}(p^2) = \sum_{i<j} \tilde{\alpha}_{\boldsymbol{c},i}(p)\tilde{\alpha}_{\boldsymbol{c},j}(p) = \frac{\tilde{\lambda}_{\boldsymbol{c}}(p)^2 - \sum_i \tilde{\alpha}_{\boldsymbol{c},i}(p)^2}{2}.$$

To prove that certain averages *exist*, we will assume Conjectures 8.2.6 (EKL) and 8.2.11 (HW2). But to satisfactorily *estimate* said averages, we will take a concrete point-counting approach (though one could use monodromy groups instead; see Remark 8.5.12 below). The result is Proposition 8.3.12 below. To concisely state the averages relevant to us, we first make a convenient archimedean definition:

Definition 8.3.11. Say that a tuple $\boldsymbol{Z} \in \mathbb{R}^m$ is *β-lopsided* if $|Z_i| \le |Z_j|^\beta$ for all $(i,j) \in [m]^2$. Then view $\{\boldsymbol{Z} \in \mathbb{R}^m : \boldsymbol{Z} \text{ is } \beta\text{-lopsided}\}$ as a topological subspace of $\mathbb{R}^m \cup \{\boldsymbol{\infty}\} \cong S^m$.

Proposition 8.3.12 (LocAv, or LA)**.** *Fix F with $m \ge 4$* even, *and* assume *(EKL) and (HW2). Now fix $(\boldsymbol{a}, n_0, \beta, n) \in \mathbb{Z}^m \times \mathbb{Z}_{\ge 1} \times \mathbb{R}_{\ge 1} \times \mathbb{Z}_{\ge 1}$. If we restrict $\boldsymbol{c}$ to the locus $F^\vee(\boldsymbol{c}) \ne 0$, then the following two limits* exist*, and are* independent *of β:*

$$\bar{\mu}_F^{\boldsymbol{a},n_0}(n) := \lim_{\boldsymbol{Z}\to\boldsymbol{\infty}} \left(\frac{1}{|\mathcal{B}(\boldsymbol{Z}) \cap \{\boldsymbol{a} \bmod n_0\}|} \cdot \sideset{}{'}\sum_{\boldsymbol{c} \equiv \boldsymbol{a} \bmod n_0} \mathbf{1}_{\boldsymbol{c} \in \mathcal{B}(\boldsymbol{Z})} \cdot \mu_{\boldsymbol{c}}(n) \right)$$

$$\bar{\mu}_{F,2}^{\boldsymbol{a},n_0}(n_1,n_2) := \lim_{\boldsymbol{Z}\to\boldsymbol{\infty}} \left(\frac{1}{|\mathcal{B}(\boldsymbol{Z}) \cap \{\boldsymbol{a} \bmod n_0\}|} \cdot \sideset{}{'}\sum_{\boldsymbol{c} \equiv \boldsymbol{a} \bmod n_0} \mathbf{1}_{\boldsymbol{c} \in \mathcal{B}(\boldsymbol{Z})} \cdot \mu_{\boldsymbol{c}}(n_1)\mu_{\boldsymbol{c}}(n_2) \right).$$

Furthermore, for some absolute constants $\delta, \delta' > 0$, the following statements hold.

(LA1) The function $n \mapsto \bar{\mu}_F^{\boldsymbol{a},n_0}(n)$ is multiplicative*. For $\Re(s) \ge 1/2 - \delta$, we have*

$$\sum_{l \ge 0} p^{-ls} \bar{\mu}_F^{\boldsymbol{a},n_0}(p^l) = 1 + (\tilde{\lambda}_V(p)p^{-s-1/2} + p^{-2s}) + O(p^{O(v_p(n_0)-1-\delta'}),$$

uniformly over $p, s, \boldsymbol{a}, n_0, \beta$.

(LA2) The function $(n_1,n_2) \mapsto \bar{\mu}_{F,2}^{\boldsymbol{a},n_0}(n_1,n_2)$ is multiplicative *(i.e. if $\gcd(n_1n_2, n_1'n_2') = 1$, then $\bar{\mu}_{F,2}^{\boldsymbol{a},n_0}(\boldsymbol{n})\bar{\mu}_{F,2}^{\boldsymbol{a},n_0}(\boldsymbol{n}') = \bar{\mu}_{F,2}^{\boldsymbol{a},n_0}(\boldsymbol{n}\boldsymbol{n}')$). For $\Re(s) \ge 1/2 - \delta$, we have*

$$\sum_{\boldsymbol{l} \ge 0} p^{-\boldsymbol{l}\cdot\boldsymbol{s}} \bar{\mu}_{F,2}^{\boldsymbol{a},n_0}(p^{\boldsymbol{l}}) = 1 + p^{-s_1-s_2} + \sum_{j\in[2]} (\tilde{\lambda}_V(p)p^{-s_j-1/2} + p^{-2s_j}) + O(p^{O(v_p(n_0))-1-\delta'}),$$

uniformly over $p, s, \boldsymbol{a}, n_0, \beta$.

Definition 8.3.13. Let $\bar{\mu}_F(n) := \bar{\mu}_F^{\mathbf{0},1}(n)$, and $\bar{\mu}_{F,2}(n_1, n_2) := \bar{\mu}_{F,2}^{\mathbf{0},1}(n_1, n_2)$.

Definition 8.3.14. For p-adic calculations only, we use the convenient convention "$\mu_{\boldsymbol{c}}, a_{\boldsymbol{c},1}, a_{\boldsymbol{c},2}, a_{\boldsymbol{c},3} := 0$" for all $\boldsymbol{c} \in \mathbb{Z}_p^m$ with $F^\vee(\boldsymbol{c}) = 0$. (In particular, for such $\boldsymbol{c}$'s, we set $\mu_{\boldsymbol{c}}(1) := 0$, etc., so $\mu_{\boldsymbol{c}}, a_{\boldsymbol{c},1}, a_{\boldsymbol{c},2}, a_{\boldsymbol{c},3}$ are *not* multiplicative in the standard sense.)

Proof of existence and multiplicativity of limits. By (EKL), [Wan21a, Corollary 6.3], and the bound $|\tilde{\alpha}(p)| \le 1$ (from GRC), it is routine to show that the p-adic averages

$$\mathbb{E}_{\boldsymbol{c}\in\mathbb{Z}_p^m \cap \{\boldsymbol{a} \bmod n_0\}}[\mu_{\boldsymbol{c}}(p^l)] \quad \text{and} \quad \mathbb{E}_{\boldsymbol{c}\in\mathbb{Z}_p^m \cap \{\boldsymbol{a} \bmod n_0\}}[\mu_{\boldsymbol{c}}(p^{l_1})\mu_{\boldsymbol{c}}(p^{l_2})]$$

are well-defined, and furthermore (by the Chinese remainder theorem) that these p-adic averages determine $\bar{\mu}_F^{\boldsymbol{a},n_0}(n), \bar{\mu}_{F,2}^{\boldsymbol{a},n_0}(n_1, n_2)$ in the obvious way. (Note that in $\mathbb{Z}^m$, the locus $F^\vee(\boldsymbol{c}) = 0$ has "density 0" in the boxes defining $\bar{\mu}_F^{\boldsymbol{a},n_0}(n), \bar{\mu}_{F,2}^{\boldsymbol{a},n_0}(n_1, n_2)$.) $\square$

Beginning of proof of required estimates. If $\delta' \le 1$, say, then by GRC, we may "absorb" the case $p \mid n_0$ into the error term $O(p^{O(v_p(n_0))}p^{-1-\delta'})$. Now assume $p \nmid n_0$. If $3 \cdot (1/2 - \delta) \ge 1 + \delta'$, then we are left with analyzing the contributions from $l \in \{1, 2\}$ in (LA1), and $|\boldsymbol{l}| \in \{1, 2\}$ in (LA2). In other words, up to re-defining δ, δ', we must prove the following.

(1) $\mathbb{E}_{\boldsymbol{c}\in\mathbb{Z}_p^m}[\mu_{\boldsymbol{c}}(p)] = \tilde{\lambda}_V(p)p^{-1/2} + O(p^{-1/2-\delta'})$, for $l = 1$.

(2) $\mathbb{E}_{\boldsymbol{c}\in\mathbb{Z}_p^m}[\mu_{\boldsymbol{c}}(p^2)] = 1 + O(p^{-\delta'})$, for $l = 2$. (Cf. "essentially symplectic" in §8.5.2, which begins near p. 95.)

(3) $\mathbb{E}_{\boldsymbol{c}\in\mathbb{Z}_p^m}[\mu_{\boldsymbol{c}}(p)^2] = 1 + O(p^{-\delta'})$, for $\boldsymbol{l} = (1, 1)$. (Cf. "essentially cuspidal and self-dual".)

(Note that (1)–(2) "cover" not just the cases "$l \in \{1, 2\}$" in (LA1), but also the cases "$|\boldsymbol{l}| \in \{1, 2\}$ with $l_1 l_2 = 0$" in (LA2).)

But in each of (1)–(3), the locus $p \mid F^\vee(\boldsymbol{c})$ fits in an $O(p^{-1})$ error term, by Lang–Weil and GRC. We may then restrict to $p \nmid F^\vee(\boldsymbol{c})$, in which case $\mu_{\boldsymbol{c}}(p) = -\tilde{\lambda}_{\boldsymbol{c}}(p) := -\tilde{\operatorname{tr}}(\mathrm{Frob}\,|_{H^{m_*}}) = \widetilde{E}_{\boldsymbol{c}}(p)$ and $\mu_{\boldsymbol{c}}(p)^2 = \widetilde{E}_{\boldsymbol{c}}(p)^2$, while $\mu_{\boldsymbol{c}}(p^2) = \frac{1}{2}(\tilde{\lambda}_{\boldsymbol{c}}(p)^2 - \sum_i \tilde{\alpha}_{\boldsymbol{c},i}(p)^2) = \frac{1}{2}(\widetilde{E}_{\boldsymbol{c}}(p)^2 + \widetilde{E}_{\boldsymbol{c}}(p^2))$.[3] (Note that these formulas for $\mu_{\boldsymbol{c}}(p^\bullet)$ are only fully correct when m_* is *odd*.)

It thus remains to prove the following; but these follow from Corollary 5.1.19.

(a) $\mathbb{E}_{\boldsymbol{c}\in\mathbb{F}_p^m}[\widetilde{E}_{\boldsymbol{c}}(p)\mathbf{1}_{p\nmid F^\vee(\boldsymbol{c})}] = \tilde{\lambda}_V(p)p^{-1/2} + O(p^{-1/2-\delta'})$.

(b) $\mathbb{E}_{\boldsymbol{c}\in\mathbb{F}_p^m}[\widetilde{E}_{\boldsymbol{c}}(p^2)\mathbf{1}_{p\nmid F^\vee(\boldsymbol{c})}] = 1 + O(p^{-\delta'})$.

(c) $\mathbb{E}_{\boldsymbol{c}\in\mathbb{F}_p^m}[\widetilde{E}_{\boldsymbol{c}}(p)^2\mathbf{1}_{p\nmid F^\vee(\boldsymbol{c})}] = 1 + O(p^{-\delta'})$.

(Note here that (a) implies (1), and (c) implies (3), while (b)–(c) imply (2).) $\square$

[3] If $p \nmid F^\vee(\boldsymbol{c})$, then $\mu_{\boldsymbol{c}}(p^2) = (-1)^2\tilde{\lambda}_{V_{\boldsymbol{c}},\bigwedge^2}(p)$, but we only need this $\bigwedge^2$ interpretation elsewhere.

8.3.3 "Deriving" the Ratios Conjectures

Assume the hypotheses, (EKL) and (HW2), of Proposition 8.3.12 (LocAv). Then for the "ratio" $1/L$ (or for "pure products" thereof), the recipe [CFZ08, §5.1]—carried over to the general setting of [SST16, pp. 534–535, Geometric Families and Remark (i)]—now makes sense for the families $\boldsymbol{c} \mapsto \pi_{\boldsymbol{c}}$ underlying Conjectures 8.2.18 and 8.2.28. We will soon "derive" these two conjectures accordingly.

Remark 8.3.15. If one were willing to directly apply a similar "recipe" to one of $\Phi_1\Phi_2, \Phi^{\mathrm{HW}}, \Phi$ (and not just $1/L$ or Φ_1), then our work would simplify correspondingly. But it is unclear how generally a naive recipe like that could hold, without a clear supporting model like a classical random matrix ensemble.

Remark 8.3.16. In our specific "geometric" Ratios Conjectures, (i) we are assuming (EKL), not just (HW2); (ii) the L-functions $L(s, V_{\boldsymbol{c}})$ are not all primitive, as [CFK⁺05] explicitly requires, and [CFZ08] thus implicitly requires; and (iii) we do not order our families by conductor (or by discriminant, for that matter). This is all OK:

(1) Regarding (i), we use (EKL) here (in §8.3.3) merely to prove that certain local averages *exist*. This is technically important, but far from the heart of RMT-type matters. In fact, it would be harmless to explicitly include (EKL) as a hypothesis in Conjectures 8.2.18 and 8.2.28, if one wanted to do so.

(2) Point (ii) is minor. Using (HW2), one can show that $L(s, V_{\boldsymbol{c}})$ is primitive for all $\boldsymbol{c}$'s outside a fairly sparse set. (See "essentially cuspidal" in §8.5.2, which begins near p. 95.) So for our purposes (locally and globally), the question of whether to include all $L(s, V_{\boldsymbol{c}})$'s, or only primitive ones, is essentially cosmetic.

   Alternatively, it may well be that primitivity is not essential, even on average; see the sentence "Non-primitive families can also be handled…" on [CFK⁺05, p. 34].

(3) Regarding (iii), we are still following the RMT-based philosophy underlying [CFZ08]. We are just indexing by different level sets, as is natural for multi-parameter families like ours; cf. [SST16, p. 535, Remark (i); and p. 560, second paragraph after (25)].

   For a more thorough discussion of the expected RMT-type models for our families, see §8.5.2, which begins near p. 95.

Deriving (R1)–(R2)

To "derive" Conjecture 8.2.18(R1), replace each term

$$L(s, \pi_{\boldsymbol{c}})^{-1} = \sum_{n \geq 1} \mu_{\boldsymbol{c}}(n) n^{-s}$$

on the "left-hand side of (R1)" with its "naive expected value over $\boldsymbol{c}$", i.e.

$$\sum_{n \geq 1} \bar{\mu}_F(n) n^{-s} = \prod_p \sum_{l \geq 0} p^{-ls} \bar{\mu}_F(p^l) = \prod_p \left[1 + (\tilde{\lambda}_V(p) p^{-s-1/2} + p^{-2s}) + O(p^{-1-\delta'})\right]$$

for $\Re(s) \geq 1/2 - \delta$. This "naive average" factors as $\zeta(2s)L(s+1/2, V)A_F(s)$, for a certain Euler product $A_F(s)$ that converges absolutely on the half-plane $\Re(s) \geq 1/2 - \delta$.

To "derive" Conjecture 8.2.18(R2), similarly replace

$$L(s_1, \pi_{\boldsymbol{c}})^{-1} L(s_2, \pi_{\boldsymbol{c}})^{-1} = \sum_{n_1, n_2 \geq 1} \mu_{\boldsymbol{c}}(n_1) \mu_{\boldsymbol{c}}(n_2) \boldsymbol{n}^{-\boldsymbol{s}},$$

for $\Re(\boldsymbol{s}) \geq 1/2 - \delta$, with

$$\prod_p \sum_{\boldsymbol{l} \geq 0} p^{-\boldsymbol{l} \cdot \boldsymbol{s}} \bar{\mu}_{F,2}(p^{\boldsymbol{l}}) = \prod_p \left[1 + p^{-s_1 - s_2} + \sum_{j \in [2]} (\tilde{\lambda}_V(p) p^{-s_j - 1/2} + p^{-2s_j}) + O(p^{-1-\delta'}) \right],$$

which factors as $A_{F,2}(\boldsymbol{s}) \zeta(s_1 + s_2) \prod_{j \in [2]} \zeta(2s_j) L(s_j + 1/2, V)$.

Remark 8.3.17. Here $\zeta(2s), L(s+1/2, V), \zeta(s_1+s_2), \ldots$ are called *polar factors*.

Remark 8.3.18. When there are L's in the numerator, the recipe must be modified using the approximate functional equation for L. (In particular, the root number, gamma factor, and conductor play a visible role on average, whereas in (R1)–(R2) they seem to be be hidden in the σ-dependence, i.e. how close s is allowed to be to the critical line.)

Remark 8.3.19. Under a quasi-GRH for $\zeta(s), L(s, V)$, the errors in (R1)–(R2) remain essentially the same (up to a factor of $O_{\sigma, \boldsymbol{\sigma}, \epsilon}(Z^{\hbar\epsilon})$) even after dividing both sides of (R1) by $\zeta(2s)L(s+1/2, V)$, or both sides of (R2) by $\prod_{j \in [2]} \zeta(2s_j) L(s_j + 1/2, V)$.

Deriving (RA1)

To "derive" Conjecture 8.2.28 (RA1), replace each term

$$L(s, \pi_{\boldsymbol{c}})^{-1} = \sum_{n \geq 1} \mu_{\boldsymbol{c}}(n) n^{-s}$$

on the "left-hand side of (RA1)", for $\Re(s) \geq 1/2 - \delta$, with its "naive expected value over $\boldsymbol{c} \equiv \boldsymbol{a} \bmod n_0$", i.e.

$$\begin{aligned} \sum_{n \geq 1} \frac{\bar{\mu}_F^{\boldsymbol{a}, n_0}(n)}{n^s} &= \prod_p \sum_{l \geq 0} \frac{\bar{\mu}_F^{\boldsymbol{a}, n_0}(p^l)}{p^{ls}} \\ &= \prod_p \left[1 + (\tilde{\lambda}_V(p) p^{-s-1/2} + p^{-2s}) + O(p^{O(v_p(n_0)) - 1 - \delta'}) \right]. \end{aligned}$$

This "naive average" factors as $\zeta(2s)L(s+1/2, V)A_F^{\boldsymbol{a}, n_0}(s)$, for a certain Euler product $A_F^{\boldsymbol{a}, n_0}(s)$ that converges absolutely on the half-plane $\Re(s) \geq 1/2 - \delta$.

8.4 Main conditional results

Let Σ_{gen} denote the contribution to the right-hand side of eq. (3.2) from the locus $F^\vee(\boldsymbol{c}) \neq 0$. (Explicitly, $\Sigma_{\text{gen}} := X^{m-3} \sum_{n \geq 1} \sum_{\boldsymbol{c} \in \mathbb{Z}^m} \mathbf{1}_{F^\vee(\boldsymbol{c}) \neq 0} \cdot n^{-(m-1)/2} \widetilde{S}_{\boldsymbol{c}}(n) \widetilde{I}_{\boldsymbol{c}}(n)$.)

Theorem 8.4.1 ([Wan21a, Theorem 3.49]). *Fix F diagonal with $m \in \{4, 6\}$, and fix $w \in C_c^\infty(\mathbb{R}^m)$ with (F, w) clean. Assume* $\operatorname{lcm}(\boldsymbol{F})$ *is cube-free. Then*

(a) $|\Sigma_{\mathrm{gen}}| \ll X^{3(m-2)/4}$ *holds* under (HWSp), ($\mathrm{SFSC}_{q,6}$+), and (R2'); *and*

(b) $|\Sigma_{\mathrm{gen}}| \ll X^{3(m-2)/4-\Omega(1)}$ *holds* under (HWSp), ($\mathrm{SFSC}_{q,6}$+), (RA1), and (EKL).

Proof sketch. We use ($\mathrm{SFSC}_{q,6}$+) precisely to ensure that Conjecture 7.3.7 (B3) holds (see §7.3). Now factor Φ using Definition 8.3.2. For (a), begin with a framework for "restriction and separation" going beyond Remark 4.1.2 (see [Wan21a, §5] for details), and then use Hölder appropriately between "good" and "bad" factors; some important ingredients are (B3), Lemma 7.1.4, Conjecture 8.2.24 (R2'), [Wan21a, Proposition 7.27], and Corollary 8.3.9 (Φ3E). For (b), we handle some ranges (namely those with large "error moduli") as in (a). Over the remaining ranges, we then decompose Σ_{gen} "adelically" into pieces—based on the polynomial H from Conjecture 8.2.6 (EKL)—up to a small exceptional set constructed in [Wan21a, §7.8] by algorithmic tree-like means. We then estimate these pieces via local calculations and Poisson summation. $\square$

Remark 8.4.2. In Theorem 8.4.1(b), the power saving $\Omega(1)$ deserves some clarification, because the situation here is not as clear-cut as that in Corollary 6.1.4.

(1) The saving $\Omega(1)$ can be safely taken to be independent of w: no *exponent* anywhere in [Wan21a] truly depends on the weight w. (In particular, (SFSC), (RA1), and (EKL) depend only on F, not on w.)

(2) *If* (SFSC), (RA1), and (EKL) are assumed to be sufficiently "exponent-uniform" over F, then one can take $\Omega(1)$ to be independent of F (but still dependent on m).

Regarding the implied constant in Theorem 8.4.1(b), one could probably work out an explicit w-dependence of the form $(1 + \operatorname{diam}(\operatorname{Supp} w) + \|w\|_{O_F(1),\infty})^{O_F(1)}$ (where $\|w\|_{k,\infty}$ denotes a Sobolev norm), given enough patience. An explicit F-dependence would probably take even more patience.

Theorem 8.4.3 ([Wan21a, Theorem 3.39]). *Fix F diagonal with $m = 6$. Assume* $\operatorname{lcm}(\boldsymbol{F})$ *is cube-free.*

(a) Say $F = x_1^3 + \cdots + x_6^3$, and assume (HWSp), ($\mathrm{SFSC}_{q,6}$+), and (R2'). *Then $N_F(X) \ll X^3$ holds as $X \to \infty$. Therefore, a positive fraction of integers lie in $\{x^3+y^3+z^3 : (x, y, z) \in \mathbb{Z}_{\geq 0}^3\}$.*

(b) Alternatively, assume (HWSp), ($\mathrm{SFSC}_{q,6}$+), (RA1), and (EKL). *Then for any* given *$w \in C_c^\infty(\mathbb{R}^m)$ with (F, w) clean, the pair (F, w) is HLH (with a power saving), in the sense of Definition 1.4.6. Therefore,*

(i) the Hasse principle holds for $V/\mathbb{Q}$; and also

(ii) asymptotically 100% *of integers $a \not\equiv \pm 4$* mod 9 *lie in $\{x^3 + y^3 + z^3 : (x, y, z) \in \mathbb{Z}^3\}$, if $F = x_1^3 + \cdots + x_6^3$.*

Proof of Theorem 8.4.3(a) assuming Theorem 8.4.1(a). By a Hölder argument (see [Wan21a] for details), $N_F(X) \ll X^3$. Consequently, Observation 2.1.3 implies *positive lower density* of $\{x^3 + y^3 + z^3 : (x, y, z) \in \mathbb{Z}^3_{\geq 0}\}$. □

Proof of Theorem 8.4.3(b) assuming Theorem 8.4.1(b). Using Corollary 6.1.4—and the cleanliness assumption in Theorem 8.4.3(b)—we find that Theorem 8.4.1(b) directly implies HLH for (F, w), in fact with (an unnecessary) power saving. Theorem 8.4.3(b)(i) then immediately follows, upon choosing a weight $w \in C_c^\infty(\mathbb{R}^6)$ with (F, w) clean and $c_{\mathrm{HLH},F,w} > 0$ (doable by hand, or via Proposition 6.2.3). On the other hand, Theorem 8.4.3(b)(ii) follows from Theorem 2.1.8 (essentially due to [Dia19]). □

Remark 8.4.4. See the introductions to [HB99, HB07] for some history, and (what is likely) the state of affairs, on the Hasse principle for diagonal cubic forms (although [HB07] also discusses non-diagonal cubic forms)—with the unconditional record being $m = 7$, due to [Bak89]. (Over global function fields of characteristic ≥ 7, the record is $m \geq 6$, achieved by geometric techniques [Tia17].)

Remark 8.4.5. In Theorem 8.4.3 (and in Theorem 8.4.1), one could relax the assumption (HWSp) to (HW2), but that would muddy the proof, with little benefit. (For a brief sketch of the necessary modifications, see [Wan21a, Appendix A.4].)

Remark 8.4.6. In Theorem 8.4.3 (and in Theorem 8.4.1), I expect that the assumption ($\mathrm{SFSC}_{q,6}$+) can be relaxed to ($\mathrm{SFSC}_{p,3}$); see Remark 7.3.13. Furthermore, I believe (the proof of) Theorem 8.4.1 would directly generalize to $\mathbb{P}^{m-1}_{\mathbb{Q}}$-smooth F if one replaced the assumption ($\mathrm{SFSC}_{q,6}$+) with Conjecture 7.3.7 (B3); however, to generalize (the proof of) Theorem 8.4.3 accordingly, one would also need to generalize Corollary 6.1.4.

Remark 8.4.7. I expect that with a lot of additional technical work (cf. Remark 7.1.6), one could replace the cleanliness condition in Theorem 8.4.3(b) with the condition that (F, w) be smooth. (If successful, this would, in particular, conditionally imply the original HLH conjecture of Hooley from Example 1.2.1.) I suspect that if (F, w) is smooth, then in the key "generic range" where $|c_1|, \ldots, |c_m|$ are all roughly of size $X^{1/2}$ (up to a factor of $X^{\pm\delta}$), a slightly deformed version of Lemma 7.1.4 should hold (because in the critical range $|u| \asymp \|\boldsymbol{v}\| \gg 1$, the zeros of $u\nabla F(\boldsymbol{x}) - \boldsymbol{v}$ would then lie relatively far from $(\operatorname{hess} F)_{\mathbb{R}}(\mathbb{R})$), and thus the proof of Theorem 8.4.1 should remain adequate. (The complementary "non-generic range" of $\boldsymbol{c}$'s can be handled under GRH, following e.g. §4.1.)

8.5 Supplementary material on L-functions

8.5.1 A discussion of (HW2) and (HWSp)

Generalities

Fix F with $m \geq 3$. (Here we *allow arbitrary* $m \geq 3$ *and* $\mathbb{P}^{m-1}_{\mathbb{Q}}$*-smooth* F, until further notice.) Fix a tuple $\boldsymbol{c} \in \mathbb{Z}^m$ with $F^\vee(\boldsymbol{c}) \neq 0$. In the notation of Definition 8.2.4 and Conjecture 8.2.11 (HW2), then define $m_*, M_{\boldsymbol{c}}, M_V$ and fix (M, S).

Remark 8.5.1. In Definition 8.2.1, we fixed an auxiliary pair (ℓ_0, ι). But we could have instead used Serre's definition of Hasse–Weil L-functions (and its extension to motives), where one *assumes* "ℓ-independence of the local factors L_p" (known for us by [Las17, Corollary 1.2],[4] since our M's are "tensor-generated" by smooth projective hypersurfaces). Note that [Hoo86b, HB98] both defined $L(s, V_{\boldsymbol{c}})$ following Serre; for $m \in \{4, 6\}$, the ℓ-independence of the $L_p(s, V_{\boldsymbol{c}})$'s was already known at the time, due to connections with abelian varieties; cf. [Tay04, paragraph after Conjecture 1.2].

Remark 8.5.2. It is conjectured that M should be semi-simple (as an ℓ_0-adic representation of $G_{\mathbb{Q}}$), in which case [Tay04, p. 100, prior to Conjectures 3.4–3.5] precisely "specifies" (HW2)'s putative π_M. Semi-simplicity is known at least for $M = M_{\boldsymbol{c}}$ when $m \in \{4, 6\}$, and thus (by a general representation-theoretic argument[5]) for all M's "tensor-generated" by such $M_{\boldsymbol{c}}$'s.

We now briefly elaborate on the (conjectural) rationale for each part of (HW2).

(a) By definition, $M_{\boldsymbol{c}}, M_V$ arise from geometry, pure of weight $m_*, 1 + m_*$, respectively. And if $M \neq M_{\boldsymbol{c}}, M_V$, then by the Künneth $G_{\mathbb{Q}}$-isomorphism, M is a subquotient of $H^{2m_*}((V_{\boldsymbol{c}})_{\overline{\mathbb{Q}}} \times (V_{\boldsymbol{c}})_{\overline{\mathbb{Q}}}, \mathbb{Q}_{\ell_0})$—whence M arises from geometry, pure of weight $2m_*$.

It is known that $\dim M_{\boldsymbol{c}}, \dim M_V \ll_m 1$ (uniformly over $\boldsymbol{c}$), so $\dim M \ll_m 1$ in general as well.

Also, M is unramified away from S (by smooth proper base change), so we should have $q(M) \mid \operatorname{rad}(S)^{O_m(1)}$—trivially if $M = M_V$, and by the "bounded depth in families" observation of [SST16, §2.11] if $M \neq M_V$.

(b) The *finiteness* of the *set* of gamma factors $L_\infty(s, M)$ (for a *given* value of m, as $\boldsymbol{c}$ varies) should follow at least formally from Hodge theory; see [Tay04, p. 79, definition of $\mathrm{HT}(-)$; p. 86, definition of $\Gamma(-, s)$; and p. 80, Conjecture 1.2].

(c) This is a precise instance of the *Langlands reciprocity conjecture*; cf. [SST16, Conjecture 4]. See also [SST16, pp. 534–535, Geometric Families] for some context.

(d) We say that a cuspidal π *satisfies GRC* if it is tempered at all places. If each cuspidal constituent of the putative isobaric π_M is tempered, then $|\tilde{\alpha}_{M,j}(p)| \leq 1$ for all p, j, and also $L_\infty(s, M)$ is holomorphic on $\{\sigma > 0\}$; but the implication, even in the cuspidal case, may not literally be an equivalence (cf. [FPRS19, Axiom 4 and Lemma 3.2]).

(e) We say that $L(s, \pi_M)$ *satisfies GRH* if all of its zeros in the region $\Re(s) > 0$ lie on the line $\Re(s) = 1/2$. (Under our "assumptions" on π_M from (HW2)(c)–(d), it would be *equivalent* to require GRH to hold for all "cuspidal constituent" L-functions. This is because all such "constituents" are known to be zero-free for $\Re(s) \geq 1$, by [IK04, Theorem 5.42].)

[4] which is based in part on Scholze's results on the weight-monodromy conjecture

[5] using (i) a representation-theoretic passage from a *topological* subgroup $H \subseteq \mathrm{GL}_{\dim M_{\boldsymbol{c}}}(\mathbb{Q}_{\ell_0})$ to its *Zariski* closure in $\mathrm{GL}_{\dim M_{\boldsymbol{c}}} / \mathbb{Q}_{\ell_0}$, based on the relative *coarseness* of the Zariski topology; and (ii) [Mil17, Corollary 22.44], a general result on *algebraic* groups over fields of characteristic zero

Remark 8.5.3. (HW2) and Godement–Jacquet imply that $L(s, M)$ has finite order and a standard functional equation (with $|\epsilon(M)| = 1$, etc.), and is holomorphic except possibly for poles at $s = 1$ corresponding to trivial constituents of π_M. (Here the *unitarity* in Definition 8.2.10 restricts poles to $\Re(s) = 1$, and the *finiteness* restricts poles to $s = 1$.) Together with (HW2)(b), GRC, and GRH, these analytic properties imply the uniform estimate $1/L(s, M) \ll_{m,\epsilon} q(M)^\epsilon(1 + |s|)^\epsilon$ over $\boldsymbol{c}, M, s$ with $\Re(s) \geq 1/2 + \epsilon$; see e.g. [IK04, Theorem 5.19 and the ensuing paragraph].

We now explain the (conjectural) rationale for (HWSp).

(1) The first part of (HWSp) is (HW2), which we have already explained.

(2) For simplicity, say $m \geq 4$, i.e. $m_* \geq 1$. Given $X \in \{V_{\boldsymbol{c}}, V\}$ of dimension $d \in \{m_*, 1+m_*\}$, consider the cup-product pairing

$$\psi\colon H^d(X_{\overline{\mathbb{Q}}}, \mathbb{Q}_{\ell_0}) \times H^d(X_{\overline{\mathbb{Q}}}, \mathbb{Q}_{\ell_0}) \to H^{2d}(X_{\overline{\mathbb{Q}}}, \mathbb{Q}_{\ell_0}) \cong \mathbb{Q}_{\ell_0}(-d),$$

where $\mathbb{Q}_{\ell_0}(-d)$ denotes the Tate motive of weight $2d$. By Poincaré duality, ψ is non-degenerate. In fact, ψ induces a non-degenerate pairing on $H^d_{\mathrm{diff}}(X) = H^d(X)/H^d(\mathbb{P}^{m-1})$, as we now recall.

Case 1: d is odd. Then $H^d(\mathbb{P}^{m-1}) = 0$, so $H^d_{\mathrm{diff}}(X) = H^d(X)$. Thus ψ can (trivially) be viewed as a non-degenerate pairing on $H^d_{\mathrm{diff}}(X)$.

Case 2: d is even. Then ψ is symmetric (and non-degenerate), and $H^d(\mathbb{P}^{m-1})$ is one-dimensional. It is known that the restriction $\psi|_{H^d(\mathbb{P}^{m-1})}$ is non-degenerate; this follows, for instance, from Poincaré duality (for a copy of $\mathbb{P}^d$ sitting in $\mathbb{P}^{m-1}$) and the functoriality of cup products. Therefore, we have an orthogonal direct sum decomposition $H^d(X) = H^d(\mathbb{P}^{m-1}) \oplus H^d(\mathbb{P}^{m-1})^\perp$, and $\psi|_{H^d(\mathbb{P}^{m-1})^\perp}$ is non-degenerate. Via the decomposition, $H^d_{\mathrm{diff}}(X) \cong H^d(\mathbb{P}^{m-1})^\perp$, so $\psi|_{H^d(\mathbb{P}^{m-1})^\perp}$ can be viewed as a non-degenerate pairing on $H^d_{\mathrm{diff}}(X)$.

It follows that there are non-degenerate pairings $M_{\boldsymbol{c}} \times M_{\boldsymbol{c}} \to \mathbb{Q}_{\ell_0}(-m_*)$ and $M_V \times M_V \to \mathbb{Q}_{\ell_0}(-1 - m_*)$, whence $M \cong M^\vee(-w)$ if $M \in \{M_{\boldsymbol{c}}, M_V\}$. Hence $M \cong M^\vee(-w)$ even if $M \notin \{M_{\boldsymbol{c}}, M_V\}$. In every case, M is self-dual up to a Tate twist of weight $2w$. So π_M should be self-dual on the nose.

(3) Say $2 \mid m$, i.e. $2 \nmid m_*$. Then the aforementioned pairing $M_{\boldsymbol{c}} \times M_{\boldsymbol{c}} \to \mathbb{Q}_{\ell_0}(-m_*)$ is *skew-symmetric*. So under Tate's global semi-simplicity conjecture, the representation $M_{\boldsymbol{c}} \wedge M_{\boldsymbol{c}}$ should decompose as $\mathbb{Q}_{\ell_0}(-m_*) \oplus M'_{\boldsymbol{c},2}$, for some representation $M'_{\boldsymbol{c},2}$ of weight $2m_*$. Then $M'_{\boldsymbol{c},2}$ should correspond to some *nice isobaric* $\phi_{\boldsymbol{c},2}$, with the expected compatibilities (a)–(b) of (HWSp).

Remark 8.5.4. There is an intuitive reason for $M_{\boldsymbol{c}}, M_V$ to be self-dual: at least at good primes, $M_{\boldsymbol{c}}, M_V$ arise from *point counts*, which are obviously always *integral* (and therefore *real*).

Remark 8.5.5. Suppose $2 \mid m$, and assume (HWSp). Then most of the standard analytic properties of $L(s, V_{\boldsymbol{c}}, \bigwedge^2)$—based on (HW2)—carry over to $L(s, \phi_{\boldsymbol{c},2}) = L(s, V_{\boldsymbol{c}}, \bigwedge^2)/\zeta(s)$. Furthermore, $L(s, \phi_{\boldsymbol{c},2})$ is zero-free for $\Re(s) \geq 1$. By GRH for $L(s, V_{\boldsymbol{c}}, \bigwedge^2)$, it follows that $1/L(s, \phi_{\boldsymbol{c},2}) \ll_{m,\epsilon} q(V_{\boldsymbol{c}})^\epsilon(1 + |s|)^\epsilon$ holds uniformly over $\boldsymbol{c}, s$ with $\Re(s) \geq 1/2 + \epsilon$.

Let $\pi_{\boldsymbol{c}} := \pi_{M_{\boldsymbol{c}}}$. Before proceeding, note that if $2 \mid m$ and $M_{\boldsymbol{c}}$ is *irreducible*, then we expect the putative $\pi_{\boldsymbol{c}}$ to be *cuspidal self-dual symplectic* as defined on [SST16, p. 533]. Out of idle curiosity, we raise the following refined question, which could be too optimistic.

Question 8.5.6. Say $2 \mid m$, and assume (HW2). Then is it true that each cuspidal constituent $\pi_{\boldsymbol{c},i}$ of $\pi_{\boldsymbol{c}}$ is *self-dual* and *symplectic*, with $L(s, \pi_{\boldsymbol{c},i}, \bigwedge^2)$ analytic except for a simple pole at $s = 1$? (If this is really true, then all $\pi_{\boldsymbol{c},i}$ must be nontrivial, i.e. $L(s, \pi_{\boldsymbol{c}})$ must be entire.)

Remark 8.5.7. The fact that $\pi_{\boldsymbol{c}}$ may not be cuspidal muddies the waters. Perhaps [Gro16] can clarify matters (with the notion of a symplectic motive).

Example 8.5.8. Say $m = 4$. Then $\pi_{\boldsymbol{c}}$ is the representation generated by the weight 2 modular cusp form of level $N_{J(V_{\boldsymbol{c}})}$ associated to the elliptic curve $J(V_{\boldsymbol{c}})/\mathbb{Q}$. Here $L(s, \pi_{\boldsymbol{c}}, \bigwedge^2) = \zeta(s)$.

Example 8.5.9. Fix F diagonal, and suppose $\boldsymbol{c} = (0, \dots, 0, 1)$. Then $V_{\boldsymbol{c}}$ is a diagonal cubic hypersurface of dimension m_*, so $L(s, V_{\boldsymbol{c}})$ is a product of *normalized* L-functions attached to algebraic Hecke characters of weight m_* on $\mathbb{Q}(\zeta_3)$. Furthermore, if $2 \mid m$, then no factors of $\zeta(s)$ can appear, so $L(s, V_{\boldsymbol{c}})$ must be entire. (These facts about $L(s, V_{\boldsymbol{c}})$ are classical; see Lemma 8.6.7 below for details.)

Alternative motivic descriptions

Fix $m, F, \boldsymbol{c}$ as before. Then $V_{\boldsymbol{c}}$ is isomorphic to a smooth projective cubic hypersurface of dimension $m_* \geq 0$.

Often the representation $M_{\boldsymbol{c}}$ can also be realized from other perspectives, at least *up to semi-simplification*. Recall that by Chebotarev and Brauer–Nesbitt, two "finitely ramified" ℓ-adic representations of $G_{\mathbb{Q}}$ agree up to semi-simplification if and only if their local L-factors agree at all but finitely many primes. In particular, the L-function $L(s, -)$ from Definition 8.2.1, at least "up to finitely many factors", is a "complete and well-defined" invariant for "such representations up to semi-simplification".

Remark 8.5.10. Semi-simplification and "finite factor-fudging" *should* be unnecessary (in view of Tate's global semi-simplicity conjecture), but to be safe (unconditionally speaking), we allow them. In any case, they are convenient.

Returning to $V_{\boldsymbol{c}}$, we now give some alternative descriptions of $M_{\boldsymbol{c}}$ for $m \in \{4, 5, 6\}$. Here we let "$\approx$" denote "equality up to finitely many Euler factors". (Experts may well know more precise information, either in general or in our specific situations; e.g. for item (1) below, see the discussion in Appendix A surrounding Proposition A.1.1.)

(1) If $m \in \{4, 6\}$, then there exists an abelian variety $A_{\boldsymbol{c}}$ of dimension $(\dim M_{\boldsymbol{c}})/2 \in \{1, 5\}$ such that $L(s, V_{\boldsymbol{c}}) \approx L(s, A_{\boldsymbol{c}})$. For $m = 4$, we can take $A_{\boldsymbol{c}}$ to be the Jacobian $J(V_{\boldsymbol{c}})$ of the genus one curve $V_{\boldsymbol{c}}$, as noted on [HB98, p. 680]. For $m = 6$, we can take the Albanese $A(F(V_{\boldsymbol{c}}))$ of the Fano surface $F(V_{\boldsymbol{c}})$ of lines on $V_{\boldsymbol{c}}$.[6]

[6]See [DLR17, Theorem 4.1] for a computational perspective, explaining at least the Fano surface connection. For a more complete discussion, see [Rei72, Appendix 4.3] or [Mur74].

(2) If $m = 5$, then there exists a 6-dimensional Artin representation $\rho_{\boldsymbol{c}}$ associated to $V_{\boldsymbol{c}}$ such that $L(s, V_{\boldsymbol{c}}) \approx L(s, \rho_{\boldsymbol{c}})$. At least at all but finitely many primes, this is explained in [Man86]. To include all primes, one could apply [Poo17, Proposition 9.2.6] to $(V_{\boldsymbol{c}})_{\overline{\mathbb{Q}}}$. (The point is that the ℓ_0-adic cycle class homomorphism [Poo17, (7.6.2)] for $(V_{\boldsymbol{c}})_{\overline{\mathbb{Q}}}$ is an isomorphism, since the cohomological Brauer group of $(V_{\boldsymbol{c}})_{\overline{\mathbb{Q}}}$ vanishes [Poo17, Corollary 6.9.11]. See also [Jah14, Remark III.4.10, Theorem III.7.9(ii), Lemma III.8.4, and Example III.8.7(ii)].)

Strictly speaking, such observations are unnecessary. But the question of reciprocity for abelian varieties (see e.g. [Gro16, BCGP21]) and Artin motives is currently better studied than that for hypersurfaces—hence worth mentioning (especially for $m \geq 5$, where automorphy seems to remains open for $M_{\boldsymbol{c}}$ in general; e.g. for $m = 6$, see Appendix A for some discussion).

Remark 8.5.11. Recall the non-degenerate pairing $\psi \colon H^{m_*}(V_{\boldsymbol{c}}) \times H^{m_*}(V_{\boldsymbol{c}}) \to H^{2m_*}(V_{\boldsymbol{c}})$.

(1) If $m_* = 1$, then ψ is essentially the Weil pairing on the elliptic curve $J(V_{\boldsymbol{c}})/\mathbb{Q}$. Here ψ is *symplectic* and $\bigwedge^2 H^1(V_{\boldsymbol{c}}) \cong \mathbb{Q}_{\ell_0}(-1)$ is trivial up to Tate twist (cf. the fact that $\tilde{\alpha}_p \tilde{\beta}_p = 1$ if $p \nmid N$).

(2) If $m_* = 2$, then ψ is the *symmetric* intersection pairing on the cubic surface $V_{\boldsymbol{c}}$. Since 2 is even, $H^2(V_{\boldsymbol{c}}, \mathbb{Q}_{\ell_0}(1)) \cong \operatorname{Pic}(V_{\boldsymbol{c}} \times \overline{\mathbb{Q}}) \otimes \mathbb{Q}_{\ell_0}$ splits as $[K_{V_{\boldsymbol{c}}}]\mathbb{Z} \oplus K_{V_{\boldsymbol{c}}}^{\perp}$. It is the *primitive* part $K_{V_{\boldsymbol{c}}}^{\perp} = [-H]^{\perp}$ (which is isomorphic to $M_{\boldsymbol{c}}$) that defines $L(s, V_{\boldsymbol{c}})$ (a degree 6 Artin L-function).

(3) If $m_* = 3$, then ψ is again *symplectic*. It is likely closely related to the Weil pairing associated to $(A_{\boldsymbol{c}}, \lambda_{\boldsymbol{c}})$, where $\lambda_{\boldsymbol{c}} \colon A_{\boldsymbol{c}} \to A_{\boldsymbol{c}}^{\vee}$ denotes a certain principal polarization defined in terms of $V_{\boldsymbol{c}}$. (For one possible construction of $\lambda_{\boldsymbol{c}}$, see [DLR17, Remark 4.2, par. 2, involving a "difference morphism" on $F(-)$].)

8.5.2 A statistical discussion of our families of L-functions

We now discuss the expected statistical nature of our families of L-functions. Fix F with $m \in \{4, 6, 8, \dots\}$, and for convenience, assume Conjecture 8.2.11 (HW2) (but not (HWSp)).

The Sarnak–Shin–Templier framework

Let $\pi_{\boldsymbol{c}} := \pi_{M_{\boldsymbol{c}}}$. Then $\pi_{\boldsymbol{c}}$ should be cuspidal for almost all $\boldsymbol{c}$. Indeed, the proof of Proposition 8.3.12 shows (unconditionally) that $p^{-m} \sum_{\boldsymbol{c} \in \mathbb{F}_p^m} \mathbf{1}_{p \nmid F^{\vee}(\boldsymbol{c})} \cdot |\tilde{\lambda}_{\boldsymbol{c}}(p)|^2 = 1 + O(p^{-\delta})$ as $p \to \infty$. It follows that as $Z \to \infty$, we have

$$\begin{aligned} 1 + O(Z^{-\delta'}) &= \frac{1}{Z^{\delta} \cdot (2Z)^m} \sum_{p \leq Z^{\delta}} (\log p) \cdot \sum_{\boldsymbol{c} \in [-Z,Z]^m} \mathbf{1}_{F^{\vee}(\boldsymbol{c}) \neq 0} \cdot |\tilde{\lambda}_{\boldsymbol{c}}(p)|^2 \\ &\geq 1 + O_{\epsilon}(Z^{-\delta/2+\epsilon}) + \frac{1}{(2Z)^m} \sum_{\boldsymbol{c} \in [-Z,Z]^m} \mathbf{1}_{F^{\vee}(\boldsymbol{c}) \neq 0} \cdot \mathbf{1}_{\pi_{\boldsymbol{c}} \text{ is not cuspidal}}, \end{aligned}$$

by a "representation-theoretic" analysis of (the poles at $s = 1$ of) the L-functions $L(s, M_{\boldsymbol{c}} \otimes M_{\boldsymbol{c}}^{\vee})$ via (HW2), Observation 8.2.15, and [IK04, §5.6's Exercise 6 and §5.7's Theorem 5.15]. So $\#\{\boldsymbol{c} \in [-Z, Z]^m : F^{\vee}(\boldsymbol{c}) \neq 0, \text{ and } \pi_{\boldsymbol{c}} \text{ is not cuspidal}\} \ll Z^{m-\delta''}$.

Similarly (still under (HW2)), via other (unconditional) local moments from the proof of Proposition 8.3.12, one can statistically analyze the poles of $L(s, M_{\boldsymbol{c}} \otimes M_{\boldsymbol{c}}), L(s, M_{\boldsymbol{c}} \wedge M_{\boldsymbol{c}})$ at $s = 1$. Such an analysis (when combined with the previous paragraph) reveals the family $\boldsymbol{c} \mapsto \pi_{\boldsymbol{c}}$ to be *essentially* cuspidal, self-dual, and symplectic (in the sense of [SST16, p. 538, (i)–(iii)]), so that $\operatorname{ord}_{s=1} L(s, \pi_{\boldsymbol{c}}, \bigwedge^2) = -1$ almost always.

Remark 8.5.12. The computations above are closely related to "vertical" or "local" monodromy. Since V is smooth of even dimension $m - 2 \in \{2, 4, \dots\}$, with hyperplane sections $V_{\boldsymbol{c}}$ of odd dimension $m_* \in \{1, 3, \dots\}$, a general result of Deligne (see [Kat04, Introduction, pp. 1–2], around the line "For n odd, the monodromy group G_d is…") shows that the "Zariski closure of the monodromy of the local system"

$$\boldsymbol{c} \mapsto H^{m_*}(V_{\boldsymbol{c}})/H^{m_*}(V) = H^{m_*}(V_{\boldsymbol{c}})$$

of rank $N_1 = \frac{2^{m_*+2}+2(-1)^{m_*}}{3} \in \{2, 10, \dots\}$ on the space of "smooth, degree $d = 1$, hypersurface sections" (in [Kat04]'s setup[7]) is the "full symplectic group" $\mathrm{Sp}(N_1)$. (Cf. [Kat13, §8, "hypersurface examples"], which are of the same spirit.)

In particular, by Deligne–Katz equidistribution ([Kat13, Theorem 5.1], as applied in [SST16, §2.11]), our family has Sato–Tate group $\mathrm{Sp}(N_1, \mathbb{C})$ in the sense of [SST16].

In any case, by [SST16, Conjecture 2], the low-lying L-function zeros associated to our family should have symmetry type consisting of $\mathrm{SO}_{\mathrm{even}}$ and $\mathrm{SO}_{\mathrm{odd}}$; cf. the discussion on [SST16, p. 549] for Dwork families of odd degree, i.e. with "n even". Thus in *all* RMT-type predictions, the family $\boldsymbol{c} \mapsto L(s, \pi_{\boldsymbol{c}})$ should have "symmetry type" composed of $\mathrm{SO}_{\mathrm{even}}$ and $\mathrm{SO}_{\mathrm{odd}}$; cf. [SST16, pp. 540–541, paragraph discussing "moments of L-values"]. (The root numbers are probably evenly distributed, but their distribution is actually irrelevant to our main results; see Remark 8.5.15 below.)

Remark 8.5.13. RMT-type predictions should apply equally well to all "natural" parameterizations of a given family, in the spirit of Remark 8.3.16(3) and [SST16, p. 535, Remark (i); and p. 560, second paragraph after (25)]. (See e.g. Conjecture 8.2.28 (RA1), a mean-value prediction over certain "nearly homogeneous" regions $\mathcal{B}(\boldsymbol{Z}) \cap \{\boldsymbol{a} \bmod n_0\}$, where one can think of $\mathcal{B}(\boldsymbol{Z})$ as being "truly homogeneous, but *with small coefficients*".) In this regard, *multi-parameter* families raise some interesting questions (involving "lopsidedness" and "singularities" of weights) that are subtler than those for *single-parameter* families.

Remarks on (R1)–(R2) and (RA1)

In §8.3.3, we applied the RMT-based heuristic behind [CFZ08, §5.1, (5.6)] to state certain Ratios Conjectures (R1)–(R2) and (RA1).

Remark 8.5.14. For a discussion of uniformity in σ, t, see [CS07, (2.11b)–(2.11c)]. In particular, (2.11b) suggests that in (R1)–(R2) and (RA1), we could take $\sigma - 1/2$ as small as $\Omega(1/\log Z)$. But it is cleaner for us to fix $\sigma > 1/2$—which is good enough for us anyways, since we assume *power-saving* error terms in (R1)–(R2) and (RA1).

[7] we take the "universal family" of smooth hyperplane sections, but a "sufficiently general Lefschetz pencil" would also suffice according to [Kat04]

Remark 8.5.15. The Ratios Conjectures involve not just the low-lying zeros of $L(s, \pi_{\boldsymbol{c}})$, but rather "all zeros up to height ≈ 1" (morally). Nonetheless, one still expects a remarkable degree of universality: see e.g. [CFZ08, Conjectures 5.3–5.4] for orthogonal examples of the Ratios Conjectures, where the presence of "L's in the denominator" on the left-hand side leads to "polar factors of $\zeta(2s)$" on the right-hand side, just as in (R1)–(R2) and (RA1).

Interestingly, root numbers and functional equations play no role in the "recipe" for $1/L$. This somehow reflects the naive intuition that $1/L$ is "more random" than L. In other aspects, though, $1/L$-moments over orthogonal families seem loosely analogous to L-moments over symplectic families (and vice versa).

8.6 Miscellaneous writeups

8.6.1 Poles given by the original variety

Now assume F is diagonal. Then one can "compute" $L(s, V)$; see Lemma 8.6.7 below.

Definition 8.6.1. Given a nontrivial multiplicative character $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}^\times$, define the standard Gauss sum $g(\chi) := \sum_{x \in \mathbb{F}_q} \chi(x) e_p(\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x))$, and let $\tilde{g}(\chi) := g(\chi)/q^{1/2}$.

Proposition 8.6.2 (See e.g. [IR90, Chapter 10, Theorem 2])**.** *Let q be a power of a prime $p \nmid 3F_1 \cdots F_m$. If $q \equiv 2 \bmod 3$, then $E(q) = 0$. If $q \equiv 1 \bmod 3$, and $\chi_3 = \chi_{3,q}\colon \mathbb{F}_q^\times \to \mu_3 \subseteq \mathbb{C}^\times$ denotes* either *of the two multiplicative characters of order 3, then*

$$E(q) = q^{-1+m/2} \sum_{\boldsymbol{a}}{}' \chi_3(F_1^{-a_1} \cdots F_m^{-a_m}) \tilde{g}(\chi_3^{a_1}) \cdots \tilde{g}(\chi_3^{a_m}),$$

where $a_i \in \{1, -1\}$ and $3 \mid a_1 + \cdots + a_m$.

Although over each prime p there are two possible choices of "compatible" $\chi_{3,q}$ (such that $\chi_{3,q^r} = \chi_{3,q} \circ N_{\mathbb{F}_{q^r}/\mathbb{F}_q}$ whenever $q \equiv 1 \bmod 3$), the cyclotomic field $K := \mathbb{Q}(\zeta_3)$ itself (in which the χ are valued) only has finitely many automorphisms—and thus provides a way to glue different p together, via the cubic residue symbol $\chi_{3,\wp}$.

Definition 8.6.3. For a prime $\wp \nmid 3$ of K with residue field $k := \mathcal{O}_K/\wp$, and a unit residue $x \in k^\times$, let $\chi_{3,\wp}(x) \in \mu_3$ with $\chi_{3,\wp}(x) \equiv x^{(N\wp - 1)/3} \bmod \wp$. Then, for each integer $r \geq 1$ and given field extension ℓ/k of degree r, let $\chi_{3,\wp^r} := \chi_{3,\wp} \circ N_{\ell/k}$ be the unique character $\ell^\times \to \mu_3$ extending $\chi_{3,\wp}\colon k^\times \to \mu_3$.

Since $g(\chi)$ is well-defined (i.e. independent of the realization of $\mathbb{F}_q$), we can make sense of $g(\chi_{3,\wp^r})$ by identifying ℓ with $\mathbb{F}_{N\wp^r}$, i.e. $g(\chi_{3,\wp^r}) := \sum_{x \in \ell} \chi_{3,\wp^r}(x) e_p(\mathrm{Tr}_{\ell/\mathbb{F}_p}(x))$. Now by the previous proposition (though there might be a more conceptual approach phrased in terms of the K-automorphisms of V_K), the indices $\boldsymbol{a}$ decompose into pairs $\{\boldsymbol{a}, -\boldsymbol{a}\}$, each defining a Hecke L-function over K. This is the content of the following result:

Corollary 8.6.4. *For good primes $p \nmid 3F_1 \cdots F_m$, we have*

$$L_p(s, V) := \exp\left((-1)^{m-2} \sum_{r \geq 1} \widetilde{E}(p^r)(p^{-s})^r / r\right) = \prod_{\{\boldsymbol{a}, -\boldsymbol{a}\}}{}' \prod_{\wp \mid p} (1 - \psi_{\boldsymbol{a}}(\wp)(N\wp)^{-s})^{-1},$$

where $\wp$ denotes a prime in $\mathcal{O}_K$ and $\psi_{\boldsymbol{a}}$ is the unique primitive *Hecke character on K satisfying $\psi_{\boldsymbol{a}}(\wp) = \chi_{3,\wp}(\boldsymbol{F}^{-\boldsymbol{a}})(-\tilde{g}(\chi_{3,\wp}^{a_1}))\cdots(-\tilde{g}(\chi_{3,\wp}^{a_m}))$ for all $\wp \nmid 3F_1\cdots F_m$.*

Remark 8.6.5. For the construction of the (classical) Hecke characters $\psi_{\boldsymbol{a}}$, see [Wei52]. (Actually, [Wei52] might only yield a possibly imprimitive character of conductor dividing $(3F_1\cdots F_m)^\infty$, but this suffices.) Strictly speaking, [Wei52, p. 489, Theorem] only addresses the product of g's, but the Artin symbol $\wp \mapsto \chi_{3,\wp}(\boldsymbol{F}^{-\boldsymbol{a}})$ is harmless by class field theory over K (cf. [Wei52, final paragraph of p. 494, regarding the L-function of a diagonal curve]).

Remark 8.6.6. Each $\psi_{\boldsymbol{a}}$ is $\overline{\mathbb{Q}}$-valued, i.e. "algebraic" (also called "type A_0"). In fact, one can show that each "un-normalized" Hecke character $N_{K/\mathbb{Q}}^{m/2}\cdot\psi_{\boldsymbol{a}}$ maps into K.

Proof. Let $T = p^{-s}$. Using the Hasse–Davenport relation [IK04, p. 275, Theorem 11.4] and $\chi_{3,\wp^r} = \chi_{3,\wp}\circ N_{\ell/k}$ as in [Wei49, p. 506, (8)] gives

$$(-1)^m\sum_{r\geq 1}\widetilde{E}(p^r)T^r/r = -{\sum_{\boldsymbol{a}}}' \log(1-\chi_{3,\wp}(\boldsymbol{F}^{-\boldsymbol{a}})(-\tilde{g}(\chi_{3,\wp}^{a_1}))\cdots(-\tilde{g}(\chi_{3,\wp}^{a_m}))T),$$

for either choice of $\wp \mid p$, if $p\equiv 1 \bmod 3$. Similarly, if $p \equiv 2 \bmod 3$ (and $\wp = (p)$), then

$$(-1)^m\sum_{r\geq 1}\widetilde{E}(p^r)T^r/r = -\frac{1}{2}{\sum_{\boldsymbol{a}}}' \log(1-\chi_{3,\wp}(\boldsymbol{F}^{-\boldsymbol{a}})(-\tilde{g}(\chi_{3,\wp}^{a_1}))\cdots(-\tilde{g}(\chi_{3,\wp}^{a_m}))T^2),$$

in which case the automorphism $x\mapsto x^p$ of k equates the contributions from $\boldsymbol{a}$ and $p\boldsymbol{a} \equiv -\boldsymbol{a} \bmod 3$ (because $\boldsymbol{F}^{-\boldsymbol{a}} \equiv \boldsymbol{F}^{-p\boldsymbol{a}} \bmod p$ and $g(\chi_{3,\wp}) = g(\chi_{3,\wp}^p)$), thus removing the factor of $1/2$. The desired formula for $L_p(s,V)$ follows. $\square$

Lemma 8.6.7. *With the above notation, $L(s,V) = \prod'_{\{\boldsymbol{a},-\boldsymbol{a}\}} L(s,\psi_{\boldsymbol{a}})$. Furthermore, $L(s,V)$ has a pole at $s=1$ of total order $r_F \geq 0$ equal to the number of pairs $\{\boldsymbol{a},-\boldsymbol{a}\}$ such that $F_1^{-a_1}\cdots F_m^{-a_m}\in(\mathbb{Q}^\times)^3$ and $\sum_{i\in[m]} a_i = 0$.*

Remark 8.6.8. For m odd, or for typical F, we have $r_F = 0$. For F Fermat with m even, $r_F = \frac{1}{2}\binom{m}{m/2} = \binom{m-1}{m/2-1}$ (so that e.g. $r_F = 3$ if $m=4$, and $r_F = 10$ if $m = 6$).

Proof. First, to fully prove the factorization on the nose, combine [And86, Theorem 8(II), Theorem 6, and Corollary 5.7.2]; here Theorem 8(II) identifies $M_V(-1)$ (and hence M_V) as a motive "potentially of complex multiplication type" (which can therefore be analyzed by Theorem 6 and Corollary 5.7.2).

It remains to compute $r_F := -\operatorname{ord}_{s=1} L(s,V)$. But work of Hecke (see e.g. [IK04, Theorem 3.8] for the imaginary quadratic case we are in) immediately implies that

$$r_F = \#\{\text{pairs } \{\boldsymbol{a},-\boldsymbol{a}\} : \psi_{\boldsymbol{a}} \text{ is trivial}\}.$$

Now fix $\boldsymbol{a}$. Since $\tilde{g}(\chi_3)\tilde{g}(\chi_3^{-1}) = 1$, the "defining formula" for $\psi_{\boldsymbol{a}}$ (from the previous corollary) simplifies, telling us that $\psi_{\boldsymbol{a}}(\wp) = \chi_{3,\wp}(\boldsymbol{F}^{-\boldsymbol{a}})(-\tilde{g}(\chi_{3,\wp}))^{\sum a_i}$ for all but finitely many primes $\wp$. If $\boldsymbol{F}^{-\boldsymbol{a}}\in(\mathbb{Q}^\times)^3$ and $\sum a_i = 0$, then certainly $\psi_{\boldsymbol{a}}$ must be trivial.

Conversely, suppose $\psi_{\boldsymbol{a}}$ is trivial; then $\chi_{3,\wp}(\boldsymbol{F}^{-\boldsymbol{a}})(-\tilde{g}(\chi_{3,\wp}))^{\sum a_i} = \psi_{\boldsymbol{a}}(\wp) = 1$ for almost all $\wp$. For such $\wp$, cubing yields $(-\tilde{g}(\chi_{3,\wp}))^{3\sum a_i} = 1$. If $\sum a_i \neq 0$, then $-\tilde{g}(\chi_{3,\wp})$ would be restricted to lie in a finite set, contradicting the known equidistribution of Kummer sum angles. (Alternatively, it should suffice to use Stickelberger's factorization [Lan90, Chapter 1, Theorem 2.2] of $g(\chi_3)$.) Thus $\sum a_i = 0$, so $\chi_{3,\wp}(\boldsymbol{F}^{-\boldsymbol{a}}) = 1$ for almost all $\wp$, from which Chebotarev implies $\boldsymbol{F}^{-\boldsymbol{a}} \in (\mathbb{Q}^\times)^3$, completing the proof. $\square$

Remark 8.6.9. For $m = 4$, a simpler treatment is possible via Artin representation theory [Jah14, pp. 213–216], since for each $\boldsymbol{a}$, the g's cancel out: $\sum a_i$ is 0 mod 3 and m mod 2, and thus 0. (For example, if $F = x_1^3 + \cdots + x_4^3$, then $L(s, V) = \zeta_{\mathbb{Q}(\zeta_3)}(s)^3$.) For $m = 6$, though, we need to address the "transcendental motive" (in the language of [SD14]) associated to $\boldsymbol{a} = \pm\boldsymbol{1}$, even though it does not ultimately contribute poles.

8.6.2 Typical analytic ranks, versus r_F

For diagonal F, Lemma 8.6.7 expresses $L(s, V)$ as a product of Hecke L-functions and determines the order r_F of the pole at $s = 1$. (For arbitrary $\mathbb{P}^{m-1}_{\mathbb{Q}}$-smooth F, the discussion below should still hold conditionally on automorphy for $L(s, V)$.) Naively, r_F should have something to do with special subvarieties of V (cf. Chapter 6). This is partly true, but the full story, via the global Tate conjecture for codimension-$(1 + m_*)/2$ cycles on V modulo suitable equivalence (e.g. $\operatorname{Pic}(V)$ for $m_* = 1$, when V is a cubic surface), seems to involve other cycles on V as well. (One should perhaps work with homological or numerical equivalence. We will not be too precise about such technical questions on algebraic cycles, including base change or rationality issues.)

When $F^\vee(\boldsymbol{c}) \neq 0$, let $r_{\boldsymbol{c}}$ be the *analytic rank*, i.e. central order of vanishing, of $L(s, \pi_{\boldsymbol{c}})$. The following may not be strictly necessary, but seems good to discuss:

Conjecture 8.6.10 (Nagao-type conjecture). *If $m \in \{4, 6\}$, then for almost all $\boldsymbol{c} \in \mathbb{Z}^m$ with $F^\vee(\boldsymbol{c}) \neq 0$, we have $r_{\boldsymbol{c}} \in \{r_F, r_F + 1\}$.*

Remark 8.6.11. For typical F, Lemma 8.6.7 says $r_F = 0$, and we certainly typically expect that $r_{\boldsymbol{c}} \in \{0, 1\}$. For specific F, comparing typical $r_{\boldsymbol{c}}$ with r_F might require BSD-on-average or similar. As some algebraic evidence for $m = 4$ (i.e. $m_* = 1$), [Sta16] has shown (for any *fixed* smooth cubic surface V) that the restriction map $\operatorname{Pic}(V) \to \operatorname{Pic}(V_{\boldsymbol{c}})$ is *typically injective* as $\boldsymbol{c}$ varies, so that typically $r_{\boldsymbol{c}} \geq r_F$.

For $m = 6$ (i.e. $m_* = 3$), is it true that the relevant restriction map is typically injective? (If not, then the conjecture above must be modified according to the dimension of the kernel.)

We now analyze the maximal linear $\mathbb{Q}$-subvarieties for $m \in \{4, 6\}$ when F is Fermat, and explain (with the $m = 4$ example) why robust numerical testing requires at least a bit of care.

Example 8.6.12. For $m = 4$ and F Fermat, the genus one curve $V_{\boldsymbol{c}}$ has three rational parametric points coming from the three rational lines of the cubic surface V, so (taking one of the three to be the origin) we certainly expect $J(V_{\boldsymbol{c}})$ to typically have (algebraic) rank at least $3 - 1 = 2$. Naive random sampling of $V_{\boldsymbol{c}}$ (with $\|\boldsymbol{c}\|_\infty \leq 6$) leads to a rank distribution of $[5, 23, 154, 166, 35, 3, 0, 0, \dots]$, at first suggesting typical (analytic) ranks $2, 3$.

However, $r_F = 3$. A more careful computation, sampling with $\|\boldsymbol{c}\|_\infty \leq 10$ and avoiding the locus $\prod_{i<j}(c_i - c_j) = 0$ (which arose from guesswork—but it would be good to check how $\operatorname{Pic}(V) \to \operatorname{Pic}(V_{\boldsymbol{c}})$ behaves for such $\boldsymbol{c}$), gives a rank distribution of $[0, 0, 10, 70, 42, 1, 0, 0, \dots]$,

suggesting instead typical (algebraic) ranks $3, 4$, in line with Conjecture 8.6.10. Probably we missed a typical third point in the image of $\operatorname{Pic}(V) \to \operatorname{Pic}(V_{\boldsymbol{c}})$, which could be computed in principle from [Bro09, p. 139, §8.3.1] (which explicitly describes $\operatorname{Pic}(V)$, or equivalently—by [Jah14, top of p. 211]—$\operatorname{Pic}(V_{\overline{\mathbb{Q}}})^{G_{\mathbb{Q}}}$, since $V(\mathbf{A}_{\mathbb{Q}}) \neq \emptyset$).

(Disclaimer: Nearly all of the ranks computed as part of the data should be correct, but a small portion may be incorrect due to compromises made by typical algorithms, or due to the fact that verifying large analytic ranks remains an open problem.)

Example 8.6.13. For $m = 6$ and F Fermat, $V_{\boldsymbol{c}}$ has 15 rational parametric lines coming from the $5 \cdot 3 \cdot 1 = 15$ rational 2-planes of the cubic fourfold V (one for each partition $\mathcal{J}$ of $[6]$ into $2 + 2 + 2$). These lines on $V_{\boldsymbol{c}}$ (i.e. points on $F(V_{\boldsymbol{c}})$) correspond to points on $A(F(V_{\boldsymbol{c}}))$ (taking one of the points to be the origin of $A_{\boldsymbol{c}}$); *what are the relations* between these points?

By [CG72, p. 286], the analysis might involve configurations like $V_{\boldsymbol{c}} \cap \{x_1 + x_2 = 0\} \cap \{x_3 + x_4 + x_5 + x_6 = 0\}$ (a union of three coplanar lines). There are $\binom{6}{2} = 15$ such *triangles*. Each line lies in 3 triangles. Thus we can sequentially "remove a line from an existing triangle" ≥ 5 times (so that at the end, the remaining ≤ 10 lines "span" or "triangulate" the rest). This may be suboptimal; it is a question of linear algebra to find a "triangle-reduced basis".

What other "typical cycles" on $V_{\boldsymbol{c}}$ (coming from V) are there?

Remark 8.6.14. In both cases, at least some of the rational parametric cycles on $V_{\boldsymbol{c}}$ come from linear spaces on V of the largest possible dimension (in the sense of §6.3).

Remark 8.6.15. For $m = 6$, it may be worth testing analytic ranks numerically, perhaps through the logarithmic derivative L'/L. (Completely computing local L-factors of cubic threefolds seems to quickly get expensive. But to test ranks, one should only need to work with small moduli.) Notably, [SD67, bottom of p. 290, with $r = 2$] provides a relevant variant of BSD "supported by unpublished results of Bombieri and Swinnerton-Dyer for the cubic threefold" [SD67, p. 291]; I cannot tell if these "unpublished results" are more numerical or theoretical, however.

Chapter 9

Variations

§9.1 discusses problems inspired by Hypothesis HW, GRH on average, and large sieves, in the spirit of §4.1.

§9.2 discusses questions related to "perturbing" the delta method—which seems to be an interesting direction of research (cf. [MV19]).

In §9.3, we speculate on some Diophantine problems that may or may not be within reach under standard hypotheses similar to those in Chapter 8.

9.1 Problems inspired by Hypothesis HW

9.1.1 Problems with a cubic flavor

The cubic convexity barrier

Beating the Hua bound (or "breaking convexity" as [Woo95] might say) lies in the following family of (probably equivalent) problems:

(1) To show that the (smoothly weighted) count of solutions $\boldsymbol{x} \in \mathbb{Z}^6$ to $x_1^3 + \cdots + x_6^3 = 0$, in the region $\|\boldsymbol{x}\| \ll X$, is $O(X^{7/2-\delta_3})$ for some $\delta_3 > 0$.

(2) To show that the analogous smooth count for $x_1^3 + \cdots + x_8^3 = 0$ satisfies a Hardy–Littlewood asymptotic $cX^5 + O(X^{5-\delta_4})$ with a power saving $\delta_4 > 0$.

(3) In general, for $s \geq 4$, to show that the analogous count for $x_1^3 + \cdots + x_{2s}^3 = 0$ satisfies a Hardy–Littlewood asymptotic of the form $cX^{2s-3} + O(X^{3s/2-1-\delta_s})$ for some $\delta_s > 0$.

These "even Fermat bounds" would all follow from an improvement of the current generic cubic Weyl sum bound of $O_\epsilon(X^{3/4+\epsilon})$, which appears quite difficult to beat; see e.g. [HB10] assuming *abc*. In any case, the beauty of [Hoo86b, Hoo97, HB98] lies in the averaging over arcs (and the further averaging over moduli going beyond Kloosterman), which is morally independent of the issue of pointwise Weyl bounds.

The equivalent family above would also imply similar bounds for odd numbers of variables, e.g. an upper bound for "$2s = 5$" (a problem raised by [Bom09]). At first glance, there does not appear to be an *equivalent* Diophantine problem involving 4 variables (where "$s = 2$"), or any odd number "$2s$", although the general "restricted arcs" form of the 4 variable problem (see problem (5) below) could in principle pare away most, if not all, but the most extreme minor arcs for $s \geq 3$.

Other problems with a cubic flavor

Besides the "familiar" or "classical" Fermat cubic problems above, other natural problems in the same vein include the following:

(4) To show that each smooth projective cubic surface $V(F)$ over $\mathbb{Q}$ has $O_{F,\epsilon}(X^{3/2+\epsilon})$ solutions $\boldsymbol{x} \ll X$ away from its rational lines—or at least $O_F(X^{12/7-\delta})$, to beat [Sal15]'s bound. (Note however that one of the nice features of [Sal15] is uniformity over F, which we do not consider here.)

(5) To show that $x_1^3 + \cdots + x_4^3 = 0$ has $O_\epsilon(X^{2+\epsilon})$ solutions, in a way that generalizes to show that uniformly over $M \leq X^{3/2}$, we have

$$\int_{\theta \in \mathfrak{M}(M)} d\theta\, |T(\theta)|^4 \ll_\epsilon X^\epsilon \cdot (X + M^2/X),$$

where $T(\theta) := \sum_{|x| \leq X} e(\theta x^3)$ and $\mathfrak{M}(M) := \bigcup_{q \leq M} \{\theta \in \mathbb{R}/\mathbb{Z} : |q\theta - a| \leq M/X^3\}$.

(Unconditionally, [Brü91] obtained $O_\epsilon(X^\epsilon) \cdot (X + M^2/X + M^{7/2}/X^3)$.)

(6) For $k \in \{3, 4\}$, to show that the equation $x_1^2 + x_2^3 + x_3^k = y_1^2 + y_2^3 + y_3^k$, with each *monomial* restricted to be of size at most N, satisfies a (smoothly weighted) Hardy–Littlewood asymptotic of the form $cN^{1+2\delta_k} + O_\epsilon(N^{1+\delta_k+\epsilon})$, where $\delta_k := (1/2 + 1/3 + 1/k) - 1$.

(7) For $k \in \{3, 4\}$, to show that $x_1^2 + x_2^3 + x_3^k$ positively represents all $n \leq N$ but an exceptional set of size $O_\epsilon(N^{1-\delta_k+\epsilon})$—which [Brü91] proved unconditionally for $k = 5$, and obtained weaker bounds towards for $k = 3, 4$.

Remark 9.1.1. Progress on (4) would already be interesting for "somewhat general" (e.g. diagonal) F's. But for cubic surfaces $V(F)$ with $\mathbb{Q}$-lines, unconditional bounds beyond [Sal15] are already known—e.g. $O_\epsilon(X^{4/3+\epsilon})$ for the Fermat cubic.

(Note that general diagonal $V(F)$'s have no $\mathbb{Q}$-lines.)

Remark 9.1.2. The bound in (5), if true, may or may not be optimal (up to ϵ). It is optimal for $M = X^{3/2}$, at least.

Remark 9.1.3. The asymptotic error term $O_\epsilon(N^{1+\delta_k+\epsilon})$ in (6), if valid, would likely be optimal up to ϵ, due to "trivial loci" such as $\boldsymbol{x} = \boldsymbol{y}$. It could be interesting to try to obtain a "secondary term" asymptotic that precisely captures the influence of such loci; cf. [Vau15, Theorem 1.4] regarding the mixed ternary forms $x_1^2 + x_2^2 + x_3^k$ for $k \geq 3$.

Remark 9.1.4. By mimicking [Brü91], one can show that the bound in (5) implies the "exceptional set bound" in (7); see [Wan21e, Theorem 1.12] for details. It is also reasonable to expect "(5) to imply (6)" and "(6) to imply (7)"—at least morally—but such expectations would need to be carefully checked.

Conditional approaches

The bounds in (1)–(3), and the bound in (4) for diagonal F, can be established conditionally under Hypothesis HW for smooth projective cubic hypersurfaces of dimension $3, 5, 2s-3$, and 1, respectively, in certain families. See [Wan21c, Theorem 1.28] (whose philosophy we sketched in §4.1) for a unified treatment, which also clarifies how one may "relax" Hypothesis HW.

Similarly, under Hypothesis HW for smooth hyperplane sections of $V(x_1^3 + \cdots + x_4^3)$, one can mimic [Wan21c, Theorem 1.28] to prove the bound in (5), and consequently the bound in (7). Again, some "relaxation" of Hypothesis HW is possible; see [Wan21e, Theorem 1.6] for details. (The situation for (6) may well be similar, but would need to be carefully analyzed and written up.) So at least in (5) and (7), a tiny improvement over [Brü91]'s bounds should already be possible using existing general automorphic large sieve inequalities; see [Wan21e, Remark 1.9] for more details. But truly satisfying progress on any of (1)–(7) will probably (and hopefully) require significant new ideas.

Remark 9.1.5. In the absence of progress on GRH itself, the "GRH-on-average philosophy" provides one of the most alluring approaches towards (1)–(7).

However, other ideas could also plausibly lead to interesting progress on (1)–(7). For example, (4) might be susceptible to clever slicing methods (e.g. those developed by Salberger or Heath-Brown), especially when combined with advances in the theory of irrational varieties (including upper-bound sieves and elliptic-curve statistics).

From another direction, Bourgain and Demeter (among others) have applied decoupling theory on curves (for instance) to prove interesting point-counting bounds, often even in the absence of translation invariance. Although the most striking applications of decoupling to *classical* Diophantine problems so far seem to have been restricted to degree ≥ 5 or codimension ≥ 2, the full power of decoupling remains to be understood, and one cannot yet rule out the possibility of applying decoupling (or one of its non-archimedean allies, as developed by Wooley and others) towards (1)–(7).

9.1.2 On families of quadrics

Aside from cubic problems, *quadratic* problems—especially those in *families*—provide an immediate opportunity for deeper exploration via the delta method. Such problems might also play a role in, or at least relate to, Diophantine analysis on varieties that are *not* complete intersections.

Example 9.1.6. Consider Manin's conjecture for the conic bundle $a_1x_1^2 + a_2x_2^2 + a_3x_3^2 = 0$ in $\mathbb{P}^2 \times \mathbb{P}^2$. ([LB15, BBS18] have expressed the opinion that this seems "out of reach" of existing techniques; the analog in $\mathbb{P}^3 \times \mathbb{P}^3$ is known [BHB20]. On the other hand, Heath-Brown may have a forthcoming proof; see [HB22, final paragraph of §1].)

The whole variety E itself (i.e. the total space) is probably not a complete intersection. But the "left-hand side" above certainly defines a *family* of ternary quadratic forms F_0.

Now recall that point counting on a *ternary* affine quadric $F_0 = b$ with F_0 fixed and definite—or with (F_0, b) fixed and F_0 indefinite—is basically understood, but varying F_0 or b in general can get tricky [FI13]. Nonetheless, in the *homogeneous* case $b = 0$ above, the situation might be better—with a certain interesting family of classical (Dirichlet or Artin) L-functions plausibly relating to Manin's conjecture for E.

9.1.3 On intersections of quadrics

For simplicity, we discuss only individual smooth projective complete intersections $W := V(F_1, \ldots, F_R)$ in $\mathbb{P}^{m-1}_{\mathbb{Q}}$ with $m \geq 5$ and $R := 2$ and $\deg F_1 = \cdots = \deg F_R = d := 2$, though there are certainly many other interesting (affine or projective) varieties one could consider (individually or on average).

Given $\boldsymbol{F} := (F_1, \ldots, F_R)/\mathbb{Z}$, fix $w \in C_c^\infty(\mathbb{R}^m)$ supported away from $\mathbf{0}$. For $X \geq 1$, let $N_{\boldsymbol{F},w}(X) := \sum_{\boldsymbol{x} \in \mathbb{Z}^m} w(\boldsymbol{x}/X) \cdot \mathbf{1}_{\boldsymbol{F}(\boldsymbol{x})=\mathbf{0}}$, and if $m = 5$, let $N'_{\boldsymbol{F},w}(X)$ denote the corresponding weighted count restricted $\boldsymbol{x} \in \mathbb{Z}^m$ not lying on any $\mathbb{Q}$-line of W.

For references on what is known about the Hasse principle for W, we refer the reader to the introduction of [Vis19]. When $m = 5$, the "Hasse principle up to Brauer–Manin" is unknown but conjectured to hold, and from a quantitative point of view the best we know in general is Salberger's (possibly unpublished) bound $N'_{\boldsymbol{F},w}(X) \ll_{\boldsymbol{F},\epsilon} X^{3/2+\epsilon}$.

Let $D \in \mathbb{Z}[\boldsymbol{c}]$ denote a suitable discriminant polynomial, so that a given hyperplane section $W_{\boldsymbol{c}}$ of W is singular if and only if $D(\boldsymbol{c}) = 0$. Let

$$S_{\boldsymbol{F},\boldsymbol{c}}(q) := \sideset{}{'}\sum_{\boldsymbol{a} \in (\mathbb{Z}/q)^R} \sum_{\boldsymbol{x} \in (\mathbb{Z}/q)^m} e_q(\boldsymbol{a} \cdot \boldsymbol{F}(\boldsymbol{x}) + \boldsymbol{c} \cdot \boldsymbol{x})$$

(cf. the complete exponential sums defined in [HBP17]), where we restrict $\boldsymbol{a}$ to be *primitive* modulo q. Then let $\widetilde{S}_{\boldsymbol{F},\boldsymbol{c}}(q) := q^{-(m+R)/2} S_{\boldsymbol{F},\boldsymbol{c}}(q)$, let $\Phi(\boldsymbol{F}, \boldsymbol{c}, s) := \sum_{q \geq 1} \widetilde{S}_{\boldsymbol{F},\boldsymbol{c}}(q) q^{-s}$, and let $\Psi_1(\boldsymbol{F}, \boldsymbol{c}, s) = \sum_{q \geq 1} b_{\boldsymbol{F},\boldsymbol{c}}(q) q^{-s}$ denote a first-order approximation of $\Phi(\boldsymbol{F}, \boldsymbol{c}, s)$ in the sense suggested by Definition 3.2.9.

Optimistic *preliminary calculations* (subject to errors or unforeseen difficulties) suggest the following conjecture:

Conjecture 9.1.7. *Fix $\boldsymbol{F}, w, \Psi_1$. Let $Y := X^{d/(1+1/R)} = X^{4/3}$.* Assume that *for all $X \geq 1$ and $Z \geq Y/X = X^{1/3}$, and for all $N \ll Y$ and intervals $I \subseteq [N/2, 2N]$, we have a uniform bound of the form*

$$\sideset{}{'}\sum_{\boldsymbol{c} \in [-Z,Z]^m} \left| \sum_{q \in I} b_{\boldsymbol{F},\boldsymbol{c}}(q) \right|^2 \ll_\epsilon Z^\epsilon \max(Z^m, Y) \cdot N,$$

where we restrict to $\boldsymbol{c}$ with $D(\boldsymbol{c}) \neq 0$. Also assume *the existence of a "fully satisfactory" two-dimensional Kloosterman method. Now let $\theta := \frac{dR(m-1-R)}{2(R+1)} = \frac{2(m-3)}{3}$. Then*

(1) $N'_{\boldsymbol{F},w}(X) \ll_\epsilon X^{\theta+\epsilon} = X^{4/3+\epsilon}$ if $m = 5$;

(2) $N_{\boldsymbol{F},w}(X) \ll_\epsilon X^{\theta+\epsilon} = X^{2+\epsilon}$ if $m = 6$; and

(3) $N_{\boldsymbol{F},w}(X) = cX^{m-dR} + O_\epsilon(X^{\theta+\epsilon}) = cX^{m-4} + O_\epsilon(X^{2(m-3)/3+\epsilon})$ if $m \geq 7$ (with $c = c_{\boldsymbol{F},w}$ being the usual Hardy–Littlewood prediction if $m \geq 8$, and a more complicated constant that also incorporates $\mathbb{Q}$-planes on W if $m = 7$).

Remark 9.1.8. The "elementary GRH on average" assumption above, and that in [Wan21c, Theorem 1.28], should belong to a wider framework of "average hypotheses" over "natural but thin" geometric families. Hopefully at least one such family is sufficiently rich yet tractable to inspire significant new ideas.

Remark 9.1.9. Over $D(\boldsymbol{c}) \neq 0$, the Hasse–Weil L-functions $L(s, W_{\boldsymbol{c}})$ are quite rich. Roughly speaking, each $W_{\boldsymbol{c}}$ (after passing to certain standard ℓ-adic Galois representations) corresponds (up to Tate twist) to an abelian variety of dimension $(m-3)/2$ if $2 \nmid m$, and to an Artin representation of dimension m if $2 \mid m$. (See e.g. [BT16, Theorem 2.1] or [Rei72].)

Remark 9.1.10. The $\boldsymbol{F}$'s considered in [HBP17] define *singular* W's (with singular loci of positive dimension, in fact), so are not included in the conjecture above. But it would be interesting to see how much further one can push [HBP17] under automorphy and GRH for varieties "related to" hyperplane sections of W. (The hyperplane sections $W_{\boldsymbol{c}}$ for [HBP17] are all singular, but some desingularization might be possible.)

Remark 9.1.11. The "two-dimensional Kloosterman method" in [HBP17] is based on "positivity" (cf. the works [Hoo86b, Hoo97] avoiding [DFI93, HB96]), and hence is not "fully satisfactory" (in that it does not readily generalize to arbitrary $\boldsymbol{F}$'s). Finding a suitable "fully satisfactory" generalization, even in the rational function field case (where the one-dimensional Kloosterman method is "trivial to set up"), seems to be an interesting but challenging problem; see [Vis19] for some fascinating progress (with fairly good cancellation—but one might optimistically hope for even more).

Remark 9.1.12. At least for complete intersections, it might be interesting to try "lifting" the Kloosterman method to the setting of universal torsors, but this could be difficult (and maybe only directly relevant to W's of small dimension, e.g. projective surfaces).

9.2 Enlarging or deforming the delta method

For a nice historical introduction to [DFI93, HB96]'s delta method, we refer the reader to [IK04, Chapter 20]. In the rest of §9.2, we discuss some possibilities—mostly far from being fully understood so far—for modifying or perturbing the delta method.

9.2.1 Deforming or enlarging the search space

Let (F, w) be a smooth pair for some $\mathbb{P}^{m-1}_{\mathbb{Q}}$-smooth homogeneous $F \in \mathbb{Z}[x_1, \dots, x_m]$ of degree ≥ 1. Then we have

$$(1 + O_A(Y^{-A})) \cdot \frac{N_{F,w}(X)}{X^{m-\deg F}} = X^{-m+\deg F} Y^{-2} \sum_{n \geq 1} \sum_{\boldsymbol{c} \in \mathbb{Z}^m} n^{-m} S_{\boldsymbol{c}}(n) I_{\boldsymbol{c},X,Y}(n)$$

for all $(X, Y) \in \mathbb{R}^2_{>0}$, where for all $(\boldsymbol{c}, n) \in \mathbb{R}^m \times \mathbb{R}_{>0}$, we let

$$I_{\boldsymbol{c},X,Y}(n) := \int_{\boldsymbol{x} \in \mathbb{R}^m} d\boldsymbol{x}\, w(\boldsymbol{x}/X) h(n/Y, F(\boldsymbol{x})/Y^2) e(-\boldsymbol{c} \cdot \boldsymbol{x}/n) \quad \text{(following [HB96]).}$$

For simplicity, we have restricted to *homogeneously* expanding weights $\boldsymbol{x} \mapsto w(\boldsymbol{x}/X)$, but it could be useful to work more generally. Note the following:

(1) given F, w, X, the weighted count $N_{F,w}(X)$ is independent of Y; and

(2) to explore the solutions to $F = 0$, we have *both* parameters w, X at our disposal.

Changing the region

At a basic level, we have already seen the possible benefits of (2) in Chapter 2 (which is based on [Dia19]). But in Chapters 6–8, we fixed $Y := X^{3/2}$, even as we let w vary; perhaps one could get sharper results by letting Y vary with w.

Changing the modulus cutoff parameter

Typically one chooses $Y \asymp_{F,w} X^{(\deg F)/2}$; what if one perturbs Y by $X^{\pm\delta}$, or $(\log X)^{\pm\delta}$, or…? [Mun15, MV19] have made a first step in this direction. These works are based on *modifying* $I_{\boldsymbol{c},X,Y}(n)$.

But we do not seem to have a deep understanding of what choices should work. Conceptually, there are at least two natural problems one could look at. (For concreteness, we fix $(\deg F, m) = (3, 6)$, but the discussion below certainly generalizes easily.)

(1) Given $X' \ll X$, count points $\boldsymbol{x} \ll X$ with $F(\boldsymbol{x}) \ll (X')^3$ such that $F(\boldsymbol{x}) = 0$. (The basic question here is whether the condition $F(\boldsymbol{x}) \ll (X')^3$ can be used to give better integral estimates, or not. A tantalizing open question is to recover the Hua bound $N_{x_1^3+\cdots+x_6^3}(X) \ll_\epsilon X^{7/2+\epsilon}$ unconditionally using the delta method.)

(2) Given a separable form $F = F_1(\boldsymbol{x}_1) - F_2(\boldsymbol{x}_2)$, and $X_1, X_2 \geq X$, count points with $\|\boldsymbol{x}_1\| \asymp X_1$ and $\|\boldsymbol{x}_2\| \asymp X_2$ with $F_1(\boldsymbol{x}_1) \ll X^3$ and $F_2(\boldsymbol{x}_2) \ll X^3$, such that $F_1(\boldsymbol{x}_1) = F_2(\boldsymbol{x}_2)$. (For variance computations when $F_1 = F_2$, we need to let both X_1, X_2 grow the same amount; for the Hasse principle for F, it suffices to simply let one of X_1, X_2 grow and keep the other fixed, e.g. $X_2 = X$. It could also be worth splitting F in different ways, e.g. not just $6 = 3+3$ or $6 = 3+1+1+1$ but also in between, like $6 = 4+1+1$.)

Question 9.2.1. In both cases, there is freedom in how we choose the modulus cutoff Y in the delta method; what are the best choices?

Remark 9.2.2. Even after one has chosen Y, it is unclear what the "best form" of the delta method is. For (1), after choosing $Y := (X')^{(\deg F)/2}$ for a certain $X' \ll X$, [MV19]—inspired by [Mun15]—take the Fourier transform of $U(X^3\xi/Y^2)h(r, X^3\xi/Y^2)$ (and then plug in $\xi = F(\tilde{\boldsymbol{x}})$) to get something closely resembling a "heuristic circle method with denominators $n \ll Y$". Can we do better by explicitly "remembering" the condition $F(\tilde{\boldsymbol{x}}) \ll Y^2/X^3$ (given that the Fourier transform "forgets" it)?

And how does the technique for (1) generalize to (2)? Should one apply the two-variable Fourier transform to $U(X^3\xi_1/Y^2)U(X^3\xi_2/Y^2)h(r, X^3[\xi_1+\xi_2]/Y^2)$ (and then plug in $\xi_i = F_i(\tilde{\boldsymbol{x}}_i)$), moving away from the classical "heuristic" circle method (but still likely having decay in $u_1, u_2, u_1 - u_2$ "localizing" towards the classical $u_1 = u_2 = u$), or should one directly reduce to the "heuristic" (smoothed) circle method and deal with the weight $\prod U(X^3F_i(\tilde{\boldsymbol{x}}_i)/Y^2)$ later in the integral estimates?

9.2.2 Smoothing or averaging

Given an expression from the delta method, one could try to include it into some kind of weighted average—whether it be over archimedean or adelic parameters, or over number fields or varieties, or perhaps (maybe by positivity) even over expressions without immediate Diophantine interpretation. Ideally, one would like to preserve as much information as possible about the original varieties of interest.

Remark 9.2.3. One might hope to somehow "deform and smooth" (in the Fourier-analytic spirit of e.g. [Qu07, Lemma 3.2] or [Sel89, p. 169 (p. 10 in Paper 12), use of Parseval]) to reduce the RMT-based input in Theorem 8.4.3 (or at least in the more *qualitative* parts thereof) to statements about *low-lying* zeros—though I currently do not see how to do so.

Note that on the "right-hand side" of the delta method, the (X, Y, w)-dependence essentially lies in the factors $(X^{\deg F}/Y^2) \cdot X^{-m} I_{\boldsymbol{c},X,Y}(n)$.

Question 9.2.4. Are there interesting examples where one can average over some or all of X, Y, w in the delta method to simplify *qualitative* Diophantine analysis?

As an oversimplified toy example, consider the following X-*averaging* of a Mellin transform:

$$\begin{aligned} \int_X d^\times X\, g(X) \int_n d^\times n\, f(n/X) n^s &= \int_{n,v} d^\times n\, d^\times v\, g(n/v) f(v) n^s \\ &= \int_v d^\times v\, f(v) \int_n d^\times n\, g(n/v) n^s, \end{aligned}$$

where $v := n/X$. Even if f is mysterious, the X-averaging moves the n-dependence away from f, leaving the Mellin decay in $|s|$ to be dictated by the smoothness of g.

Remark 9.2.5. In fact, such "convolution creation" has appeared before in work on delta-like methods.

In order to prove $N_{F,w}(X) \ll_\epsilon X^{3+\epsilon}$ under Hypothesis HW, [Hoo97, p. 180, (17)] uses dyadic averaging in X to avoid the interesting but harder derivative estimates of [HB98] (who does not smooth at all). The point is that without smoothing, the *trivial* n-derivative bounds on $I_{\boldsymbol{c}}(n)$ would have a fatal loss as $n \to 0$ (i.e. at the end of the argument, the final n-exponent would become negative, whereas it was zero before differentiating).[1] However, with averaging, the derivative on n can be moved to a harmless "smoothing" factor,[2] without Heath-Brown's *improved* n-derivative bounds.

Similarly, one could likely simplify parts of [Wan21c] (especially in archimedean aspects, at least in the m-even Fermat case) by dyadic (or slightly larger) smoothing.

Remark 9.2.6. We call the toy example above "oversimplified" because while the factor $f(n/X)$ only depends on n/X, the actual factor $(X^{\deg F}/Y^2) \cdot X^{-m} I_{\boldsymbol{c},X,Y}(n)$ of interest depends on $n/Y, X\boldsymbol{c}/n, X^{\deg F}/Y^2$ (and F, w)—though at least if $\deg F = 2$ and we restrict to $Y = X$, then $(X^{\deg F}/Y^2) \cdot X^{-m} I_{\boldsymbol{c},X,Y}(n)$ depends only on $n/X, \boldsymbol{c}$ (and F, w).

[1] This is why [Hoo86b] needed to separately consider small n, i.e. "junior arcs" [Hoo86b, p. 81, §9], getting a final exponent of $60/19 + \epsilon = 3.15\ldots + \epsilon$ instead of $3 + \epsilon$. See [Hoo97, p. 176] for Hooley's commentary.

[2] Strictly speaking, Hooley does not smooth, but rather uses a positivity-based design, together with the "convolution-like sub-structure" created by averaging, to "eliminate" the n-dependence.

Remark 9.2.7. The limitations of averaging—and the extent of resulting internal cancellation (if any)—are unclear, due to the (currently) mysterious nature of the multi-parameter families $I_{\boldsymbol{c},X,Y}(n)$ and $I^{\vee}_{\boldsymbol{c},X,Y}(s)$ (with F, w dependence suppressed). Experimentation, whether theoretical or computational, could be enlightening.

9.3 Problems inspired by the Ratios Conjectures

9.3.1 What is the optimal power saving?

We have not attempted to optimize the power saving $X^{\Omega(1)}$ in Theorem 8.4.1(b), since there could be a more efficient approach waiting to be understood. However, a sufficiently large power saving might bring the Hasse principle for $m = 5$ within reach (perhaps under GLH and [CFK+05]'s standard moment conjectures, applied to certain Artin L-functions).

From the opposite angle, it would also be very interesting to determine any hard limits on what power savings can be expected, e.g. coming from Brauer–Manin obstructions for $m = 4$ (but there the $F^{\vee}(\boldsymbol{c}) = 0$ contribution could at least sometimes muddy the analysis).

Remark 9.3.1. When F is non-diagonal, $m \le 5$, *or* $\deg F \ge 4$, even the $F^{\vee}(\boldsymbol{c}) = 0$ contribution itself deserves to be better understood (and it may be the most tractable starting point); see the discussion in Remark 6.4.3, Example 6.5.3, [Wan21d, Remarks 1.19 and 5.1], and Question 6.5.5. And when $m = 4$, even the $\boldsymbol{c} = \boldsymbol{0}$ contribution deserves to be better understood; see e.g. Remark 3.4.2.

Question 9.3.2 (Cf. Remark 6.1.5)**.** When the Manin–Peyre prediction for $m = 4$ *differs* from the "$\mathbb{Q}$-lines plus major arcs" heuristic of [Bro09, §8.3.3]—as is the case when $F = x_1^3 + \cdots + x_4^3$ [Bro09, p. 149], for instance—where does the *difference* lie, precisely, in the delta method?

Besides $V(x_1^3 + \cdots + x_4^3)$, another example for Question 9.3.2 is $V(5x_1^3 + 12x_2^3 + 9x_3^3 + 10x_4^3)$ (due to [CG66]), which in fact has a *Brauer–Manin obstruction to rational points* over $\mathbb{Q}$. The latter example is arguably cleaner, in that it has $V(F^{\vee})(\mathbb{Q}) = \emptyset$.

Proof that $V(F^{\vee})(\mathbb{Q}) = \emptyset$. Let $m := 4$, let $\boldsymbol{F} := (5, 12, 9, 10)$, let $F := \sum_{i\in[m]} F_i x_i^3$, and let $V := V(F)/\mathbb{Q}$. Fix $\boldsymbol{c} \in \mathbb{Q}^m \setminus \{\boldsymbol{0}\}$ with $F^{\vee}(\boldsymbol{c}) = 0$, i.e. $F^{\vee}(\boldsymbol{c}) = 0$. Then by the classical "factorization" of $F^{\vee}(\boldsymbol{c})$ (see (6.1) in §6.3.2), we have $\sum_{i\in[m]} \pm(c_i^3/F_i)^{1/2} = 0$ for some choice of signs. But $\boldsymbol{c} \ne \boldsymbol{0}$, so in particular, there must exist a nonempty set $\mathcal{I} \subseteq [m]$, and rationals $d, x_i \in \mathbb{Q}^{\times}$ for $i \in \mathcal{I}$, such that (i) $c_i^3/F_i = F_i^2 x_i^6 d^3 \in d \cdot (\mathbb{Q}^{\times})^2$ for all $i \in \mathcal{I}$, and (ii) $\sum_{i\in\mathcal{I}} F_i x_i^3 = 0$. But $V(\mathbb{Q}) = \emptyset$, and $(x_i)_{i\in\mathcal{I}} \in \mathbb{Q}^{\mathcal{I}} \setminus \{\boldsymbol{0}\}$, so (ii) is impossible. Thus $\boldsymbol{c}$ cannot exist. □

In view of Remark 6.1.5, such an example naturally raises the following question:

Question 9.3.3. Given a smooth cubic surface $V = V(F)/\mathbb{Q}$ with $V(\mathbf{A}_{\mathbb{Q}}) \ne \emptyset$ and $V(\mathbf{A}_{\mathbb{Q}})^{\mathrm{Br}} = \emptyset$ (in the notation of [Jah14, Definition IV.1.1 and Notation IV.2.5]), is it necessarily true that $V(F^{\vee})(\mathbb{Q}) = \emptyset$, i.e. $V^{\vee}(\mathbb{Q}) = \emptyset$ (in the notation of §6.2)?

9.3.2 Hooley's critical ternary triumvirate

Note that

$$1 = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = \frac{1}{2} + \frac{1}{4} + \frac{1}{4} = \frac{1}{2} + \frac{1}{3} + \frac{1}{6}.$$

Let $R_4(a), R_6(a)$ denote $\#\{y, \boldsymbol{x} \ge 0 : y^2 + x_1^4 + x_2^4 = a\}$ and $\#\{x, y, z \ge 0 : x^2 + y^3 + z^6 = a\}$, respectively. [Hoo86a] made *precise asymptotic ℓ^2 conjectures* for R_4, R_6 alongside the analogous [Hoo86a, Conjecture 2] for r_3. Hooley's ternary problems belong to the framework of a "*cba* conjecture" (with a goal—opposite the *abc* conjecture—of *producing points*); cf. [Har17, Conjecture 1.11].

We focus on R_4 (for technical reasons to arise later). Just like for r_3, the first moment of R_4 grows linearly: $\sum_{a \le B} R_4(a) \propto (B^{1/2})(B^{1/4})^2(1 + o(1)) = B(1 + o(1))$ as $B \to \infty$. But now, ℓ^2 is in some sense easier than for r_3: for R_4, the divisor bound easily implies the following result.

Proposition 9.3.4. Unconditionally, $\sum_{a \le B} R_4(a)^2 \ll_\epsilon B^{1+\epsilon}$.

Remark 9.3.5. [Hoo79] proved $\sum_{a \le B} R_4(a)^2 \ll_\epsilon B(\log B)^{4/\pi - 1 + \epsilon}$ by introducing and analyzing "Δ-functions" (i.e. certain divisor sums). [Rob11] improved [Hoo79]'s bound to (something a bit stronger than) $O_\epsilon(B(\log B)^\epsilon)$; but $O(B)$ seems to remain open.

Question 9.3.6. Can the bound $O_\epsilon(B^{1+\epsilon})$ be improved to $O(B)$ for R_4 (or for R_6)?

Observation 9.3.7. If $\#\{(\boldsymbol{x}, \boldsymbol{y}) \in [0, X]^4 \times [X^2/2, X^2]^2 : y_1^2 + x_1^4 + x_2^4 = y_2^2 + x_3^4 + x_4^4\} \ll X^4$, then $\{a \in \mathbb{Z}_{\ge 0} : R_4(a) \ne 0\}$ has *positive lower density.*

Proof. Use double counting and Cauchy as for r_3, but with the large variable "y" restricted to $[X^2/2, X^2]$. □

For the question above (restricted to $\|\boldsymbol{y}\| \asymp X^2$ for simplicity), we now describe a possible conditional approach whose roots lie close in spirit to work of Hooley (e.g. [Hoo81, Hoo84]). Let $F_4(\boldsymbol{x}) = F_4(x_1, \ldots, x_4) := x_1^4 + x_2^4 - x_3^4 - x_4^4$. Let $n := y_1 + y_2$ and $N := X^{\deg F_4/2} = X^2$. Then

$$\#\{(\boldsymbol{x}, \boldsymbol{y}) \in [0, X]^4 \times [N/2, N]^2 : (y_2 - y_1)(y_2 + y_1) = F_4(\boldsymbol{x})\} \le \sum_{n \in [N, 2N]} \sum_{\boldsymbol{x} \in [0, X]^4} \mathbf{1}_{n \mid F_4(\boldsymbol{x})}.$$

By positivity, it is harmless to over-extend the right-hand side to a smooth sum

$$\mathcal{S}_4 := \sum_{n \ge 1} D(n/N) \sum_{\boldsymbol{x} \in \mathbb{Z}^4} \mathbf{1}_{n \mid F_4(\boldsymbol{x})} \cdot w(\boldsymbol{x}/X)$$

with $D \in C_c^\infty(\mathbb{R}_{>0})$ and $w \in C_c^\infty(\mathbb{R}^4)$. Given n, we then expand $\mathbf{1}_{n \mid F_4(\boldsymbol{x})}$ over "vertices" $a \bmod n$ on the circle, using additive characters $t \mapsto e_n(at)$, to get

$$\mathcal{S}_4 = \sum_{n \ge 1} D(n/N) n^{-1} \sum_{a \in \mathbb{Z}/n} \sum_{\boldsymbol{x} \in \mathbb{Z}^4} e_n(aF_4(\boldsymbol{x})) \cdot w(\boldsymbol{x}/X).$$

(One might call this the "polygon(s) method" for divisor-type problems; it is very close in spirit to the circle method.) Now let $S_{4,\boldsymbol{c}}(n) := \sum_{a\in\mathbb{Z}/n}\sum_{\boldsymbol{z}\in(\mathbb{Z}/n)^4} e_n(aF_4(\boldsymbol{z}) + \boldsymbol{c}\cdot\boldsymbol{z})$. Then by Poisson summation, one can show that

$$\mathcal{S}_4 = \sum_{n\geq 1} D(n/N)n^{-1}\sum_{\boldsymbol{c}\in\mathbb{Z}^4} S_{4,\boldsymbol{c}}(n)\cdot(X/n)^4\hat{w}(X\boldsymbol{c}/n).$$

Conjecture 9.3.8. *Automorphy and GRH for smooth hyperplane sections of the form* $V_{\mathbb{P}^3}(F_4, \boldsymbol{c}\cdot\boldsymbol{x})_{\mathbb{Q}}$ *(with* $\boldsymbol{c}\in\mathbb{Z}^4\setminus\{\boldsymbol{0}\}$*) should recover a "*$B^{1+\epsilon}$*-like" bound,* $\mathcal{S}_4 \ll_\epsilon X^{4+\epsilon}$*.*

Remark 9.3.9. These hyperplane sections are essentially *smooth plane quartic curves*. In general, a smooth projective curve over $\mathbb{Q}$ has the same Hasse–Weil L-function as its Jacobian (an abelian variety of dimension equal to the genus of the curve). So the hypotheses above concern the L-functions of certain abelian varieties of dimension 3.

Assuming the question of proving $\mathcal{S}_4 \ll X^4$ is indeed open—and not susceptible to hypersurface slicing methods—the above approach is interesting in that one might be able to *remove the* ϵ conditionally (in the fashion of Chapter 8).

9.3.3 On quartics

A "standard" application of the delta method, without cancellation over $\boldsymbol{c}$, is (at least morally) limited in scope to quadratic and cubic problems—at least as far as hypersurfaces are concerned. It is amusing, but probably incorrect, to compare this situation to the following classical fact: every Diophantine equation over $\mathbb{Z}$ (or $\mathbb{Q}$) is equivalent to a system of quadratic equations, and thus to a single quartic equation; an analogous statement holds for homogeneous Diophantine equations over $\mathbb{Z}$ (or $\mathbb{Q}$), by [Mum70, Theorem 1] (a result based on Veronese embeddings).

In "natural" quartic and higher-degree equations, there should be more structure that remains to be uncovered. For instance, [MV19] uses "averaging over a mod q" to get the best known results on general quartic projective hypersurfaces, although it uses a "non-standard modulus cutoff" in the delta method. Furthermore, Question 9.3.6 presents a "borderline quartic" example where the "standard delta method" should provide a reasonable conditional avenue forward. There is also some hope of using RMT-type predictions to analyze quartics beyond [MV19], but any attempt at a fully rigorous conditional analysis might face serious difficulties at the moment.

At the same time, it is very natural to wonder if there might be a better (more efficient) way to "complete exponential sums" for quartics like F_4. So it cannot hurt to list a few half-baked thoughts, in the hope of inspiring further discussion or creativity:

(1) The tantalizing Question 9.3.6 may provide a good "critical" testing ground for old and new quartic techniques alike.

(2) At least for smooth projective cubic hypersurfaces, one expects all "special subvarieties" to be *linear*. But at least *some* smooth projective quartic hypersurfaces contain special *quadrics*: see e.g. Example 6.5.4.

The contrast between cubics and quartics seems vaguely parallel to the difference between Szemerédi's theorem for 3-term and 4-term arithmetic progressions—the former being of a "linear" nature, and the latter of a "quadratic" nature. (But maybe "quadratic" means something else here, in that the set of 4-term arithmetic progressions is cut out by a system of 2 linear equations.)

(3) For homogeneous quartics, maybe L-functions associated to more complicated slices (e.g. degree 2 hypersurfaces) should be relevant, not just (linear) hyperplane slices as in the standard Kloosterman method for homogeneous equations?

(4) The standard Kloosterman method is based on *interpolation*—i.e. *completing* incomplete sums—via Nyquist–Shannon sampling. (Given a weight $w \in C_c^\infty(\mathbb{R})$, a modulus $n \gg X$, and a residue $z \bmod n$, one replaces $\sum_{x \equiv z \bmod n} w(x/X)$ with a "dual" sum of length $\lessapprox n/X$—a certain linear combination of *linear* phases $e_n(c \cdot z)$.)

Is there a useful "nonlinear" version of Nyquist–Shannon sampling?

(5) One might try to replace "exact" sums with "approximate" sums. For example, in the context of $\mathcal{S}_4$ (as defined after Question 9.3.6), one could write

$$\sum_{x^2 \equiv z^2 \bmod n} w(x/X) = \sum_{y \equiv z^2 \bmod n} \mathbf{1}_{y=\square} \cdot w(y^{1/2}/X) \quad \text{for any given } z \bmod n,$$

and then try to detect $\mathbf{1}_{y=\square}$ "statistically" (in the spirit of the "square sieve").

(6) Perhaps one could seek inspiration from quadratic Fourier analysis, or from formulas in sphere packing that contain $f(\sqrt{n})$'s, or from [Kum18, pp. 25–27, §2.2 Other versions of the δ-method]'s discussion of Jutila's and Munshi's ideas, or from. . . .

Appendix A

Modularity questions

In this appendix, we fix a $\mathbb{P}^{m-1}_{\mathbb{Q}}$-smooth cubic form $F/\mathbb{Z}$ in $m \in \{4, 6\}$ variables, and consider the Hasse–Weil L-functions $L(s, V_{\boldsymbol{c}})$ appearing in Example 3.2.10. Fix $\boldsymbol{c}$. (Here we restrict attention to $\boldsymbol{c} \in \mathbb{Z}^m$ with $V_{\boldsymbol{c}}$ smooth of dimension m_*.)

A.1 A general discussion

Let us first recall some useful definitions and facts. Let ℓ be a prime, and K a field of characteristic $\neq \ell$. Let X be a smooth, projective, geometrically integral variety over K. Let $P_{X/K}$ denote the *Picard variety* of X (following [Poo17, §5.7.3], say), and $A_{X/K}$ the *Albanese variety* of X (following [Poo17, Example 5.12.11], say). It is known that $P_{X/K}, A_{X/K}$ are *dual* abelian varieties over K [Poo17, Theorem 5.7.20].

Proposition A.1.1 (Standard)**.** *In this setting, we have a canonical isomorphism $H^1(X_{\overline{K}}, \mathbb{Q}_\ell) \cong V_\ell(P_{X/K})(-1) \cong V_\ell(A_{X/K})^\vee$ of ℓ-adic G_K-representations.*

Proof. To get the statement, one can combine, for instance, [Pet11, Remark] and [Bel13, par. 1]. (Alternatively, see [GL02, Proposition 9.6], which however assumes "for simplicity" that X has a K-rational point.) The standard proof follows [Tom13, Piotr Achinger's comment and Adel Betina's answer]. By ℓ-adic Kummer theory (cf. [Poo17, §7.6.3]) and [Poo17, Proposition 6.6.1], one obtains isomorphisms $H^1(X_{\overline{K}}, (\mathbb{Z}/\ell^n)(1)) \cong \operatorname{Pic}(X_{\overline{K}})[\ell^n]$ for $n \geq 1$, compatible with multiplication by ℓ (as n varies). The result follows. (See [Poo17, §§7.5.3–7.5.4] for some useful background on ℓ-adic cohomology, Tate modules, and Tate twists.) $\square$

We now return to the original setting of this appendix. If $m = 4$, then $V_{\boldsymbol{c}}$ is a genus one curve, so $L(s, V_{\boldsymbol{c}})$ is known to be the L-function of a weight 2 modular cusp form (since $L(s, V_{\boldsymbol{c}}) = L(s, J(V_{\boldsymbol{c}}))$ by Proposition A.1.1 and [Poo17, Example 5.12.12], and elliptic curves over $\mathbb{Q}$ are known to be modular).

Now suppose $m = 6$. Let $A_{\boldsymbol{c}}$ denote the Albanese variety of $F(V_{\boldsymbol{c}})/\mathbb{Q}$, the Fano surface of lines on $V_{\boldsymbol{c}}$. If we consider the ℓ-adic $G_{\mathbb{Q}}$-representations $M_{\boldsymbol{c}} := H^3(V_{\boldsymbol{c}} \times \overline{\mathbb{Q}}, \mathbb{Q}_\ell)$ and $V_\ell(A_{\boldsymbol{c}})^\vee = H^1(A_{\boldsymbol{c}} \times \overline{\mathbb{Q}}, \mathbb{Q}_\ell)$ defining $L(s, V_{\boldsymbol{c}}), L(s, A_{\boldsymbol{c}})$, respectively, then $M_{\boldsymbol{c}}(1) \cong H^1(F(V_{\boldsymbol{c}}) \times \overline{\mathbb{Q}}, \mathbb{Q}_\ell) \cong V_\ell(A_{\boldsymbol{c}})^\vee$, by [Rei72, Appendix 4.3, Corollaries 4.29] (or [CP15, paragraph containing diagram (5)]; cf. [DLR17, first paragraph of the proof of Theorem 4.1]) and Proposition A.1.1. In particular, $L(s, V_{\boldsymbol{c}}) = L(s, A_{\boldsymbol{c}})$.

Here $\dim A_{\boldsymbol{c}} = 5$. But at least in general, abelian varieties $A/\mathbb{Q}$ of dimension 5 are not yet known to be automorphic. (In this connection, the current state of the art is [BCGP21], which "potentially addresses" abelian *surfaces* over certain fields like $\mathbb{Q}$.)

However, one might ask if (at least) for certain F's, like $x_1^3 + \cdots + x_6^3$, the $A_{\boldsymbol{c}}$'s might have some special structure that could "reduce" the complexity of $L(s, A_{\boldsymbol{c}})$. In this vein, recall that the Hasse–Weil L-function of a diagonal projective hypersurface over $\mathbb{Q}$ of degree $d \geq 2$ always factors into L-functions of degree $\phi(d)$ associated to certain Hecke characters on $\mathbb{Q}(\zeta_d)$; at least up to bad factors, this observation is due to Weil [Wei52].

Now let $F := x_1^3 + \cdots + x_6^3$. Does some loose analog of Weil's observation hold for the "codimension-2 diagonal systems" $V_{\boldsymbol{c}}$? In the most direct sense, the answer to this basic question is probably negative in general (see Proposition A.2.1 below). So even when $F = x_1^3 + \cdots + x_6^3$, the conjectured automorphy of $V_{\boldsymbol{c}}$ (or $A_{\boldsymbol{c}}$) does not seem to easily follow from existing progress on the Langlands program; one may well need interesting new ideas in order to probe the family $\boldsymbol{c} \mapsto L(s, V_{\boldsymbol{c}})$ from an automorphic point of view.

A.2 Notes on an explicit diagonal system

Armed with the theory of abelian varieties,[1] we can prove the following statement:

Proposition A.2.1. *Let* $F := x_1^3 + \cdots + x_6^3$ *and* $\boldsymbol{c} := (1, 2, 3, 4, 5, 6)$. *Then for each prime* ℓ *and number field* $K/\mathbb{Q}$, *the* ℓ*-adic* $G_{\mathbb{Q}}$*-representations* $M_{\boldsymbol{c}}, V_\ell(A_{\boldsymbol{c}})$ *are irreducible as* ℓ*-adic* G_K*-representations.*

Remark A.2.2. Proposition A.2.1 does not rule out the possibility that $L(s, V_{\boldsymbol{c}})$ could be "deconstructed" in a subtler way, e.g. as the 9th symmetric power L-function of a GL_2 representation.

Remark A.2.3. It may or may not be natural to contrast Proposition A.2.1 with [Pat97, Theorem 3.1], a result showing that the "one-dimensional constituents of $S_{\boldsymbol{c}}(q)$" (before taking mixed 6th moments over $a \bmod q$) are "nice" over $\mathbb{Q}(\zeta_3)$ (and in fact, amenable to statistical analysis via cubic metaplectic forms [Lou14]).

The conclusion of Proposition A.2.1 is equivalent to the statement that $\operatorname{End}(A_{\boldsymbol{c}} \times \overline{\mathbb{Q}}) = \mathbb{Z}$. Thus Proposition A.2.1 naturally suggests the following conjecture:

Conjecture A.2.4. *For* almost all $\boldsymbol{c} \in \mathbb{Z}^m$ *with* $F^\vee(\boldsymbol{c}) \neq 0$, *we have* $\operatorname{End}(A_{\boldsymbol{c}} \times \overline{\mathbb{Q}}) = \mathbb{Z}$.

Remark A.2.5. By [Mas96], the conjecture above should be equivalent to a question about the generic fiber of the family $A_{\boldsymbol{c}}$. Furthermore, a "typical" special fiber should control the possible behavior at the generic fiber. But Proposition A.2.1 does not seem to imply the conjecture "for free" (without further calculation).

Proof of Proposition A.2.1. For convenience, let $A := A_{\boldsymbol{c}}$. Now fix ℓ. We know that $M_{\boldsymbol{c}}(1) \cong V_\ell(A)^\vee$ as ℓ-adic $G_{\mathbb{Q}}$-representations. So by Faltings' isogeny and semi-simplicity theorems (and the fact that endomorphism rings of abelian varieties are torsion-free as $\mathbb{Z}$-modules),

[1] Thanks to Will Sawin for initial help in this direction.

our desired irreducibility statement is equivalent to the statement that $\operatorname{End}(A_K) = \mathbb{Z}$ holds for all $K/\mathbb{Q}$.

We now prove the latter statement—which is independent of ℓ. To begin, we re-define $\ell := 2$ for convenience. By Appendix A.3.1, $V_{\boldsymbol{c}}$ has good reduction at all primes $p \in [5, 100]$, so by smooth proper base change, $M_{\boldsymbol{c}}$ is unramified at all $p \in [5, 100]$. Since $M_{\boldsymbol{c}}(1) \cong V_\ell(A)^\vee$, it follows that $V_\ell(A)$ is unramified at all $p \in [5, 100]$. So by the Néron–Ogg–Shafarevich criterion, A has good reduction at all $p \in [5, 100]$.

For an abelian variety B over a field, let $\operatorname{End}^0(B) := \operatorname{End}(B) \otimes_{\mathbb{Z}} \mathbb{Q}$; then $\operatorname{End}(B)$ is always a subring of $\operatorname{End}^0(B)$, finitely generated as a $\mathbb{Z}$-module. In particular, to achieve our original goal, it suffices to prove that $\operatorname{End}^0(A_K) = \mathbb{Q}$ holds for all $K/\mathbb{Q}$.

Now fix $K/\mathbb{Q}$, and fix a Néron model $\mathcal{A}_K/\mathcal{O}_K$ of A_K. Then for each prime $\wp$ of $\mathcal{O}_K$, let $k_\wp := \mathcal{O}_K/\wp$ and $f_\wp := [k_\wp : \mathbb{F}_p]$. Finally, for each A_K-*good* prime $\wp$, let $A_\wp := \mathcal{A}_K \times k_\wp$ denote the reduction of A_K modulo $\wp$; let $\pi_\wp \in \operatorname{End}(A_\wp)$ denote the "geometric Frobenius" *morphism* $A_\wp \to A_\wp$ over $k_\wp$ (as in [Sta22, Tag 03SQ]); and *if* $\wp \nmid \ell$, let $P_\wp \in 1 + t\mathbb{Z}[t]$ denote the "*reverse* characteristic polynomial" of $\pi_\wp$. Then given $\wp \mid p \in [5, 100]$, we have $P_\wp(t) = \prod(1 - \beta_i^{f_\wp} t)$, where $P_{p\mathbb{Z}}(t) = \prod(1 - \beta_i t)$, where for $p \in \{5, 7, 11, 13\}$ we know by Appendix A.3.1 that

(1) $P_{5\mathbb{Z}}(t) = (5t^2 + 1) \cdot (625t^8 + 100t^6 + 4t^2 + 1)$ at $p = 5$,

(2) $P_{7\mathbb{Z}}(t) = (7t^2 + t + 1) \cdot (7t^2 + 4t + 1)^4$ at $p = 7$,

(3) $P_{11\mathbb{Z}}(t) = 161051t^{10} + 14641t^9 + 3993t^8 + 2420t^7 + 539t^6 + 14t^5 + 49t^4 + 20t^3 + 3t^2 + t + 1$ at $p = 11$, and

(4) $P_{13\mathbb{Z}}(t) = (13t^2 + 1) \cdot (13t^2 + 7t + 1) \cdot (13t^2 + 4t + 1)^3$ at $p = 13$.

In order to put the data above to use, note that for each A_K-good prime $\wp$, the theory of Néron models furnishes a canonical *reduction map* $\operatorname{End}(A_K) \to \operatorname{End}(A_\wp)$, which is known to be *injective* (see e.g. [Hui16]). On the other hand, for each A_K-good $\wp$, the results [Tat66, Theorems 1(a) and 2(a)] (combined with the fact, noted in [Tat66, p. 141, par. 2], that over $k_\wp$, "isogeny-class factorizations" correspond to "simple E-algebra factorizations") imply that

$$\operatorname{End}^0(A_\wp) \cong E_{\wp,1} \times \cdots,$$

where $E_{\wp,1}, \ldots$ are certain *central simple algebras* (CSA's) over the *number field factors* $F_{\wp,1}, \ldots$ of the commutative subalgebra $\mathbb{Q}[\pi_\wp] \subseteq \operatorname{End}^0(A_\wp)$, and where

$$F_{\wp,j} \cong \mathbb{Q}[T]/(T^{10} P_{\wp,j}(1/T)) \cong \mathbb{Q}[t]/(P_{\wp,j}(t)) \quad \text{and} \quad [E_{\wp,j} : F_{\wp,j}] = m_{\wp,j}^2$$

correspond to an arbitrary factorization $P_\wp(t) = P_{\wp,1}(t)^{m_{\wp,1}} \cdots$ of $P_\wp(t)$ into irreducibles over $\mathbb{Q}$ of the form $1 + t\mathbb{Z}[t]$.

Finally, fix $\wp \mid 11$ and $\wp' \mid 5$. Then by Lemma A.2.7(1), $P_\wp$ is irreducible over $\mathbb{Q}$. So $\operatorname{End}^0(A_\wp) = \mathbb{Q}[\pi_\wp] \cong F_{\wp,1} \cong L_{11}$. The algebra $\operatorname{End}^0(A_K)$ is thus isomorphic to a number field $F_0 \subseteq L_{11}$. Now on the one hand, $F_0 \subseteq L_{11}$ automatically implies $[F_0 : \mathbb{Q}] \in \{1, 5, 10\}$, by Lemma A.2.7(1). On the other hand, by Lemma A.2.7(2)–(3) (applying (2) if $2 \nmid f_{\wp'}$, and (3) if $2 \mid f_{\wp'}$), we must have $\operatorname{End}^0(A_{\wp'}) \cong E_{\wp',1} \times E_{\wp',2}$, where each $E_{\wp',j}$ is a CSA of dimension

$m^2_{\wp',j} \in \{1,4\}$ over a number field $F_{\wp',j}$ of degree $[F_{\wp',j} : \mathbb{Q}] \in \{1,2,4,8\}$. Now fix $j \in [2]$, and let F'_j denote the image of the composition $F_0 \cong \operatorname{End}^0(A_K) \to \operatorname{End}^0(A_{\wp'}) \twoheadrightarrow E_{\wp',j}$. Let $F_{\wp',j}\langle F'_j\rangle$ denote the $F_{\wp',j}$-subalgebra of $E_{\wp',j}$ generated by F'_j. Then certainly $[F_{\wp',j}\langle F'_j\rangle : F_{\wp',j}] \le [E_{\wp',j} : F_{\wp',j}] = m^2_{\wp',j} \le 4$. But $F_{\wp',j}\langle F'_j\rangle$ is an F'_j-module, and hence a *free* F'_j-module (since F'_j is a field). (In fact, $F_{\wp',j}\langle F'_j\rangle$ is a *commutative* F'_j-algebra, since $F_{\wp',j}$ is *central* in $E_{\wp',j}$.) Therefore, $[F'_j : \mathbb{Q}] \mid [F_{\wp',j}\langle F'_j\rangle : \mathbb{Q}] \perp 5$. But the surjection $F_0 \twoheadrightarrow F'_j$ is *nonzero*, because $E_{\wp',j} \neq 0$ (and all of our ring maps, including $F_0 \to E_{\wp',j}$, are *unital*). Also, F_0 is a *field*. Therefore, $F'_j \cong F_0$. It follows that $[F'_j : \mathbb{Q}] = [F_0 : \mathbb{Q}]$, whence $[F_0 : \mathbb{Q}] \in \{1,5,10\} \cap \mathbb{Z}_5^\times = \{1\}$. So $[\operatorname{End}^0(A_K) : \mathbb{Q}] = [F_0 : \mathbb{Q}] = 1$, as desired. $\square$

Remark A.2.6. One could slightly shorten the proof, by replacing K with a suitable quadratic extension K' (if necessary) to reduce to the case where there exists $\wp \mid 5$ with $2 \mid f_\wp$.

Let $Q_5(t) := P_{5\mathbb{Z}}(t)$ and $Q_{11}(t) := P_{11\mathbb{Z}}(t)$ be the specific polynomials at $p \in \{5, 11\}$ written above. We need the following lemma:

Lemma A.2.7. *Given $p \in \{5, 11\}$ and $f \ge 1$, write $Q_p(t) = \prod(1 - \beta_i t)$, let*

$$Q_{p^f}(t) := \prod(1 - \beta_i^f t),$$

let $Q_{p^f,1}(t)^{m(p^f,1)} \cdots$ denote a factorization of $Q_{p^f}(t)$ into irreducibles over $\mathbb{Q}$ of the form $1 + t\mathbb{Z}[t]$, and let $\sqrt{Q_{p^f}} := Q_{p^f,1} \cdots$ denote a "radical" of Q_{p^f}.

(1) The polynomial $Q_{11}(t)$ is irreducible, and the field $L_{11} := \mathbb{Q}[t]/Q_{11}$ contains exactly 1 *nontrivial subfield L'_{11}. Here $[L_{11} : \mathbb{Q}] = 10$ and $[L'_{11} : \mathbb{Q}] = 5$. Furthermore, for each $f \ge 1$, the polynomial $Q_{11^f}(t)$ is irreducible, and $\mathbb{Q}[t]/Q_{11^f} \cong L_{11}$.*

(2) We can write $Q_5 = Q_{5,1}Q_{5,2}$ with $\deg Q_{5,1} = 2$ and $\deg Q_{5,2} = 8$. Under these conventions, the fields $F_{5,j} := \mathbb{Q}[t]/Q_{5,j}$ have degree $\deg Q_{5,j}$. Furthermore, for each odd *$f \ge 1$, we have $Q_{5^f} = \sqrt{Q_{5^f}}$ and $\mathbb{Q}[t]/Q_{5^f} \cong F_{5,1} \times F_{5,2}$.*

(3) We can write $\sqrt{Q_{5^2}} = Q_{5^2,1}Q_{5^2,2}$ with $\deg Q_{5^2,1} = 1$ and $\deg Q_{5^2,2} = 4$. Under these conventions, the fields $F_{5^2,j} := \mathbb{Q}[t]/Q_{5^2,j}$ have degree $\deg Q_{5^2,j}$. Furthermore, for each even *$f \ge 2$, we have $Q_{5^f} = (\sqrt{Q_{5^f}})^2$ and $\mathbb{Q}[t]/\sqrt{Q_{5^f}} \cong F_{5^2,1} \times F_{5^2,2}$.*

Remark A.2.8. It is important to view $L_{11}, F_{5,2}$ as "abstract" (rather than "embedded") number fields, since they are not Galois (even though $F_{5,1}, F_{5^2,1}, F_{5^2,2}$ happen to be Galois).

Before proceeding, note in general that $Q_q(t) \mid Q_{q^f}(t^f)$, so the algebra map $\mathbb{Q}[t] \to \mathbb{Q}[t]/Q_q$ given by $t \mapsto t^f$ factors through a "canonical" map

$$\phi_{q,f} \colon \mathbb{Q}[t]/Q_{q^f} \to \mathbb{Q}[t]/Q_q$$

between two 10-dimensional algebras. In general, $\phi_{q,f}$ need not be bijective (or equivalently, injective or surjective), but it is always nonzero (because it sends 1 to 1). But we can still analyze $Q_{q^f}(t)$ reasonably well, using the Galois-theoretic correspondence between irreducible polynomials over $\mathbb{Q}$ on the one hand, and $G_{\mathbb{Q}}$-orbits in $\overline{\mathbb{Q}}$ on the other.

Proof of (1). Fix $f \geq 1$. Let $\phi := \phi_{11,f}$. By Appendix A.3.1, we know that Q_{11} is irreducible, i.e. L_{11} is a field. So $G_{\mathbb{Q}}$ acts *transitively* on the roots of Q_{11} in $\overline{\mathbb{Q}}$, and hence also on the roots of Q_{11^f}. Thus $\sqrt{Q_{11^f}}$ is irreducible—i.e. $\mathbb{Q}[t]/Q_{11^f}$ is a *local* Artinian ring. Since $\phi \neq 0$, it follows that $\mathbb{Q}[t]/Q_{11^f} \cong L_{11}$ holds if and only if $Q_{11^f} = \sqrt{Q_{11^f}}$.

So to prove (1), it remains to show that $Q_{11^f}(t)$ is square-free (and thus irreducible), and compute the subfields of L_{11}. (We can do these tasks in either order.)

For the first task, write $Q_{11}(t) = \prod(1 - \beta_i t)$ with $\beta_1, \ldots, \beta_{10} \in \overline{\mathbb{Q}}$; then we need to show that $\beta_1^f, \ldots, \beta_{10}^f$ are pairwise distinct. Assume for contradiction that $\beta_i^f = \beta_j^f$ for some distinct i, j. Then β_i/β_j is a *primitive* nth root of unity for some $n \mid f$. In particular, $\zeta_n \in \mathbb{Q}(\beta_i, \beta_j)$. Since $[L_{11} : \mathbb{Q}] = 10$, it follows that $\phi(n) = [\mathbb{Q}(\zeta_n) : \mathbb{Q}] \mid 10d$ for some $d \in \{1, 2, \ldots, 9\}$. But in general, $\phi(n) \geq \sqrt{n/2}$, so $90 \geq \sqrt{n/2}$, whence $n < 2 \cdot 100^2$. By Appendix A.3.2, we conclude that $n \leq 330$. Hence by Appendix A.3.2, $[\mathbb{Q}(\beta_i^n) : \mathbb{Q}] = 10$, so the Galois conjugates $\beta_1^n, \ldots, \beta_{10}^n$ must be pairwise distinct—contradicting the fact that $(\beta_i/\beta_j)^n = 1$.

For the second task, see Appendix A.3.2. $\square$

Proof of (2). By Appendix A.3.1, we can (and must) take $Q_{5,1} := 5t^2 + 1$ and $Q_{5,2} := 625t^8 + 100t^6 + 4t^2 + 1$. (These are irreducible, i.e. $F_{5,1}$ and $F_{5,2}$ are fields.)

Now fix $f \geq 1$ with $2 \nmid f$. Let $\phi := \phi_{5,f}$. For convenience, write $Q_{5,1} = \prod_{i \leq 2}(1 - \beta_i t)$ and $Q_{5,2} = \prod_{i \geq 3}(1 - \beta_i t)$ with $\beta_1, \ldots, \beta_{10} \in \overline{\mathbb{Q}}$. Then let $Q_{5,1,f} := \prod_{i \leq 2}(1 - \beta_i^f t)$ and $Q_{5,2,f} := \prod_{i \geq 3}(1 - \beta_i^f t)$. For each j, Galois theory implies that $\sqrt{Q_{5,j,f}}$ is irreducible; cf. the proof of (1). Furthermore, $Q_{5,1,f}$ is clearly square-free, since $\beta_1^f = -\beta_2^f \neq \beta_2^f$.

In fact, $Q_{5,2,f}$ is also square-free. Assume not; then arguing as in (1), we obtain distinct $i, j \geq 3$, and an integer $n \leq 330$, such that $n \mid f$ and $(\beta_i/\beta_j)^n = 1$. But here $2 \nmid f$, so $2 \nmid n$. So by Appendix A.3.2, $[\mathbb{Q}(\beta_i^n) : \mathbb{Q}] = 8$, so the Galois conjugates $\beta_3^n, \ldots, \beta_{10}^n$ must be pairwise distinct—a contradiction.

From the above information, we conclude that $Q_{5,j,f}$ is irreducible for each j. Since $\deg Q_{5,1,f} \neq \deg Q_{5,2,f}$, it then follows that the polynomial $Q_{5^f} = Q_{5,1,f} Q_{5,2,f}$ is square-free. Consequently, by the Chinese remainder theorem modulo Q_{5^f}, Q_5, we can identify ϕ with the product of the canonical maps $\psi_j \colon \mathbb{Q}[t]/Q_{5,j,f} \to \mathbb{Q}[t]/Q_{5,j}$. But ψ_1, ψ_2 are nonzero maps between *fields*—and thus isomorphisms. So ϕ is an isomorphism itself, whence $\mathbb{Q}[t]/Q_{5^f} \cong \mathbb{Q}[t]/Q_5 \cong F_{5,1} \times F_{5,2}$, as desired. $\square$

Proof of (3). Following (2) and its proof, write $Q_{5,1} = \prod_{i \leq 2}(1 - \beta_i t)$ and $Q_{5,2} = \prod_{i \geq 3}(1 - \beta_i t)$ with $\beta_1, \ldots, \beta_{10} \in \overline{\mathbb{Q}}$. The polynomials $Q_{5,1}, Q_{5,2}$ are *even*, so we may assume that $\beta_1 + \beta_2 = \beta_3 + \beta_4 = \cdots = 0$. Then in particular, $Q_{5^2}(t) = Q_{5^2,1}(t)^2 Q_{5^2,2}(t)^2$, where $Q_{5^2,1} := \prod_{2 \mid i \leq 2}(1 - \beta_i^2 t) = Q_{5,1}(t^{1/2}) = 5t + 1$ and $Q_{5^2,2} := \prod_{2 \mid i \geq 3}(1 - \beta_i^2 t) = Q_{5,2}(t^{1/2}) = 625t^4 + 100t^3 + 4t^1 + 1$. (Note that $Q_{5^2,j}(t^2) = Q_{5,j}(t)$, and each $Q_{5,j}$ is irreducible, so each $Q_{5^2,j}$ is certainly irreducible.)

Now fix $f \geq 1$ with $2 \mid f$. Let $\phi := \phi_{5^2,f/2}$. Let $Q_{5^2,1,f/2} := \prod_{2 \mid i \leq 2}(1 - \beta_i^f t)$ and $Q_{5^2,2,f/2} := \prod_{2 \mid i \geq 3}(1 - \beta_i^f t)$, and let $Q_{5,1,f} := \prod_{i \leq 2}(1 - \beta_i^f t) = Q_{5^2,1,f/2}^2$ and $Q_{5,2,f} := \prod_{i \geq 3}(1 - \beta_i^f t) = Q_{5^2,2,f/2}^2$. For each j, Galois theory implies that $\sqrt{Q_{5^2,j,f/2}}$ is irreducible; cf. the proof of (1). Furthermore, $Q_{5^2,1,f/2}$ is clearly square-free, since $\deg Q_{5^2,1,f/2} = 1$.

In fact, $Q_{5^2,2,f/2}$ is also square-free. Assume not; then arguing as in (1), we obtain distinct *even* indices $i, j \geq 3$, and an integer $n \leq 330$, such that $n \mid f/2$ and $(\beta_i^2/\beta_j^2)^n = 1$. So by Appendix A.3.2, $[\mathbb{Q}((\beta_i^2)^n):\mathbb{Q}]=4$, so the Galois conjugates $(\beta_4^2)^n,\dots,(\beta_{10}^2)^n$ must be pairwise distinct—a contradiction.

We conclude that $Q_{5^2,1,f/2}$ and $Q_{5^2,2,f/2}$ are irreducible. Since $\deg Q_{5^2,1,f/2} \neq \deg Q_{5^2,2,f/2}$, it then follows that $\sqrt{Q_{5^f}} = Q_{5^2,1,f/2}Q_{5^2,2,f/2}$ and $Q_{5^f} = (\sqrt{Q_{5^f}})^2$. By studying the "canonical" maps $\phi'\colon \mathbb{Q}[t]/\sqrt{Q_{5^f}} \to \mathbb{Q}[t]/\sqrt{Q_{5^2}}$ and $\psi_j'\colon \mathbb{Q}[t]/Q_{5^2,j,f/2} \to \mathbb{Q}[t]/Q_{5^2,j}$, we obtain an isomorphism $\mathbb{Q}[t]/\sqrt{Q_{5^f}} \cong \mathbb{Q}[t]/\sqrt{Q_{5^2}} \cong F_{5^2,1}\times F_{5^2,2}$, as desired; cf. the proof of (2). $\square$

A.3 Supporting code

A.3.1 Local zeta polynomials

Running (in SageMath)

```
def Delta(c_1,c_2,c_3,c_4,c_5,c_6):
    return Integer(expand(
        3*prod(sqrt(c_1)^3 + i[0]*sqrt(c_2)^3 + i[1]*sqrt(c_3)^3
        + i[2]*sqrt(c_4)^3 + i[3]*sqrt(c_5)^3 + i[4]*sqrt(c_6)^3
    for i in cartesian_product([[-1,1],[-1,1],
                                [-1,1],[-1,1],[-1,1]]))));
factor(Delta(1,2,3,4,5,6))
```

yields `3^13 * 996001 * 1898591 * 107541241 * 1722583559`. So in particular, $V_{\boldsymbol{c}}$ in Proposition A.2.1 has good reduction at all primes $p \in [5,100]$. Now in [DLR17]'s supplementary code [Laf16, `algorithms.txt`], re-define the function `ZetaG` as follows:

```
// Input: (c,q) with q an odd prime power,
//              c[6] ne 0 in Fq, and V_c/Fq smooth of dimension 3
// Output: the characteristic polynomial P_1(F(V_c))
ZetaG := function(c,q)
 Fq := FiniteField(q);
 P4<[x]> := ProjectiveSpace(Fq,4);
 X := Scheme(P4,c[6]^3*(&+[x[i]^3 : i in [1..5]])
               - (&+[c[i]*x[i] : i in [1..5]])^3);
 L := Scheme(P4,[x[2]+x[4], x[3]+x[5],
                            c[6]*x[1] - &+[c[i]*x[i] : i in [1..5]]]);
 R<t> := PolynomialRing(Rationals());
 n := [Mr(X,L,r) : r in [1..5]];
 u := -&+[n[r]/r*t^r : r in [1..5]];
 ll := Coefficients(R!&+[u^k/Factorial(k) : k in [0..5]])[1..6];
 cf := ll cat [q^i*ll[6-i] : i in [1..5]];
 return &+[cf[i]*t^(i-1) : i in [1..11]];
end function;
```

Then from the Magma session

```
> load "algorithms.txt";
```

```
> ZetaG([1,2,3,4,5,6],5);
3125*t^10 + 1125*t^8 + 100*t^6 + 20*t^4 + 9*t^2 + 1
> ZetaG([1,2,3,4,5,6],7);
16807*t^10 + 40817*t^9 + 50421*t^8 + 40572*t^7 + 23478*t^6
+ 10182*t^5 + 3354*t^4 + 828*t^3 + 147*t^2 + 17*t + 1
> ZetaG([1,2,3,4,5,6],11);
161051*t^10 + 14641*t^9 + 3993*t^8 + 2420*t^7 + 539*t^6
+ 14*t^5 + 49*t^4 + 20*t^3 + 3*t^2 + t + 1
> ZetaG([1,2,3,4,5,6],13);
371293*t^10 + 542659*t^9 + 432809*t^8 + 234572*t^7 + 94718*t^6
+ 29666*t^5 + 7286*t^4 + 1388*t^3 + 197*t^2 + 19*t + 1
```

and the (SageMath) factorizations (over $\mathbb{Z}[t]$)

(1) `(5*t^2 + 1) * (625*t^8 + 100*t^6 + 4*t^2 + 1)` at $p = 5$,

(2) `(7*t^2 + t + 1) * (7*t^2 + 4*t + 1)^4` at $p = 7$,

(3) `irreducible` at $p = 11$, and

(4) `(13*t^2 + 1) * (13*t^2 + 7*t + 1) * (13*t^2 + 4*t + 1)^3` at $p = 13$,

we obtain the polynomials $P_{5\mathbb{Z}}, P_{7\mathbb{Z}}, P_{11\mathbb{Z}}, P_{13\mathbb{Z}}$ as quoted in the proof of Proposition A.2.1.

A.3.2 Number field subfield computations

The SageMath code

```
max_tested = 0; max_discovered = 0
for n in range(1,2*100^2):
    max_tested = n
    for d in range(1,10):
        if (10*d) % euler_phi(n) == 0:
            max_discovered = n; break
print(max_tested,max_discovered)
```

yields `19999 330`, showing that

$$\max\{n \in [1, 2 \cdot 100^2 - 1] : \exists\, d \in \{1, 2, \ldots, 9\} \text{ such that } \phi(n) \mid 10d\} = 330.$$

We now study Q_{11}. The SageMath code

```
max_tested = 0; R.<t> = QQ[]
Q_11 = ZetaG([1,2,3,4,5,6],11); K.<a> = NumberField(Q_11)
for n in range(1,500):
    max_tested = n; L = K.subfield(a^n)[0]
    if L.absolute_degree() != 10:
        print(n)
print("Tested up to",max_tested)
```

yields `Tested up to 499`, showing that for each positive integer $n < 500$, the field $\mathbb{Q}[t]/Q_{11}$ is generated by $\bar{t}^n$ (i.e. $\bar{t}^n$ is a *primitive element* for $\mathbb{Q}[t]/Q_{11}$). Meanwhile,

```
R.<t> = QQ[]; Q_11 = ZetaG([1,2,3,4,5,6],11)
K.<a> = NumberField(Q_11)
print(K.is_galois(),[Z[0].absolute_degree() for Z in K.subfields()])
```

yields `False [1, 5, 10]`, showing that the field $L_{11} := \mathbb{Q}[t]/Q_{11}$ is non-Galois, with exactly 1 nontrivial subfield L'_{11}—and furthermore, $[L'_{11} : \mathbb{Q}] = 5$.

We now study $Q_{5,2}$. The SageMath code

```
max_tested = 0; R.<t> = QQ[]
Q_52 = 625*t^8 + 100*t^6 + 4*t^2 + 1; K.<a> = NumberField(Q_52)
for n in range(1,500,2):
    max_tested = n; L = K.subfield(a^n)[0]
    if L.absolute_degree() != 8:
        print(n)
print("Tested odds up to",max_tested)
```

yields `Tested odds up to 499`, showing that for each positive *odd* integer $n < 500$, the field $\mathbb{Q}[t]/Q_{5,2}$ is generated by $\bar{t}^n$.

Finally, we study $Q_{5^2,2}$. The SageMath code

```
max_tested = 0; R.<t> = QQ[]
Q_522 = 625*t^4 + 100*t^3 + 4*t^1 + 1; K.<a> = NumberField(Q_522)
for n in range(1,500):
    max_tested = n; L = K.subfield(a^n)[0]
    if L.absolute_degree() != 4:
        print(n)
print("Tested up to",max_tested)
```

yields `Tested up to 499`, showing that for each positive integer $n < 500$, the field $\mathbb{Q}[t]/Q_{5^2,2}$ is generated by $\bar{t}^n$.

Appendix B

Sharp cutoffs

Let $F \in \mathbb{Z}[\boldsymbol{x}] = \mathbb{Z}[x_1, \ldots, x_6]$ be a $\mathbb{P}^5_{\mathbb{Q}}$-smooth 6-variable cubic form. Recall, from Definition 1.4.6, the singular series $\mathfrak{S}_F$, the set $C(\mathrm{SSV})$, and the real density $\sigma_{\infty,F,w} := \lim_{\epsilon\to 0} (2\epsilon)^{-1} \int_{|F(\boldsymbol{x})|\le\epsilon} d\boldsymbol{x}\, w(\boldsymbol{x})$ for each $w \in C_c^\infty(\mathbb{R}^6)$. Let $C(T) := \bigcup_{L\in C(\mathrm{SSV})} L \subseteq \mathbb{Q}^6$. For each $w \in C_c^\infty(\mathbb{R}^6)$, let $N'_{F,w}(X) := \sum_{\boldsymbol{x}\in\mathbb{Z}^6\setminus C(T)} w(\boldsymbol{x}/X) \cdot \mathbf{1}_{F(\boldsymbol{x})=0}$.

Let K be a compact semi-algebraic subset of $\mathbb{R}^6$. In analogy with $N'_{F,w}(X)$, let

$$N'_{F,K}(X) := \#\{\boldsymbol{x} \in (\mathbb{Z}^6 \cap XK) \setminus C(T) : F(\boldsymbol{x}) = 0\}.$$

For $\delta > 0$ and $\boldsymbol{a} \in \mathbb{R}^6$, let $B_\delta(\boldsymbol{a})$ denote the closed Euclidean ball of radius δ centered at $\boldsymbol{a}$. Now—and this is a subtle point if K is not "sufficiently transverse" to $\{\boldsymbol{x} \in \mathbb{R}^6 : F(\boldsymbol{x}) = 0\} = C(V)(\mathbb{R})$—let $K_\delta := \bigcup_{\boldsymbol{a}\in K} B_\delta(\boldsymbol{a})$, and define $\sigma_{\infty,F,K}$ to be the iterated limit

$$\lim_{\delta\to 0} \left(\lim_{\epsilon\to 0} (2\epsilon)^{-1} \operatorname{vol}\{\boldsymbol{x} \in K_\delta : |F(\boldsymbol{x})| \le \epsilon\} \right).$$

I believe that $\sigma_{\infty,F,K}$ should exist and be finite, and that the following statement should hold, but have not checked carefully.

Conjecture B.0.1. *Fix $\eta > 0$. Suppose $\lim_{X\to\infty}(N'_{F,w}(X)/X^3) = \sigma_{\infty,F,w}\mathfrak{S}_F$ holds for all $w \in C_c^\infty(\mathbb{R}^6)$ with $\operatorname{Supp} w \subseteq K_\eta$. Then $\lim_{X\to\infty}(N'_{F,K}(X)/X^3) = \sigma_{\infty,F,K}\mathfrak{S}_F$.*

Plausible proof sketch. For simplicity, assume $\mathbf{0} \notin K$. (I believe the case $\mathbf{0} \in K$ should follow from the case $\mathbf{0} \notin K$ without too much trouble.)

To prove $\limsup_{X\to\infty}(N'_{F,K}(X)/X^3) \le \sigma_{\infty,F,K}\mathfrak{S}_F$, take decreasing opens $U_i \to K$ and use the hypothesis for suitable weights $w_i \in C_c^\infty(\mathbb{R}^6)$ that are 1 on K and 0 outside U_i; I expect that $\lim_{i\to\infty} \sigma_{\infty,F,w_i} = \sigma_{\infty,F,K}$.

To finish, it should suffice to bound the "error" $(N'_{F,w_i}(X) - N'_{F,K}(X))/X^3$ by $o_{i\to\infty}(1)$ as $X \to \infty$. To do so, take $\epsilon_i \to 0$ and use weights w'_i that are 1 on $\overline{U_i \cap (C(V)(\mathbb{R}) \setminus K)}$ and 0 outside an ϵ_i-neighborhood thereof. Then

$$0 \le N'_{F,w_i}(X) - N'_{F,K}(X) \le N'_{F,w'_i}(X).$$

I expect that (since $C(V)$ is smooth away from $\mathbf{0}$, and K is semi-algebraic) the boundary of $C(V)(\mathbb{R}) \setminus K$ in $C(V)(\mathbb{R})$ is *relatively* null (for any reasonable measure on $C(V)(\mathbb{R})$), and that consequently $\sigma_{\infty,F,w'_i} = o_{i\to\infty}(1)$. □

Remark B.0.2. When choosing a region K, we must avoid the situation $K = \mathbb{Q}^6 \cap [-1, 1]^6$, and possibly also (though I have not checked carefully) the situation where $\{\boldsymbol{x} \in K : F(\boldsymbol{x}) = 0\}$ is like a fat Cantor set, say. So some restriction on K (like the present "compact semi-algebraic" assumption) is necessary or at least convenient.

Remark B.0.3. If correct, the ideas in this appendix should apply much more generally to most Manin-type settings. But of course, the hard part is still counting with *any* reasonable weight at all! I just wanted to record some thoughts and subtleties that I have not seen thoroughly discussed elsewhere.

Bibliography

[AC93] P. Aluffi and F. Cukierman, *Multiplicities of discriminants*, Manuscripta Math. **78** (1993), no. 3, 245–258. MR1206155 ↑59

[Alu95] P. Aluffi, *Singular schemes of hypersurfaces*, Duke Math. J. **80** (1995), no. 2, 325–351. MR1369396 ↑59

[And86] G. W. Anderson, *Cyclotomy and an extension of the Taniyama group*, Compositio Math. **57** (1986), no. 2, 153–217. MR827351 ↑98

[Bak89] R. C. Baker, *Diagonal cubic equations. II*, Acta Arith. **53** (1989), no. 3, 217–250. MR1032826 ↑10, 91

[BB21] D. Bonolis and T. D. Browning, *Uniform bounds for rational points on hyperelliptic fibrations* (September 27, 2021), available at `https://arxiv.org/abs/2007.14182v2`. ↑37

[BBS18] V. Blomer, J. Brüdern, and P. Salberger, *The Manin-Peyre formula for a certain biprojective threefold*, Math. Ann. **370** (2018), no. 1-2, 491–553. MR3747494 ↑103

[BC20] S. Bettin and J. B. Conrey, *Averages of long Dirichlet polynomials* (June 22, 2020), available at `https://arxiv.org/abs/2002.09466v2`. ↑82

[BCGP21] G. Boxer, F. Calegari, T. Gee, and V. Pilloni, *Abelian surfaces over totally real fields are potentially modular*, Publ. Math. Inst. Hautes Études Sci. **134** (2021), 153–501. MR4349242 ↑95, 113

[Bel13] J. Bellaïche, *etale cohomology of an abelian variety and its dual*, MathOverflow, 2013. URL: `https://mathoverflow.net/q/132271` (version: 2013-05-29). ↑112

[BGZZK21] S. Basu, S. Guo, R. Zhang, and P. Zorin-Kranich, *A stationary set method for estimating oscillatory integrals* (October 12, 2021), available at `https://arxiv.org/abs/2103.08844v3`. ↑70

[Bha14] M. Bhargava, *The geometric sieve and the density of squarefree values of invariant polynomials* (January 31, 2014), available at `https://arxiv.org/abs/1402.0031v1`. ↑73, 75

[BHB06] T. D. Browning and D. R. Heath-Brown, *The density of rational points on non-singular hypersurfaces. II*, Proc. London Math. Soc. (3) **93** (2006), no. 2, 273–303. With an appendix by J. M. Starr. MR2251154 ↑60, 63

[BHB20] ———, *Density of rational points on a quadric bundle in $\mathbb{P}^3 \times \mathbb{P}^3$*, Duke Math. J. **169** (2020), no. 16, 3099–3165. MR4167086 ↑103

[Bir04] B. Birch, *Heegner points: the beginnings*, Heegner points and Rankin L-series, 2004, pp. 1–10. MR2083207 ↑9

[BKS19] T. D. Browning, V. V. Kumaraswamy, and R. S. Steiner, *Twisted Linnik implies optimal covering exponent for S^3*, Int. Math. Res. Not. IMRN **1** (2019), 140–164. MR3897426 ↑36

[Bom09] E. Bombieri, *Problems and results on the distribution of algebraic points on algebraic varieties*, J. Théor. Nombres Bordeaux **21** (2009), no. 1, 41–57. MR2537702 ↑67, 101

[Boo19] A. R. Booker, *Cracking the problem with 33*, Res. Number Theory **5** (2019), no. 3, Paper No. 26, 6. MR3983550 ↑8

[Brü91] J. Brüdern, *Ternary additive problems of Waring's type*, Math. Scand. **68** (1991), no. 1, 27–45. MR1124817 ↑102, 103

[Bri15] M. Bright, *Intersection of two components and Jacobian criterion*, MathOverflow, 2015. URL: `https://mathoverflow.net/q/211153` (version: 2015-07-09). ↑46

[Bro09] T. D. Browning, *Quantitative arithmetic of projective varieties*, Progress in Mathematics, vol. 277, Birkhäuser Verlag, Basel, 2009. MR2559866 ↑31, 67, 100, 108

[BS21] A. R. Booker and A. V. Sutherland, *On a question of Mordell*, Proc. Natl. Acad. Sci. USA **118** (2021), no. 11, Paper No. 2022377118, 11. MR4279690 ↑8

[BSD67] E. Bombieri and P. Swinnerton-Dyer, *On the local zeta function of a cubic threefold*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3) **21** (1967), 1–29. MR212019 ↑46, 54

[BT16] M. Bernardara and G. Tabuada, *Study of Chow groups of intersections of quadrics via homological projective duality and (Jacobians of) non-commutative motives*, Izv. Ross. Akad. Nauk Ser. Mat. **80** (2016), no. 3, 3–22. MR3507384 ↑105

[BW19] J. Brüdern and T. D. Wooley, *An instance where the major and minor arc integrals meet*, Bull. Lond. Math. Soc. **51** (2019), no. 6, 1113–1128. MR4041016 ↑67

[CFK+05] J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein, and N. C. Snaith, *Integral moments of L-functions*, Proc. London Math. Soc. (3) **91** (2005), no. 1, 33–104. MR2149530 ↑82, 88, 108

[CFLS22] P. J. Cho, D. Fiorilli, Y. Lee, and A. Södergren, *Omega results for cubic field counts via lower-order terms in the one-level density* (March 7, 2022), available at `https://arxiv.org/abs/2102.08077v2`. ↑82

[CFZ08] J. B. Conrey, D. W. Farmer, and M. R. Zirnbauer, *Autocorrelation of ratios of L-functions*, Commun. Number Theory Phys. **2** (2008), no. 3, 593–636. MR2482944 ↑76, 81, 82, 88, 96, 97

[CG66] J. W. S. Cassels and M. J. T. Guy, *On the Hasse principle for cubic surfaces*, Mathematika **13** (1966), 111–120. MR211966 ↑8, 9, 108

[CG72] C. H. Clemens and P. A. Griffiths, *The intermediate Jacobian of the cubic threefold*, Ann. of Math. (2) **95** (1972), 281–356. MR302652 ↑100

[CML09] S. Casalaina-Martin and R. Laza, *The moduli space of cubic threefolds via degenerations of the intermediate Jacobian*, J. Reine Angew. Math. **633** (2009), 29–65. MR2561195 ↑44

[CP15] F. Charles and A. Pirutka, *La conjecture de Tate entière pour les cubiques de dimension quatre*, Compos. Math. **151** (2015), no. 2, 253–264. MR3314826 ↑112

[CS07] J. B. Conrey and N. C. Snaith, *Applications of the L-functions ratios conjectures*, Proc. Lond. Math. Soc. (3) **94** (2007), no. 3, 594–646. MR2325314 ↑82, 96

[CS20] D. Corwin and T. Schlank, *Brauer and Etale Homotopy Obstructions to Rational Points on Open Covers* (August 16, 2020), available at `https://arxiv.org/abs/2006.11699v3`. ↑9

[CTW12] J.-L. Colliot-Thélène and O. Wittenberg, *Groupe de Brauer et points entiers de deux familles de surfaces cubiques affines*, Amer. J. Math. **134** (2012), no. 5, 1303–1327. MR2975237 ↑9

[CTWX20] J.-L. Colliot-Thélène, D. Wei, and F. Xu, *Brauer-Manin obstruction for Markoff surfaces*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **21** (2020), 1257–1313. MR4288633 ↑9

[Dao10] H. Dao, *Can projective hypersurfaces contain linear spaces? How big?*, MathOverflow, 2010. URL: `https://mathoverflow.net/q/15111` (version: 2010-02-12). ↑60

[Dao22] ———, *When is the product of two ideals equal to their intersection?*, MathOverflow, 2022. URL: `https://mathoverflow.net/q/49299` (version: 2022-01-04). ↑47

[Dav05] H. Davenport, *Analytic methods for Diophantine equations and Diophantine inequalities*, Second, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2005. With a foreword by R. C. Vaughan, D. R. Heath-Brown and D. E. Freeman, Edited and prepared for publication by T. D. Browning. MR2152164 ↑19

[Dav39] ———, *On Waring's problem for cubes*, Acta Math. **71** (1939), 123–143. MR0000026 ↑8, 15

[Deb03] O. Debarre, *Lines on smooth hypersurfaces* (November 2003), available at `https://www.math.ens.fr/~debarre/Lines_hypersurfaces.pdf`. ↑60

[Dem66] V. A. Demjanenko, *Sums of four cubes*, Izv. Vysš. Učebn. Zaved. Matematika **1966** (1966), no. 5 (54), 64–69. MR0202663 ↑8

[DFI12] W. Duke, J. B. Friedlander, and H. Iwaniec, *Weyl sums for quadratic roots*, Int. Math. Res. Not. IMRN **11** (2012), 2493–2549. MR2926988 ↑75

[DFI93] ———, *Bounds for automorphic L-functions*, Invent. Math. **112** (1993), no. 1, 1–8. MR1207474 ↑6, 7, 11, 23, 24, 36, 75, 77, 105

[DHL06] J.-M. Deshouillers, F. Hennecart, and B. Landreau, *On the density of sums of three cubes*, Algorithmic number theory, 2006, pp. 141–155. MR2282921 ↑14

[Dia19] S. Diaconu, *On admissible integers of cubic forms*, Senior Thesis, Princeton University, 2019. URL: `http://arks.princeton.edu/ark:/88435/dsp0112579w10h`. ↑10, 13, 14, 15, 18, 19, 20, 21, 91, 106

[Die17] R. Dietmann, *On the h-invariant of cubic forms, and systems of cubic forms*, Q. J. Math. **68** (2017), no. 2, 485–501. MR3667211 ↑56

[Dim90] A. Dimca, *Betti numbers of hypersurfaces and defects of linear systems*, Duke Math. J. **60** (1990), no. 1, 285–298. MR1047124 ↑46

[Dja12] G. Djanković, *A larger GL_2 large sieve in the level aspect*, Cent. Eur. J. Math. **10** (2012), no. 2, 748–760. MR2886570 ↑35

[DL62] H. Davenport and D. J. Lewis, *Exponential sums in many variables*, Amer. J. Math. **84** (1962), 649–665. MR144862 ↑56

[DL64] ———, *Non-homogeneous cubic equations*, J. London Math. Soc. **39** (1964), 657–671. MR167458 ↑56

[DLR17] O. Debarre, A. Laface, and X. Roulleau, *Lines on cubic hypersurfaces over finite fields*, Geometry over nonclosed fields, 2017, pp. 19–51. Refer to the (slightly expanded and corrected) version of January 28, 2021: `https://arxiv.org/abs/1510.05803v5`. MR3644249 ↑50, 51, 52, 54, 94, 95, 112, 117

[Dol12] I. V. Dolgachev, *Classical algebraic geometry*, Cambridge University Press, Cambridge, 2012. A modern view. MR2964027 ↑58, 59, 60, 61

[Dol16] I. Dolgachev, *Corrado Segre and nodal cubic threefolds*, From classical to modern algebraic geometry, 2016, pp. 429–450. MR3776662 ↑47

[DR21] A. Dunn and M. Radziwiłł, *Bias in cubic Gauss sums: Patterson's conjecture* (September 15, 2021), available at `https://arxiv.org/abs/2109.07463v1`. ↑35

[DSP90] W. Duke and R. Schulze-Pillot, *Representation of integers by positive ternary quadratic forms and equidistribution of lattice points on ellipsoids*, Invent. Math. **99** (1990), no. 1, 49–57. MR1029390 ↑15

[DW21] A. Diaconu and I. Whitehead, *On the third moment of $L(\frac{1}{2}, \chi_d)$ II: the number field case*, J. Eur. Math. Soc. (JEMS) **23** (2021), no. 6, 2051–2070. MR4244522 ↑82

[EH87] D. Eisenbud and J. Harris, *On varieties of minimal degree (a centennial account)*, Algebraic geometry, Bowdoin, 1985 (Brunswick, Maine, 1985), 1987, pp. 3–13. MR927946 ↑48

[Elk01] N. Elkies, *Complete cubic parametrization of the Fermat cubic surface $w^3 + x^3 + y^3 + z^3 = 0$*, 2001. URL: `https://people.math.harvard.edu/~elkies/4cubes.html` (accessed 2022-04-13); based on a 1997-06-19 post archived at `https://web.archive.org/web/20030113052115/http://www.math.niu.edu/~rusin/known-math/96/4cubes`. ↑14

[FI13] J. B. Friedlander and H. Iwaniec, *Small representations by indefinite ternary quadratic forms*, Number theory and related fields, 2013, pp. 157–164. MR3081039 ↑103

[FPRS19] D. W. Farmer, A. Pitale, N. C. Ryan, and R. Schmidt, *Analytic L-functions: definitions, theorems, and connections*, Bull. Amer. Math. Soc. (N.S.) **56** (2019), no. 2, 261–280. MR3923345 ↑79, 92

[GL02] S. R. Ghorpade and G. Lachaud, *Étale cohomology, Lefschetz theorems and number of points of singular varieties over finite fields*, 2002, pp. 589–631. Dedicated to Yuri I. Manin on the occasion of his 65th birthday. Refer to the (corrected, revised and updated) version of August 15, 2008: `https://arxiv.org/abs/0808.2169v1`. MR1988974 ↑49, 112

[Gol96] E. P. Golubeva, *On nonhomogeneous Waring equations*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **226** (1996), no. Anal. Teor. Chisel i Teor. Funktsiĭ. 13, 65–68, 236. Refer to the English translation in J. Math. Sci. (New York) **89** (1998), no. 1, 955–957. MR1433348 ↑36

[Gra98] A. Granville, *ABC allows us to count squarefrees*, Internat. Math. Res. Notices **19** (1998), 991–1009. MR1654759 ↑73

[Gro16] B. H. Gross, *On the Langlands correspondence for symplectic motives*, Izv. Ross. Akad. Nauk Ser. Mat. **80** (2016), no. 4, 49–64. MR3535358 ↑94, 95

[Gro72] A. Grothendieck, *Groupes de monodromie en géométrie algébrique. I*, Lecture Notes in Mathematics, Vol. 288, Springer-Verlag, Berlin-New York, 1972. Séminaire de Géométrie Algébrique du Bois-Marie 1967–1969 (SGA 7 I), Dirigé par A. Grothendieck. Avec la collaboration de M. Raynaud et D. S. Rim. MR0354656 ↑50

[GS17] A. Ghosh and P. Sarnak, *Integral points on Markoff type cubic surfaces* (September 27, 2017), available at `https://arxiv.org/abs/1706.06712v2`. ↑13

[GS21] L. Grimmelt and W. Sawin, *Representation of squares by nonsingular cubic forms*, Israel J. Math. **242** (2021), no. 2, 501–547. MR4282090 ↑44

[Har13] A. J. Harper, *Sharp conditional bounds for moments of the Riemann zeta function* (May 20, 2013), available at `https://arxiv.org/abs/1305.4618v1`. ↑82

[Har17] Y. Harpaz, *Geometry and arithmetic of certain log K3 surfaces*, Ann. Inst. Fourier (Grenoble) **67** (2017), no. 5, 2167–2200. MR3732687 ↑8, 109

[Has96] B. Hassett, *Special cubic hypersurfaces of dimension four*, Ph.D. Thesis, Harvard University, 1996. MR2694332 ↑47

[HB07] D. R. Heath-Brown, *Cubic forms in 14 variables*, Invent. Math. **170** (2007), no. 1, 199–230. MR2336082 ↑39, 91

[HB10] ———, *Bounds for the cubic Weyl sum*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **377** (2010), no. Issledovaniya po Teorii Chisel. 10, 199–216, 244–245. MR2753656 ↑101

[HB22] ———, *The Distribution of Rational Points on Conics* (March 25, 2022), available at `https://arxiv.org/abs/2203.13889v1`. ↑103

[HB83] ———, *Cubic forms in ten variables*, Proc. London Math. Soc. (3) **47** (1983), no. 2, 225–257. MR703978 ↑29

[HB92] ———, *The density of zeros of forms for which weak approximation fails*, Math. Comp. **59** (1992), no. 200, 613–623. MR1146835 ↑7, 9, 15

[HB95] ———, *A mean value estimate for real character sums*, Acta Arith. **72** (1995), no. 3, 235–275. MR1347489 ↑35

[HB96] ———, *A new form of the circle method, and its application to quadratic forms*, J. Reine Angew. Math. **481** (1996), 149–206. MR1421949 ↑6, 7, 11, 15, 16, 23, 24, 30, 31, 36, 70, 75, 77, 105

[HB98] ———, *The circle method and diagonal cubic forms*, R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci. **356** (1998), no. 1738, 673–699. MR1620820 ↑7, 23, 24, 25, 27, 28, 29, 32, 33, 40, 64, 66, 70, 72, 77, 85, 92, 94, 101, 107

[HB99] ———, *The solubility of diagonal cubic Diophantine equations*, Proc. London Math. Soc. (3) **79** (1999), no. 2, 241–259. MR1702242 ↑91

[HBP17] D. R. Heath-Brown and L. B. Pierce, *Simultaneous integer values of pairs of quadratic forms*, J. Reine Angew. Math. **727** (2017), 85–143. MR3652248 ↑104, 105

[Hee52] K. Heegner, *Diophantische Analysis und Modulfunktionen*, Math. Z. **56** (1952), 227–253. MR53135 ↑9

[HL25] G. H. Hardy and J. E. Littlewood, *Some problems of 'Partitio numerorum' (VI): Further researches in Waring's Problem*, Math. Z. **23** (1925), no. 1, 1–37. MR1544728 ↑6, 13, 14

[Hoo14] C. Hooley, *On octonary cubic forms*, Proc. Lond. Math. Soc. (3) **109** (2014), no. 1, 241–281. MR3237742 ↑26, 30, 59, 74

[Hoo15] ———, *On octonary cubic forms: II*, Bull. Lond. Math. Soc. **47** (2015), no. 1, 85–94. MR3312967 ↑67

[Hoo16] ———, *On the representation of numbers by quaternary and quinary cubic forms: I*, Acta Arith. **173** (2016), no. 1, 19–39. MR3494132 ↑8, 13

[Hoo79] ———, *On a new technique and its applications to the theory of numbers*, Proc. London Math. Soc. (3) **38** (1979), no. 1, 115–151. MR520975 ↑109

[Hoo81] ———, *On Waring's problem for two squares and three cubes*, J. Reine Angew. Math. **328** (1981), 161–207. MR636202 ↑109

[Hoo84] C. Hooley, *Some recent advances in analytical number theory*, Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Warsaw, 1983), 1984, pp. 85–97. MR804678 ↑109

[Hoo86a] C. Hooley, *On some topics connected with Waring's problem*, J. Reine Angew. Math. **369** (1986), 110–153. MR850631 ↑7, 10, 67, 109

[Hoo86b] ———, *On Waring's problem*, Acta Math. **157** (1986), no. 1-2, 49–97. MR857679 ↑7, 8, 25, 26, 27, 28, 29, 32, 33, 40, 70, 72, 77, 92, 101, 105, 107

[Hoo88] ———, *On nonary cubic forms*, J. Reine Angew. Math. **386** (1988), 32–98. MR936992 ↑10, 59, 60

[Hoo91] ———, *On the number of points on a complete intersection over a finite field*, J. Number Theory **38** (1991), no. 3, 338–358. With an appendix by Nicholas M. Katz. MR1114483 ↑41, 42, 43, 49

[Hoo97] ———, *On Hypothesis K^* in Waring's problem*, Sieve methods, exponential sums, and their applications in number theory (Cardiff, 1995), 1997, pp. 175–185. MR1635754 ↑7, 8, 14, 25, 27, 28, 32, 33, 40, 70, 72, 77, 101, 105, 107

[HT10] B. Hassett and Y. Tschinkel, *Flops on holomorphic symplectic fourfolds and determinantal cubic hypersurfaces*, J. Inst. Math. Jussieu **9** (2010), no. 1, 125–153. MR2576800 ↑47

[Hua38] L.-K. Hua, *On Waring's problem*, Quart. J. Math. Oxford Ser. **9** (1938), no. 1, 199–202. ↑7

[Hui16] J. Huisman, *Endomorphism ring of an abelian variety and its reduction mod $\mathfrak{p}$*, Mathematics Stack Exchange, 2016. URL: `https://math.stackexchange.com/q/1834526` (version: 2016-06-21). ↑114

[IK04] H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004. MR2061214 ↑6, 35, 37, 81, 92, 93, 95, 98, 105

[IR90] K. Ireland and M. Rosen, *A classical introduction to modern number theory*, Second, Graduate Texts in Mathematics, vol. 84, Springer-Verlag, New York, 1990. MR1070716 ↑97

[Jah14] J. Jahnel, *Brauer groups, Tamagawa measures, and rational points on algebraic varieties*, Mathematical Surveys and Monographs, vol. 198, American Mathematical Society, Providence, RI, 2014. MR3242964 ↑67, 95, 99, 100, 108

[Kat04] N. M. Katz, *Larsen's alternative, moments, and the monodromy of Lefschetz pencils*, Contributions to automorphic forms, geometry, and number theory, 2004, pp. 521–560. MR2058618 ↑96

[Kat13] ———, *Sato-Tate in the higher dimensional case: elaboration of 9.5.4 in Serre's $N_X(p)$ book*, Enseign. Math. **59** (2013), no. 3-4, 359–377. MR3189042 ↑96

[Kis99] M. Kisin, *Local constancy in p-adic families of Galois representations*, Math. Z. **230** (1999), no. 3, 569–593. MR1680032 ↑76, 78

[Klo22] R. Kloosterman, *Maximal families of nodal varieties with defect*, Math. Z. **300** (2022), no. 2, 1141–1156. MR4363772 ↑46

[Klo26] H. D. Kloosterman, *On the representation of numbers in the form $ax^2 + by^2 + cz^2 + dt^2$*, Acta Math. **49** (1926), no. 3-4, 407–464. MR1555249 ↑11

[Kol02] J. Kollár, *Unirationality of cubic hypersurfaces*, J. Inst. Math. Jussieu **1** (2002), no. 3, 467–476. MR1956057 ↑8

[Kon02] A. Kontogeorgis, *Automorphisms of Fermat-like varieties*, Manuscripta Math. **107** (2002), no. 2, 187–205. MR1894739 ↑63

[Kow21] J. M. Kowalski, *On the proportion of squarefree numbers among sums of cubic polynomials*, Ramanujan J. **54** (2021), no. 2, 343–354. MR4204760 ↑75

[Kum18] V. V. Kumaraswamy, *Zeros of quadratic forms and the delta method*, Ph.D. Thesis, University of Bristol, 2018. URL: `https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.761126`. ↑111

[KW01] R. Kiehl and R. Weissauer, *Weil conjectures, perverse sheaves and l'adic Fourier transform*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 42, Springer-Verlag, Berlin, 2001. MR1855066 ↑49

[Laf16] A. Laface, *Counting $\mathbb{F}_q$-points on smooth cubic threefolds with a given rational line*, GitHub, 2016. URL: `https://github.com/alaface/CubLin` (version: `commit 6707b3e`). ↑51, 117

[Lan15] G. Lando, *Square-free values of polynomials evaluated at primes over a function field*, Q. J. Math. **66** (2015), no. 3, 905–924. MR3396097 ↑73

[Lan90] S. Lang, *Cyclotomic fields I and II*, second, Graduate Texts in Mathematics, vol. 121, Springer-Verlag, New York, 1990. With an appendix by Karl Rubin. MR1029028 ↑99

[Las17] A. Laskar, *On local monodromy filtrations attached to motives*, Int. J. Number Theory **13** (2017), no. 3, 801–817. MR3606955 ↑92

[Law19] S. Lawton, *Polynomials (or analytic functions) vanishing on a real algebraic set*, MathOverflow, 2019. URL: `https://mathoverflow.net/q/331205` (version: 2019-05-10). ↑60

[LB15] P. Le Boudec, *Density of rational points on a certain smooth bihomogeneous threefold*, Int. Math. Res. Not. IMRN **21** (2015), 10703–10715. MR3456025 ↑103

[Lin20] N. Lindner, *Hypersurfaces with defect*, J. Algebra **555** (2020), 1–35. MR4081491 ↑42, 44

[Lin43] U. V. Linnik, *On the representation of large numbers as sums of seven cubes*, Rec. Math. [Mat. Sbornik] N. S. **12(54)** (1943), 218–224. MR0009388 ↑9, 36

[LM21] D. Loughran and V. Mitankin, *Integral Hasse principle and strong approximation for Markoff surfaces*, Int. Math. Res. Not. IMRN **18** (2021), 14086–14122. MR4320804 ↑9

[Lou14] B. Louvel, *On the distribution of cubic exponential sums*, Forum Math. **26** (2014), no. 4, 987–1028. MR3228924 ↑35, 113

[LST19] B. Lehmann, A. K. Sengupta, and S. Tanimoto, *Geometric consistency of Manin's Conjecture* (June 3, 2019), available at `https://arxiv.org/abs/1805.10580v2`. ↑67

[LX15] Q. Liu and F. Xu, *Very strong approximation for certain algebraic varieties*, Math. Ann. **363** (2015), no. 3-4, 701–731. MR3412340 ↑9

[Mah36] K. Mahler, *Note on Hypothesis K of Hardy and Littlewood*, J. London Math. Soc. **11** (1936), no. 2, 136–138. MR1574761 ↑13, 40

[Man86] Yu. I. Manin, *Cubic forms*, Second, North-Holland Mathematical Library, vol. 4, North-Holland Publishing Co., Amsterdam, 1986. Algebra, geometry, arithmetic, Translated from the Russian by M. Hazewinkel. MR833513 ↑9, 95

[Mas96] D. W. Masser, *Specializations of endomorphism rings of abelian varieties*, Bull. Soc. Math. France **124** (1996), no. 3, 457–476. MR1415735 ↑113

[Mil04] S. J. Miller, *One- and two-level densities for rational families of elliptic curves: evidence for the underlying group symmetries*, Compos. Math. **140** (2004), no. 4, 952–992. MR2059225 ↑73, 77

[Mil17] J. S. Milne, *Algebraic groups*, Cambridge Studies in Advanced Mathematics, vol. 170, Cambridge University Press, Cambridge, 2017. The theory of group schemes of finite type over a field. MR3729270 ↑92

[MM21] Yu. I. Manin and M. Marcolli, *Homotopy Spectra and Diophantine Equations* (May 2, 2021), available at `https://arxiv.org/abs/2101.00197v2`. ↑77

[Mum70] D. Mumford, *Varieties defined by quadratic equations*, Questions on Algebraic Varieties (C.I.M.E., III Ciclo, Varenna, 1969), 1970, pp. 29–100. MR0282975 ↑110

[Mun15] R. Munshi, *Pairs of quadrics in 11 variables*, Compos. Math. **151** (2015), no. 7, 1189–1214. MR3371491 ↑106

[Mur74] J. P. Murre, *Some results on cubic threefolds*, Classification of algebraic varieties and compact complex manifolds, 1974, pp. 140–160. Lecture Notes in Math., Vol. 412. MR0374145 ↑94

[Mus11] M. Mustaţă, *Zeta functions in algebraic geometry*, University of Michigan, 2011. URL: `http://www-personal.umich.edu/~mmustata/zeta_book.pdf` (accessed 2022-04-13). ↑62

[Mus17] ———, *Lecture 2. Cubic hypersurfaces I: Rationality of certain cubic hypersurfaces*, University of Michigan, 2017. URL: `http://www-personal.umich.edu/~mmustata/lecture2_rationality.pdf` (accessed 2022-03-29). ↑49

[MV19] O. Marmon and P. Vishe, *On the Hasse principle for quartic hypersurfaces*, Duke Math. J. **168** (2019), no. 14, 2727–2799. MR4012347 ↑24, 36, 38, 39, 101, 106, 110

[Pat97] S. J. Patterson, *The asymptotic distribution of Kloosterman sums*, Acta Arith. **79** (1997), no. 3, 205–219. MR1438824 ↑113

[Pet11] S. Petersen, *Semisimplicity of étale cohomology representations*, MathOverflow, 2011. URL: `https://mathoverflow.net/q/56378` (version: 2011-02-23). ↑112

[Poo03] B. Poonen, *Squarefree values of multivariable polynomials*, Duke Math. J. **118** (2003), no. 2, 353–373. MR1980998 ↑73

[Poo17] ———, *Rational points on varieties*, Graduate Studies in Mathematics, vol. 186, American Mathematical Society, Providence, RI, 2017. MR3729254 ↑8, 9, 49, 51, 95, 112

[PP97] A. Perelli and J. Pomykała, *Averages of twisted elliptic L-functions*, Acta Arith. **80** (1997), no. 2, 149–163. MR1450922 ↑35

[PS20] B. Poonen and M. Stoll, *The valuation of the discriminant of a hypersurface* (August 26, 2020), available at `https://math.mit.edu/~poonen/papers/discriminant.pdf`. ↑43, 44, 72

[Qu07] Y. Qu, *Selberg's normal density theorem for automorphic L-functions for* GL_m, Acta Math. Sin. (Engl. Ser.) **23** (2007), no. 10, 1903–1908. MR2352305 ↑107

[Rei72] M. Reid, *The complete intersection of two or more quadrics*, Ph.D. Thesis, University of Cambridge, 1972. URL: `https://homepages.warwick.ac.uk/~masda/3folds/qu.pdf`. ↑94, 105, 112

[Rob11] O. Robert, *Sur le nombre des entiers représentables comme somme de trois puissances*, Acta Arith. **149** (2011), no. 1, 1–21. MR2805539 ↑109

[Sal15] P. Salberger, *Uniform bounds for rational points on cubic hypersurfaces*, Arithmetic and geometry, 2015, pp. 401–421. MR3467132 ↑102

[Sar00] P. Sarnak, *Kloosterman, quadratic forms and modular forms*, Nieuw Arch. Wiskd. (5) **1** (2000), no. 4, 385–389. MR1804617 ↑5

[SD01] P. Swinnerton-Dyer, *The solubility of diagonal cubic surfaces*, Ann. Sci. École Norm. Sup. (4) **34** (2001), no. 6, 891–912. MR1872424 ↑9

[SD14] ———, *Diagonal hypersurfaces and the Bloch-Kato conjecture, I*, J. Lond. Math. Soc. (2) **90** (2014), no. 3, 845–860. MR3291803 ↑99

[SD67] ———, *An application of computing to class field theory*, Algebraic Number Theory (Proc. Instructional Conf., Brighton, 1965), 1967, pp. 280–291. MR0219514 ↑100

[Seg43] B. Segre, *A note on arithmetical properties of cubic surfaces*, J. London Math. Soc **18** (1943), 24–31. MR0009471 ↑8

[Sel89] A. Selberg, *Collected papers. Vol. I*, Springer-Verlag, Berlin, 1989. With a foreword by K. Chandrasekharan. MR1117906 ↑107

[Ser00] J.-P. Serre, *Local algebra*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2000. Translated from the French by CheeWhye Chin and revised by the author. MR1771925 ↑47

[Shi88] T. Shioda, *Arithmetic and geometry of Fermat curves*, Algebraic Geometry Seminar (Singapore, 1987), 1988, pp. 95–102. MR966448 ↑63

[Sko92] A. N. Skorobogatov, *Exponential sums, the geometry of hyperplane sections, and some Diophantine problems*, Israel J. Math. **80** (1992), no. 3, 359–379. MR1202578 ↑49

[Spe10] D. E. Speyer, *When is the product of two ideals equal to their intersection?*, MathOverflow, 2010. URL: `https://mathoverflow.net/q/49261` (version: 2010-12-13). ↑47

[SST16] P. Sarnak, S. W. Shin, and N. Templier, *Families of L-functions and their symmetry*, Families of automorphic forms and the trace formula, 2016, pp. 531–578. MR3675175 ↑76, 80, 81, 88, 92, 94, 96

[Sta16] J. Starr, *Restriction of the Picard group of a surface to a curve*, MathOverflow, 2016. URL: `https://mathoverflow.net/q/233170` (version: 2016-03-09). ↑99

[Sta22] T. Stacks project authors, *The Stacks project*, 2022. URL: `https://stacks.math.columbia.edu` (accessed 2022-03-30). ↑42, 114

[Tat66] J. Tate, *Endomorphisms of abelian varieties over finite fields*, Invent. Math. **2** (1966), 134–144. MR206004 ↑114

[Tay04] R. Taylor, *Galois representations*, Ann. Fac. Sci. Toulouse Math. (6) **13** (2004), no. 1, 73–119. MR2060030 ↑77, 78, 92

[Ter18] Y. Terakado, *The determinant and the discriminant of a complete intersection of even dimension*, Math. Res. Lett. **25** (2018), no. 1, 255–280. MR3818622 ↑43

[Tia17] Z. Tian, *Hasse principle for three classes of varieties over global function fields*, Duke Math. J. **166** (2017), no. 17, 3349–3424. MR3724220 ↑6, 91

[Tom13] Tom163, *Is first etale cohomology of a variety always (dual to) a Tate Module?*, MathOverflow, 2013. URL: `https://mathoverflow.net/q/135648` (version: 2013-07-03). ↑112

[Vau15] R. C. Vaughan, *Squares: additive questions and partitions*, Int. J. Number Theory **11** (2015), no. 5, 1367–1409. MR3376217 ↑102

[Vau20] ———, *On some questions of partitio numerorum: Tres cubi*, Glasgow Mathematical Journal (2020), 1–22. ↑7, 40

[Vau86] ———, *On Waring's problem for cubes*, J. Reine Angew. Math. **365** (1986), 122–170. MR826156 ↑7, 9

[Vau97] ———, *The Hardy-Littlewood method*, Second, Cambridge Tracts in Mathematics, vol. 125, Cambridge University Press, Cambridge, 1997. MR1435742 ↑30, 31

[vdGK10] G. van der Geer and A. Kouvidakis, *A note on Fano surfaces of nodal cubic threefolds*, Algebraic and arithmetic structures of moduli spaces (Sapporo 2007), 2010, pp. 27–45. MR2676156 ↑44

[Vis19] P. Vishe, *Rational points on complete intersections over* $\mathbb{F}_q(t)$ (July 16, 2019), available at `https://arxiv.org/abs/1907.07097v1`. ↑104, 105

[VW02] R. C. Vaughan and T. D. Wooley, *Waring's problem: a survey*, Number theory for the millennium, III (Urbana, IL, 2000), 2002, pp. 301–340. MR1956283 ↑36

[VW95] ———, *On a certain nonary cubic form and related equations*, Duke Math. J. **80** (1995), no. 3, 669–735. MR1370112 ↑10, 67

[Wan21a] V. Y. Wang, *Approaching cubic Diophantine statistics via mean-value L-function conjectures of Random Matrix Theory type* (August 7, 2021), available at `https://arxiv.org/abs/2108.03398v1`. ↑7, 8, 12, 69, 70, 71, 72, 73, 74, 75, 76, 82, 87, 90, 91

[Wan21b] ———, *Axiomatizing Bombieri–Huxley* (August 7, 2021), available at `https://wangyangvictor.github.io/thesis_links`. ↑32

[Wan21c] ———, *Diagonal cubic forms and the large sieve* (August 7, 2021), available at `https://arxiv.org/abs/2108.03395v1`. ↑25, 27, 32, 33, 34, 35, 64, 74, 103, 105, 107

[Wan21d] ———, *Isolating special solutions in the delta method: The case of a diagonal cubic equation in evenly many variables over* $\mathbb{Q}$ (August 7, 2021), available at `https://arxiv.org/abs/2108.03396v1`. ↑7, 25, 30, 56, 57, 58, 59, 66, 75, 108

[Wan21e] ———, *Restricted moments of cubic Weyl sums* (August 7, 2021), available at `https://wangyangvictor.github.io/thesis_links`. ↑102, 103

[Wan21f] ———, *Tools for analyzing cubic threefolds over finite fields*, GitHub, 2021. URL: `https://github.com/wangyangvictor/singular_cubic_threefolds` (version: `commit 5971de6`). ↑53

[Wan21g] ———, *When does density beat Hua?* (August 7, 2021), available at `https://wangyangvictor.github.io/thesis_links`. ↑29, 32

[Wan22] ———, *Dichotomous point counts over finite fields* (February 21, 2022), available at `https://arxiv.org/abs/2202.10427v1`. ↑25, 41, 43, 44, 46, 47, 48, 55

[Wei49] A. Weil, *Numbers of solutions of equations in finite fields*, Bull. Amer. Math. Soc. **55** (1949), 497–508. MR29393 ↑27, 45, 98

[Wei52] ———, *Jacobi sums as "Grössencharaktere"*, Trans. Amer. Math. Soc. **73** (1952), 487–495. MR51263 ↑98, 113

[Wil22] F. Wilsch, *Integral points of bounded height on a certain toric variety* (February 22, 2022), available at `https://arxiv.org/abs/2202.10909v1`. ↑8

[Woo00] T. D. Wooley, *Sums of three cubes*, Mathematika **47** (2000), no. 1-2, 53–61 (2002). MR1924487 ↑8

[Woo13] ———, *On Waring's problem: some consequences of Golubeva's method*, J. Lond. Math. Soc. (2) **88** (2013), no. 3, 699–715. MR3145127 ↑36

[Woo15] ———, *Sums of three cubes, II*, Acta Arith. **170** (2015), no. 1, 73–100. MR3373831 ↑8

[Woo19] ———, *A slice or two of a diagonal cubic: arithmetic stratification via the circle method*, Joint IAS/Princeton University Number Theory Seminar, 2019. URL: `https://youtu.be/r1nzdaFCMjA?t=1453` (accessed 2020-11-08). ↑41, 68

[Woo95] ———, *Breaking classical convexity in Waring's problem: sums of cubes and quasi-diagonal behaviour*, Invent. Math. **122** (1995), no. 3, 421–451. MR1359599 ↑8, 101