Integers expressible as the sum of two rational cubes

Levent Alpöge, Manjul Bhargava, and Ari Shnidman

(with an appendix by Ashay Burungale and Christopher Skinner)

Abstract

We prove that a positive proportion of integers are expressible as the sum of two rational cubes, and a positive proportion are not so expressible. More generally, we prove that a positive proportion (in fact, at least one sixth) of elliptic curves in any cubic twist family have rank 0, and a positive proportion (in fact, at least one sixth) of elliptic curves with good reduction at 2 in any cubic twist family have rank 1.

Our method involves proving that the average size of the 2-Selmer group of elliptic curves in any cubic twist family, having any given root number, is 3. We accomplish this by generalizing a parametrization, due to the second author and Ho, of elliptic curves with extra structure by pairs of binary cubic forms. We then count integer points satisfying suitable congruence conditions on a quadric hypersurface in the space of real pairs of binary cubic forms in a fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z}) \times \mathrm{SL}_2(\mathbb{Z})$, using a combination of geometry-of-numbers methods and the circle method, building on earlier work of Ruth and the first author. We give a new interpretation of the singular integral and singular series arising in the circle method in terms of real and p-adic integrals with respect to a natural $(\mathrm{SL}_2 \times \mathrm{SL}_2)$-invariant measure. A uniformity estimate and sieve then shows that the average size of the 2-Selmer group over the full cubic twist family is 3. After suitably partitioning the subset of curves in the family with given root number, we execute a further sieve to show that the root number is equidistributed and that the same average, now taken over only those curves of given root number, is also 3. Finally, we apply the p-parity theorem of Dokchitser–Dokchitser and a p-converse theorem of Burungale–Skinner to conclude the proof.

Contents

1 Introduction

It has long been known which numbers can be expressed as the sum of two rational squares. As was first observed by Girard in 1625 and Fermat in 1638, and finally proven by Euler in 1749 [16, pp. 227–231], they are those positive integers whose prime factorizations have all primes that are congruent to 3 (mod 4) occurring with even exponent. Nowadays, this can also be deduced from the Hasse–Minkowski local-global principle for quadratic forms. Using this precise description, we see that a density of 0% of integers are the sum of two rational squares. Moreover, an integer is the sum of two rational squares if and only if it is the sum of two integer squares.

In contrast, the integers that are the sum of two rational cubes do not seem to follow any simple pattern:

$$1, 2, 6, 7, 8, 9, 12, 13, 15, 16, 17, 19, 20, 22, 26, 27, 28, 30, 31, 33, 34, 35, \ldots$$

It is conjectured that these integers have positive density; indeed, based on predictions of Goldfeld [19], Katz–Sarnak [24], and Bektemirov–Mazur–Stein–Watkins [4], it is natural to conjecture that the integers that can be expressed as the sum of two rational cubes should have natural density exactly 1/2. However, it has not previously been known whether this density is even greater than 0 or even less than 1.

Unlike the case of the sum of two rational/integer squares, it is possible for an integer to be the sum of two rational cubes but *not* the sum of two integer cubes, the smallest example being

$$6 = \left(\frac{17}{21}\right)^3 + \left(\frac{37}{21}\right)^3.$$

In fact, it is easy to see that the integers that can be expressed as the sum of two integer cubes have density zero.[1]

The purpose of this paper is to prove that the density of integers expressible as the sum of two rational cubes is strictly positive and strictly less than 1. We note that there is never any local obstruction for an integer to be the sum of two rational cubes, so proving this theorem must necessarily involve global arguments.

1.1 Main results

We prove the following theorem:

Theorem 1.1. *When ordered by their absolute values, a positive proportion of integers are the sum of two rational cubes, and a positive proportion are not.*

More precisely, we prove that

$$\liminf_{X\to\infty} \frac{\#\{n \in \mathbb{Z} : |n| < X \text{ and } n \text{ is the sum of two rational cubes}\}}{\#\{n \in \mathbb{Z} \colon |n| < X\}} \geq \frac{2}{21} \tag{1.1}$$

and

$$\liminf_{X\to\infty} \frac{\#\{n \in \mathbb{Z} : |n| < X \text{ and } n \text{ is not the sum of two rational cubes}\}}{\#\{n \in \mathbb{Z} \colon |n| < X\}} \geq \frac{1}{6}. \tag{1.2}$$

In fact we will prove the stronger claim that among the cubic twists $x^3 + y^3 = nz^3$ of the Fermat cubic, at least 1/6 of twists have rank 0 and at least 2/21 have rank 1.

[1] If $|x^3+y^3| = |x+y| \cdot |x^2-xy+y^2| \leq X$ with $|x| \geq |y|$, then $|x+y| \ll X/|x|^2$; hence the number of $(x, y) \in \mathbb{Z} \times \mathbb{Z}$ with $0 \neq |x^3 + y^3| \leq X$ is $\ll \sum_{|x| \ll X^{1/3}} X^{1/3} + \sum_{X^{1/3} \ll |x| \ll X^{1/2}} X/|x|^2 \ll X^{\frac{2}{3}}$.

More generally, we consider general families of cubic twists of elliptic curves. Any cubic twist family of elliptic curves over $\mathbb{Q}$ takes the form $E_{d,n} : y^2 = x^3 + dn^2$, where $d \in \mathbb{Z}$ is fixed and $n \in \mathbb{Z}$ varies. Since the elliptic curve $x^3 + y^3 = n$ can be expressed in Weierstrass form as $y^2 = x^3 - 432n^2$, the family of twists of the Fermat cubic corresponds to the case $d = -432$.[2]

We prove the following generalization of Theorem 1.1.

Theorem 1.2. *Fix $d \neq 0$. Then, when n varies ordered by $|n|$, at least $1/6$ of the elliptic curves in the cubic twist family $E_{d,n} : y^2 = x^3 + dn^2$ have rank 0, and at least $1/6$ of the elliptic curves $E_{d,n}$ with good reduction at 2 have rank 1. In particular, if the squarefree part of d is congruent to $1 \pmod 4$, then a proportion of at least $\frac{1}{21}2^{r-1}$ of the curves $E_{d,n}$ have rank 1, where r is the least residue of $v_2(d)/2$ modulo 3.*

We prove Theorem 1.2 via a determination of the average size of the 2-Selmer group of elliptic curves (satisfying any finite—or any acceptable infinite—set of congruence conditions) in a cubic twist family. Let us say that a subset $\Sigma \subset \mathbb{Z}$ is *acceptable* if it is defined by congruence conditions modulo prime powers, where for sufficiently large p, the congruence conditions include all integers with p-adic valuation at most 1. Then we prove the following theorem.

Theorem 1.3. *Fix $d \neq 0$, and let $\Sigma \subset \mathbb{Z}$ be any acceptable subset. Then, when $n \in \Sigma$ varies ordered by $|n|$, the average size of $\mathrm{Sel}_2(E_{d,n})$ is 3.*

It follows immediately from Theorem 1.3 that the average rank of elliptic curves in the cubic twist family $E_{d,n}$ ($n \in \Sigma$) is bounded. Indeed, since $\operatorname{rk} E_{d,n} \leq 2^{\operatorname{rk} E_{d,n}-1} \leq \frac{1}{2}\#\mathrm{Sel}_2(E_{d,n})$, it follows that the average rank of $E_{d,n}$ is less than $\frac{3}{2} = 1.5$.

This bound on the average rank can be improved via an analysis of root numbers. Indeed, for any fixed $d \neq 0$, we prove that the set of n such that the elliptic curve $E_{d,n}$ has a given root number is a countable union of acceptable sets. Moreover, a density of $1/2$ of elliptic curves $E_{d,n}$ have root number $+1$ and $1/2$ have root number -1.

Theorem 1.4. *Fix $d \neq 0$. The density of elliptic curves in the cubic twist family $E_{d,n}$ that have root number $+1$ (resp. -1) is $1/2$. The average size of the 2-Selmer group of just those elliptic curves $E_{d,n}$ having root number $+1$ (resp. -1) is 3.*

Theorems 1.1 and 1.2 are deduced from Theorem 1.4, using the p-parity theorem of Dokchitser–Dokchitser [17][3] and the p-converse theorem of Burungale–Skinner in the Appendix. See §2 for more details.

Theorem 1.4 also implies the following bounds on (the limsup and the liminf of) the average rank of elliptic curves in cubic twist families:

Theorem 1.5. *Fix $d \neq 0$, and let $\Sigma \subset \mathbb{Z}$ be any acceptable subset. The average rank in the cubic twist family of elliptic curves $E_{d,n}$ ($n \in \Sigma$) is at most $4/3$. Furthermore, if the squarefree part of d is congruent to $1 \pmod 4$, then the average rank in the cubic twist family of elliptic curves $E_{d,n}$ ($n \in \mathbb{Z}$) is at least $\frac{1}{21}2^{r-1}$, where r is the least residue of $v_2(d)/2$ modulo 3.*

Theorem 1.5 shows, for the first time, the boundedness (and, in many cases, the positivity) of the average rank in cubic twist families. The question of the boundedness of the average rank in twist families of elliptic curves has been studied extensively. The unique sextic twist family was

[2] Note that $E_{d,n}$ and $E_{-27d,n}$ are 3-isogenous, so their ranks (and 2-Selmer ranks) agree.

[3] Many important cases of the p-parity theorem were proved by Kim [25] and by Nekovář [33]; in fact, we only use the case $p = 2$ which was proved by Monsky [30].

handled by Elkies and the second and third authors [6]. The quadratic case has been studied by many authors (see, e.g., [20, 36, 42, 10, 45, 22, 26]), and most recently by Smith [44], whose work covers most quadratic twist families. Meanwhile, significant progress on the unique quartic twist family was made by Kane and Thorne [23].

1.2 Variations and related results

One may also ask which positive integers can be expressed as the sum of two *positive* rational cubes.

Theorem 1.6. *A positive proportion of positive integers are expressible as the sum of two positive rational cubes, and a positive proportion are not.*

Indeed, the same lower bounds on the proportions as in (1.1) and (1.2) hold for Theorem 1.6: if an elliptic curve over $\mathbb{Q}$ has positive rank, then the rational points are dense in the real component of the identity; thus if a non-cube positive integer n is the sum of two rational cubes, then it is also the sum of two positive rational cubes, because the elliptic curve $x^3+y^3=n$ then has positive rank and possesses an arc of real points in the positive quadrant.

Our methods also imply the following result about integers that are the product of three rational numbers in arithmetic progression:

Theorem 1.7. *A positive proportion of integers are expressible as the product of three rational numbers in arithmetic progression, and a positive proportion are not.*

Again, by the same arguments, the same lower bounds on the proportions in Theorem 1.7 hold as in (1.1) and (1.2); and the same lower bounds on the proportions hold for the set of positive integers that are the product of three *positive* rational numbers in arithmetic progression.

More generally, our results imply that a positive proportion of integers cannot be represented by any given reducible binary cubic form over $\mathbb{Q}$.

Theorem 1.8. *Let $f(x,y)$ be any binary cubic form over $\mathbb{Q}$ with a linear factor. Then, when ordered by absolute value, a positive proportion of integers cannot be expressed as $f(x,y)$ with $x,y\in\mathbb{Q}$. Furthermore, if the squarefree part of $\operatorname{Disc}(f)$ is $1 \pmod 4$, then a positive proportion of integers can be expressed as $f(x,y)$ with $x,y\in\mathbb{Q}$.*

Theorems 1.1 and 1.7 are the special cases of Theorem 1.8 where we set $f(x,y)=x^3+y^3$ and $f(x,y)=x(x+y)(x+2y)$, respectively. Theorem 1.8 follows from Theorem 1.2, since the elliptic curve $f(x,y)=n$ is isomorphic to the curve $E_{d,n}$ where $d=16\operatorname{Disc}(f)$.

When $f(x,y)$ is irreducible, the curve $f(x,y)=nz^3$ is not necessarily an elliptic curve, as it then often fails to even have local points. Indeed, in the irreducible case, the density of integers n, such that $f(x,y)=n$ has points everywhere locally, is 0. More precisely:

Theorem 1.9. *Let $f(x,y)$ be an irreducible binary cubic form over $\mathbb{Q}$. The number of integers n with $|n|<X$ such that the curve $f(x,y)=n$ has points everywhere locally is on the order of either $X/\log^{1/3}X$ or $X/\log^{2/3}X$, depending on whether $\operatorname{Disc}(f)$ is or is not a square.*

Indeed, for sufficiently large p, the curve $f(x,y)=n$ fails to have a solution over $\mathbb{Q}_p$ precisely when $v_p(n)$ is not a multiple of 3 and $f(x,y)$ is irreducible (mod p). By the Chebotarev density theorem, the density of primes p such that $f(x,y)$ is irreducible (mod p) is $\frac{2}{3}$ or $\frac{1}{3}$, depending on whether the Galois group of f is C_3 or S_3. The theorem then follows from standard counting results of Selberg–Delange type (e.g., Theorem 9.1).

In cases where it does have points locally, it is natural to ask how often the curve $f(x,y)=n$ has a global rational point. The genus one curve $f(x,y)=n$ naturally corresponds to an element of the "$\sqrt{-3}$-Selmer group" of $E_{d,n}$, where again $d=16\,\mathrm{Disc}(f)$. Indeed, the cubic twist families $E_{d,n}: y^2=x^3+dn^2$ have traditionally been studied via the Selmer groups associated to a natural isogeny $\sqrt{-3}: E_{d,n}\to E_{-27d,n}$ defined over $\mathbb{Q}$.

This approach allows one to determine the 3-Selmer group of any such curve. In 1879, Sylvester [46, §2] (see also Selmer [38]) used this $\sqrt{-3}$-descent to show that the 3-Selmer rank of $x^3+y^3=p$ for p a prime is 0 if $p\equiv 2,5 \pmod 9$ and is 1 if $p\equiv 4,7,8 \pmod 9$. This proved that primes $p\equiv 2,5 \pmod 9$ are not the sum of two cubes, and led Sylvester to conjecture that primes $p\equiv 4,7,8 \pmod 9$ are the sum of two cubes, a proof of which was recently announced by Kriz [27].

However, the $\sqrt{-3}$-Selmer group is not so useful in studying $x^3+y^3=n$ for general integers n, as the size of the $\sqrt{-3}$-Selmer tends to grow with the number of prime factors of n. For any fixed nonsquare $d\in\mathbb{Z}$, we show that the average number of such curves $f(x,y)=n$ locally having a point *grows* with $|n|$.

Theorem 1.10. *Suppose $d\in\mathbb{Z}$ is not a square. Then* $\mathrm{avg}_n \#\mathrm{Sel}_3(E_{d,n})=\infty$.

Since the average rank of $E_{d,n}$ remains bounded by Theorem 1.2 as n varies, we conclude that for most cubic twist families, the average size of $Ш(E_{d,n})[3]$ is unbounded.

Corollary 1.11. *Suppose $d\in\mathbb{Z}$ is not a square. Then* $\mathrm{avg}_m \#Ш(E_{d,n})[3]=\infty$.

In [1], we proved a similar but less extreme result for the entire twist family $E^k: y^2=x^3+k$. There we found that, for any integer r, a positive proportion of the curves E^k have $\#Ш(E^k)[3]>3^r$, but the average size of $Ш(E^k)[3]$ is still bounded.[4]

Corollary 1.11 can be reformulated in more concrete terms as follows:

Corollary 1.12. *Suppose $d\in\mathbb{Z}$ is not a square. When binary cubic forms $f(x,y)$ over $\mathbb{Z}$ of discriminant dn^2 are ordered by $|n|$, 100% of the plane curves $z^3=f(x,y)$ that are locally soluble fail the Hasse principle.*

1.3 A higher-dimensional example

We prove Theorem 1.3 in the more general context of cubic twist families of abelian varieties having a μ_3-action, where μ_3 is the group of third roots of unity; see Theorem 2.5.

As an application of this more general result, let $C: y^3=x^4+ax^2+b$ be a smooth plane quartic curve with μ_6-action. The quotient C/μ_2 is the elliptic curve $E: y^3=x^2+ax+b$. Let $A=\ker(\mathrm{Jac}(C)\to E)$ be the Prym surface associated to the double cover. For each n, consider the cubic twist $C_n: ny^3=x^4+ax^2+b$, which admits a double cover to the elliptic curve $E_n: ny^3=x^2+ax+b$. We prove:

Theorem 1.13. *Let A_n be the Prym abelian surface corresponding to the genus 3 bielliptic curve $C_n: ny^3=x^4+ax^2+b$. As $|n|\to\infty$, the average rank of $A_n(\mathbb{Q})$ is at most 3.*

The abelian surfaces A were previously studied by the third author and Weiss in [41], where it was shown that the average rank of the sextic (as opposed to cubic) twists of such an A is bounded, under the additional assumption that A admits a 3-torsion point. We remark that $\mathrm{End}_{\overline{\mathbb{Q}}}(A)$ is generically a quaternion ring over $\mathbb{Z}$ of discriminant 6, and hence A is generically simple. See also the recent work of Laga [28], who proves an average rank bound in a universal family of Prym surfaces.

[4] In the special case where d is a square, our method for proving Theorem 1.10 does not apply and we are not sure what to expect for the average size of $Ш(E_{d,n})[3]$.

2 Sketch of the proofs of Theorems 1.1–1.5

1. Parametrization of 2-Selmer elements of elliptic curves in cubic twist families

To prove Theorems 1.1–1.5, we use a parametrization of 2-Selmer elements in the family of elliptic curves $E(a_1, a_3) : y^2 + a_1xy + a_3y = x^3$ via pairs (F_1, F_2) of integral binary cubic forms, as studied by the second author and Ho [8, 9]. The curves $E = E(a_1, a_3)$ form the universal family of elliptic curves having a marked rational 3-torsion point. Elements of $\mathrm{Sel}_2(E)$ can be represented by pairs (C, D), where C is a genus 1 curve that is *locally soluble*, i.e., $C(\mathbb{Q}_p) \neq \emptyset$ for all p, $\mathrm{Pic}^1(C) \simeq E$, and D is a degree 2 divisor on C.

Let V denote the space of pairs of binary cubic forms. We say that a pair $(F_1, F_2) \in V(\mathbb{Q})$ is *locally soluble* if the genus one curve $z^2 = \mathrm{Disc}_{X,Y}(xF_1(X, Y) - yF_2(X, Y))$ has points everywhere locally. Note that the isomorphism class of this curve is invariant under the action of $\mathrm{SL}_2^2(\mathbb{Q})$ on $V(\mathbb{Q})$. There are two polynomial invariants A_1 and A_3 for the action of SL_2^2 on V, having degrees 2 and 6, respectively.

Theorem 2.1 ([9, Thm. 4.2])**.** *The elements in the 2-Selmer group* $\mathrm{Sel}_2(E)$ *of*

$$E : y^2 + a_1xy + a_3y = x^3$$

are in bijection with $\mathrm{SL}_2^2(\mathbb{Q})$*-equivalence classes of locally soluble pairs of integral binary cubic forms having invariants* $A_1 = Ma_1$ *and* $A_3 = M^3a_3$ *for some fixed nonzero integer* M*.*

Restricting Theorem 2.1 to the case where $A_1 = a_1 = 0$ yields the following corollary.

Corollary 2.2. *The elements in the 2-Selmer group* $\mathrm{Sel}_2(E_n)$ *of* $E_n : y^2 + ny = x^3$ *are in bijection with* $\mathrm{SL}_2^2(\mathbb{Q})$*-equivalence classes of locally soluble pairs of integral binary cubic forms satisfying* $A_1 = 0$ *and* $A_3 = M^3n$ *for some fixed nonzero integer* M*.*

The family E_n above is isomorphic to the family $E_{16,n} \colon y^2 = x^3 + 16n^2$ from the introduction. To handle general cubic twist families $E_{d,n}$, we first prove a generalization of Corollary 2.2 for E_n that gives some flexibility in the local conditions used to define the Selmer group inside $H^1(\mathbb{Q}, E_n)[2]$. We then use the fact that $E_{d,n} : y^2 = x^3 + dn^2$ is the d-th quadratic twist of E_{2d^2n}, and hence we have isomorphisms $E_{d,n}[2] \simeq E_{2d^2n}[2]$ and $H^1(\mathbb{Q}, E_{d,n}[2]) \simeq H^1(\mathbb{Q}, E_{2d^2n}[2])$. We prove:

Theorem 2.3. *The elements in the 2-Selmer group* $\mathrm{Sel}_2(E)$ *of* $E_{d,n} : y^2 = x^3 + dn^2$ *are in bijection with* $\mathrm{SL}_2^2(\mathbb{Q})$*-equivalence classes of integral pairs of binary cubic forms satisfying certain congruence conditions with* $A_1 = 0$ *and* $A_3 = M^3n$ *for some fixed nonzero integer* M *depending only on* d*.*

We prove this result in the more general context of cubic twist families of abelian varieties A admitting a μ_3-action (see Section 3). We assume A carries an ample line bundle L fixed by the μ_3-action and such that the corresponding polarization $\lambda \colon A \to \widehat{A}$ has kernel of size 4. For each such A, we show that there exists an elliptic curve $E = E(0, a_3)$ and an isomorphism of central extensions $\Theta(L) \simeq \Theta(\mathcal{O}_E(2\infty))$, where $\Theta(L) = \mathrm{Aut}(L/A)$ is Mumford's theta group [32, §23]. (When A is an elliptic curve, this isomorphism comes from quadratic twisting as above, but in higher dimension quadratic twisting is not helpful.) We use this isomorphism to view elements of $\mathrm{Sel}_{\lambda_n}(A_n)$ as elements of $H^1(\mathbb{Q}, E_n[2])$, which we then show correspond to orbits of pairs of integral binary cubic forms satisfying certain congruence conditions, in a manner that is analogous to the statement of Theorem 2.3.

2. The number of $\mathrm{SL}_2(\mathbb{Z})^2$-orbits on the invariant quadric $A_1 = 0$ with $|A_3| < X$

Theorem 2.3 shows that in order to determine the average size of the 2-Selmer group in a family of cubic twists $E_{d,n}$, where $d \in \mathbb{Z}$ is fixed and n varies over an acceptable set of integers, we must solve a certain counting problem. Namely, we must estimate the (weighted) number of $G(\mathbb{Z})$-orbits on $V(\mathbb{Z})$ lying on the quadric $Y = \{A_1 = 0\} \subset V$ such that $|A_3| < X$ and with A_3 satisfying certain congruence conditions. Here, we take $G := \mathrm{SL}_2^2/\mu_2$ to be the group acting faithfully on V.

In this direction, we prove the following result. We say that $(F_1, F_2) \in Y(\mathbb{Z})$ is *irreducible* if the corresponding binary quartic form $\mathrm{Disc}(xF_1 - yF_2)$ does not have a linear factor over $\mathbb{Q}$.

Theorem 2.4. *Let $\Sigma \subset \mathbb{Z}$ be an acceptable subset, let $S(\Sigma)$ denote the set of $y \in Y(\mathbb{Z}_p)$ such that $A_3(y) \in \Sigma$, and let $S_p(\Sigma)$ denote the p-adic closure of $S(\Sigma)$ in $Y(\mathbb{Z}_p)$. Let $N(S(\Sigma); X)$ denote the number of irreducible $G(\mathbb{Z})$-orbits of $y \in Y(\mathbb{Z})$ such that $A_3(y) \in \Sigma$ and $|A_3(y)| < X$. Then*

$$N(S(\Sigma); X) = X \cdot \int_{\substack{y \in G(\mathbb{Z})\backslash Y(\mathbb{R}) \\ |A_3(y)|<1}} dy \cdot \prod_p \int_{y \in S_p(\Sigma)} dy \; + \; O_\Sigma(X^{1-c}) \tag{2.1}$$

for an absolute constant $c > 0$; here, dy is the $G(\mathbb{R})$-invariant (resp. $G(\mathbb{Z}_p)$-invariant) measure on $Y(\mathbb{R})$ (resp. $Y(\mathbb{Z}_p)$) given by $dr_2\, dr_3 \cdots dr_8/(\partial A_1/\partial r_1)$, where $r_1, \ldots, r_8$ are the coordinates on V.

To prove Theorem 2.4, we use geometry-of-numbers techniques introduced by the second author in [5] to count these integer points in this unbounded but finite-volume region, along with the circle/smoothed delta method of Ramanujan–Hardy–Littlewood–Kloosterman–Duke–Friedlander–Iwaniec–Heath-Brown [20] to restrict the count to integer points on the quadric. This combination of techniques was first studied by Ruth [37], while a generally applicable method was given by the first author [2]. We take the method of [2] further in this article in order to carry out this count.

Specifically, in the "main body" of the fundamental domain the circle method is applicable and we use it to count integral zeroes of our quadratic form in an approximate cube in $\mathbb{R}^8$. Meanwhile, in the majority of the cusp of the fundamental domain, we apply a divisor bound to bound the point count by something which is sufficiently sharp (because the volume integral for the fundamental domain converges quickly). Finally, deep in the cusp, we show that the integral points on the quadric correspond to identity elements of the Selmer group, which can be counted separately.

One novel aspect of our use of the circle method is that we give a new interpretation of the singular integral and singular series arising in the circle method in terms of real and p-adic integrals with respect to a natural G-invariant measure dy.

We also prove a generalization of Theorem 2.4 where we allow weighted counts, where the weights are defined by suitable congruence conditions. The local densities are then computed for general such weight functions in terms of integrals with respect to this G-invariant measure dy, using our aforementioned interpretations for the singular integral and singular series.

These expressions for weighted counts in terms of integrals with respect to the real and p-adic G-invariant measures play a key role, e.g., in the application to average sizes of Selmer groups.

3. The average size of the 2-Selmer group of elliptic curves in cubic twist families

For applications to Selmer groups, we choose the weight function to be the characteristic function of the locally soluble orbits, weighted appropriately to account for the number of $G(\mathbb{Q})$-orbits in a given $G(\mathbb{Z})$-orbit. The corresponding local densities then take a particularly nice form, allowing us to combine Theorems 2.3 and 2.4.

To prove an exact average instead of just an upper bound, we must prove a uniformity estimate for the number of elements with $A_1 = 0$ and A_3 divisible by the square of a large prime. This can be a difficult problem in general, and in fact its analogue for the larger family $y^2 + a_1xy + a_3y = x^3$ is not known [9, §8]. We prove a suitable estimate in the case $A_1 = 0$ by using a geometric sieve for quadrics due to Browning and Heath-Brown [14] in those cases where A_3 is a multiple of p^2 for "mod p reasons"; we then use the condition $A_1 = 0$ to construct an A_3-preserving transformation that changes the condition that A_3 is a multiple of p^2 for "mod p^2 reasons" to being so for "mod p reasons", thereby reducing to the cases already handled.

Applying the uniformity estimate and a sieve then shows that the average size of the 2-Selmer group in any cubic twist family $E_{d,n}$ is 3, even when n varies within acceptable sets in $\mathbb{Z}$. More generally, we prove:

Theorem 2.5. *Let A be an abelian variety over $\mathbb{Q}$ with a degree 4 polarization $\lambda\colon A \to \widehat{A}$ induced by a symmetric line bundle $L \in \mathrm{Pic}(A)$. Suppose the pair (A, L) admits a fixed-point-free μ_3-action. For each nonzero $n \in \mathbb{Z}$, let $\lambda_n\colon A_n \to \widehat{A}_n$ be the cubic twist of λ. Let $\Sigma \subset \mathbb{Z}$ be any acceptable set. Then, as $|n| \to \infty$, the average size of $\#\mathrm{Sel}_{\lambda_n}(A_n)$ over $n \in \Sigma$ is 3.*

As a special case, we deduce Theorem 1.3.

4. Analysis of root numbers

Theorem 1.3 alone is not sufficient to deduce Theorems 1.1 and 1.2. To proceed further, we carry out an analysis of the root numbers $w_{d,n} \in \{\pm 1\}$ of the elliptic curves $E_{d,n}$. Recall that $w_{d,n}$ is the sign appearing in the functional equation for the L-function of $E_{d,n}$, and the parity conjecture predicts that $(-1)^{\mathrm{rk}\, E_{d,n}(\mathbb{Q})} = w_{d,n}$. We prove:

Theorem 2.6. *Fix d, and let Σ be any acceptable subset of $\mathbb{Z}$ defined by prime-to-3 congruence conditions. Then the root number $w_{d,n}$ is equidistributed as $n \in \Sigma$ goes to infinity. In other words,*

$$\sum_{n \in \Sigma : |n| \le X} w_{d,n} = o_{d,S}(X).$$

We will in fact prove Theorem 2.6 for more general Σ, namely those Σ defined by prime-to-3 congruence conditions which have a natural density. Moreover, for Σ defined by finitely many congruence conditions mod m we will obtain a bound of shape $\ll d^{O(1)} m^{O(1)} X^{1-c}$ for an explicit absolute constant $c \in \mathbb{R}^+$.

From this equidistribution result, we see that the parity conjecture would imply that 50% of twists $E_{d,n}$ have odd rank and 50% of twists $E_{d,n}$ have even rank (with d fixed and n varying), even if we impose certain congruence conditions on n. However, the prime-to-3 hypothesis in the theorem is crucial, since one can construct acceptable subsets Σ such that the root number $w_{d,n}$ is constant for $n \in \Sigma$. Indeed, we prove:

Theorem 2.7. *Fix d and let Σ be an acceptable subset of $\mathbb{Z}$. The set $\Sigma_+ \subset \Sigma$ (resp., Σ_-) of $n \in \Sigma$ such that $E_{d,n}$ has root number $+1$ (resp., -1) is a countable union of acceptable sets.*

These theorems are proved using an explicit formula for the root numbers $w_{d,n}$ due to Varilly-Alvarado (based on work of Rohrlich) [47]. Aside from multiplicative factors at 2 and 3, the root number of $y^2 = x^3 - 432n^2$ is equal to $(-1)^{\omega'(n)}$ where $\omega'(n)$ is the number of primes $p \equiv 2 \pmod 3$ dividing n. Theorem 2.6 follows from suitable asymptotic control of $\sum_n (-1)^{\omega'(n)}$ over arithmetic progressions.

Theorem 1.4 follows from combining Theorem 1.3 with Theorems 2.6 and 2.7.

5. The proportion of curves with rank 0 and rank 1 in cubic twist families

The parity conjecture is still open, so our equidistribution results for root numbers do not immediately imply anything about ranks of elliptic curves. However, we may instead apply the p-parity theorem of Dokchitser–Dokchitser [17].

Theorem 2.8 (p-parity)**.** *Let $E/\mathbb{Q}$ be an elliptic curve and let $w(E)$ be its root number. Then for every prime p, we have*

$$w(E) = (-1)^{\dim_{\mathbb{F}_p} \mathrm{Sel}_p(E) + \dim_{\mathbb{F}_p} E[p](\mathbb{Q})}.$$

It is easy to see that in any cubic twist family, we have $E_{d,n}[2](\mathbb{Q}) = 0$ for 100% of integers n. For such n, the case $p = 2$ of the p-parity theorem reads $w_{d,n} = (-1)^{\dim_{\mathbb{F}_2} \mathrm{Sel}_2(E_{d,n})}$. Across the root number $+1$ curves, the 2-Selmer group size will be an even power of 2 (i.e., 1, 4, 16, etc.), while for the root number -1 curves, the 2-Selmer group size will be an odd power of 2 (i.e., 2, 8, 32, etc.). Since the average size of $\mathrm{Sel}_2(E_{d,n})$ is 3 across those $E_{d,n}$ having root number $+1$, we conclude that at least $1/3$ of these curves must have 2-Selmer rank 0 (note that 3 is the average of $1, 4, 4$). Since $\mathrm{Sel}_2(E) = 0$ implies $\mathrm{rk}\, E(\mathbb{Q}) = 0$, we deduce that at least $1/6$ of elliptic curves in every cubic twist family have rank 0.

Similarly, since the average size of $\mathrm{Sel}_2(E_{d,n})$ is 3 across those $E_{d,n}$ having root number -1, we conclude that at least $5/6$ of these curves must have 2-Selmer rank 1 (note that 3 is the average of $2, 2, 2, 2, 2, 8$). We therefore deduce:

Corollary 2.9. *Fix $d \neq 0$. As n varies ordered by absolute value, at least $5/12$ of the elliptic curves $E_{d,n}$ have 2-Selmer rank 1.*

Moreover, we prove a similar result as n varies over sets of integers defined by congruence conditions that are prime to 3. We then apply the following p-converse theorem of Burungale and Skinner (Corollary A.4 of the Appendix).

Theorem 2.10. *Let p be prime and let $E/\mathbb{Q}$ be a CM elliptic curve with supersingular reduction at p. If $\#\mathrm{Sel}_p(E) = p$ and the localization map $\mathrm{Sel}_p(E) \to E(\mathbb{Q}_p)/pE(\mathbb{Q}_p)$ is injective, then $\mathrm{rk}\, E(\mathbb{Q}) = 1$.*

This allows us to prove that a positive proportion of cubic twists $E_{d,n}$ with good reduction at 2 have rank 1. We handle the extra 2-adic condition by proving an appropriate equidistribution theorem for 2-Selmer elements under the 2-adic localization map (Theorem 5.4), though this causes our lower bound on the proportion of rank 1 twists to drop from $5/12$ to $1/6$.

This completes the sketch of the proofs of Theorems 1.1, 1.2, and 1.5. If the theorem of Burungale–Skinner is eventually generalized to cover any CM elliptic curve with *potentially* supersingular reduction satisfying $\#\mathrm{Sel}_p(E) = p$ (so that bad reduction is allowed and without any hypothesis on the localization map), then our results would imply that at least $5/12$ of cubic twists have rank 1 in any cubic twist family of elliptic curves, and thus that at least $5/12$ of integers are the sum of two rational cubes.

3 Pairs of binary cubic forms and 2-Selmer elements in cubic twist families

In [8, §6.3.2], the second author and Ho gave a functorial bijection between orbits of pairs of binary cubic forms and isomorphism classes of genus 1 curves with extra data. We review this parametrization and then specialize it to certain genus one curves with j-invariant 0. We then

generalize this specialization to arbitrary cubic twist families of polarized abelian varieties. We use this parametrization to classify certain 2-Selmer elements in terms of orbits of pairs of cubic forms and show that orbits corresponding to 2-Selmer elements have integral representatives.

3.1 Parametrizing degree two genus one curves via pairs of binary cubics

Let F be a ground field of characteristic neither 2 nor 3. We consider the two groups $\widetilde{G} = \mathrm{GL}_2^2$ and $G = \mathrm{SL}_2^2/\mu_2$, where we view $\mu_2 \subset \mathrm{SL}_2^2$ via the diagonal scalar embedding. Both groups act on $V = F^2 \otimes \mathrm{Sym}^3 F^2$, with the first copy of GL_2 (resp. SL_2) acting on F^2 by the standard representation and the second copy acting by the symmetric cube of the standard representation. It is known that the ring of polynomial invariants for the action of G on V is freely generated by two invariants A_1 and A_3, of degrees 2 and 6 respectively [8]. These are only *relative* invariants for the action of $\widetilde{G}$. Let $Y \subset V$ be the quadric hypersurface defined by the equation $A_1 = 0$.

For each $n \in F^\times$, let $Y(F)_n$ be the subset of $y \in Y(F)$ having A_3-invariant n. The group $G(F)$ acts on both $Y(F)$ and $Y(F)_n$. We will make use of the following bijection due to the second author and Ho [8, Thm. 2.3], which relates the $G(F)$-orbits on $Y(F)_n$ to isomorphism classes of pairs (C, L), where C is a genus one curve whose Jacobian $\mathrm{Pic}^0(C)$ is the elliptic curve $E_n \colon y^2 + ny = x^3$, and where L is a degree 2 line bundle on C. Such a genus one curve C is automatically an E_n-torsor [43, §X].

Theorem 3.1. *Fix $n \in F^\times$. There is a natural bijection between $G(F)$-orbits on $Y(F)_n$ and isomorphism classes of pairs (C, L), where C is an E_n-torsor and L is a degree two line bundle on C. If $v \in Y(F)_n$ corresponds to (C, L), then $\mathrm{Stab}_{\widetilde{G}}(v) \simeq \mathrm{Aut}(C, L)$ and $\mathrm{Stab}_G(v) \simeq E_n[2]$.*

Proof. This follows from the more general result cited earlier (see also [9, Thm. 3.1]). Given a pair of cubic forms $v = (F_1, F_2) \in V(F)$, the corresponding curve C is the hyperelliptic curve $z^2 = \mathrm{Disc}_{X,Y}(xF_1(X,Y) - yF_2(X,Y))$,[5] and the line bundle L is the pullback of $\mathcal{O}_{\mathbb{P}^1}(1)$ along the map $z \colon C \to \mathbb{P}^1$. The Jacobian of C is the elliptic curve $E \colon y^2 + A_1(v)xy + A_3(v)y = x^3$. Setting $A_1(v) = 0$ and $A_3(v) = n$, we obtain an $y^2 + ny = x^3$. □

For each $n \in F^\times$, let $v_n \in Y(F)_n$ be a pair of cubic forms corresponding under the bijection of Theorem 3.1 to the pair (E_n, L_n), where $L_n = \mathcal{O}_{E_n}(2\infty)$ is the line bundle corresponding to the divisor $2\infty = 2[0 \colon 1 \colon 0]$. We refer to the $G(F)$-orbit of v_n as the reducible orbit having A_3-invariant n. Explicitly, we may take v_n to be the pair $(xy^2, x^3 + ny^3)$.

The stabilizer $\mathrm{Stab}_{\widetilde{G}}(v_n)$ is isomorphic to Mumford's theta group $\Theta(L_n) := \mathrm{Aut}(E_n, L_n)$ of automorphisms of the total space of L_n lifting those of E_n. This is an infinite noncommutative group scheme, best described via the exact sequence

$$0 \to \mathbb{G}_m \to \Theta(L_n) \to E_n[2] \to 0. \tag{3.1}$$

The subgroup $\mathbb{G}_m$ corresponds to the automorphisms of L_n given by scalar multiplication in each fiber, and the map $\Theta(L_n) \to E_n[2]$ records the underlying automorphism of E_n, which is necessarily translation by a point in $E_n[2]$.

Applying the long exact sequence in cohomology to (3.1) gives the *obstruction map*

$$\mathrm{ob} \colon H^1(F, E_n[2]) \to H^2(F, \mathbb{G}_m).$$

This is only a map of pointed sets, despite the fact that both source and target are abelian groups. By Hilbert's Theorem 90, we have $H^1(F, \mathbb{G}_m) = 0$, and hence the kernel of the obstruction map is $H^1(F, \Theta(L_n))$, viewed as a subset of $H^1(F, E_n[2])$.

[5] Here, $\mathrm{Disc}_{x,y}$ is -27 times the usual polynomial discriminant, following the conventions used in [8, 9].

Corollary 3.2. *There is a functorial bijection between the $G(F)$-orbits on $Y(F)_n$ and the pointed set $H^1(F, \Theta(L_n))$. Under this bijection, the reducible orbit corresponds to the trivial class.*

Proof. By arithmetic invariant theory [7, Prop. 1], the set of $G(F)$-orbits on $Y(F)_n$ is in bijection with the kernel of $H^1(F, \mathrm{Stab}_G(v_n)) \to H^1(F, \mathrm{SL}_2^2/\mu_2)$. Since $H^1(F, \mathrm{SL}_2^2) = 0$, we have $H^1(F, \mathrm{SL}_2^2/\mu_2) \simeq H^2(F, \mu_2)$. Thus the $G(F)$-orbits on $Y(F)_n$ are in bijection with the kernel of $H^1(F, E_n[2]) \to H^2(F, \mu_2)$, through which the obstruction map factors. Alternatively, we may apply the theory to $\widetilde{G}$ and use that $\Theta(L_n) \simeq \mathrm{Stab}_{\widetilde{G}}(v_n)$ and $H^2(F, \widetilde{G}) = 0$. □

3.2 Parametrizing $A[\lambda]$-torsors via pairs of binary cubics

We next generalize Theorem 3.1 to more general cubic twist families of elliptic curves. Since it is not any harder, we will allow abelian varieties of higher dimension as well. We will make explicit the case of elliptic curves, for the benefit of readers who are not comfortable with the language of abelian varieties, but also to emphasize that even in this case we are doing something new. Indeed, the particular cubic twist family $E_n \colon y^2 + ny = x^3$ is special in the sense that it admits the rational point $(0, 0)$ of order 3 (which is crucial for the parametrization in [8, §6.3.2]), whereas in a general cubic twist family of elliptic curves $E_{d,n} \colon y^2 = x^3 + dn^2$, with d a fixed non-square, the curve $E_{d,n}$ has no rational 3-torsion point. An easy computation shows that E_n is isomorphic to $E_{16,n}$.

In this paper, a *polarized abelian variety over F* is a pair (A, L), where A is an abelian variety over F and L is an ample line bundle on A.

Remark 3.3. Every abelian variety A admits a dual abelian variety $\widehat{A} = \mathrm{Pic}^0(A)$ which parametrizes algebraically trivial line bundles on A. If $A = E$ is an elliptic curve, then algebraically trivial is equivalent to degree 0, and there is an isomorphism $E \simeq \widehat{E} = \mathrm{Pic}^0(E)$ defined by $P \mapsto \mathcal{O}_E(P - \infty)$. In higher dimension A and $\widehat{A}$ need not be isomorphic, but any ample line bundle L on A induces a map $\lambda_L \colon A \to \widehat{A}$, sending a point P to $t_P^* L \otimes L^{-1}$, where $t_P \colon A \to A$ is translation by P. The map λ_L is called the *polarization* associated to (A, L).

If (A, L) is a polarized abelian variety, we write $\mathrm{Aut}(A, L)$ for the group of automorphisms $\alpha \in \mathrm{Aut}(A)$ such that $\alpha^* L \simeq L$. Let μ_3 be the group of third roots of unity. A *μ_3-action* on (A, L) is an inclusion of F-group schemes $\iota \colon \mu_3 \hookrightarrow \mathrm{Aut}(A, L)$. The action has *isolated fixed points* if for each nontrivial $\zeta \in \mu_3$, the endomorphism $1 - \iota(\zeta)$ has finite kernel, or in other words, is an isogeny. This condition is automatically satisfied if A is simple, e.g., if A is an elliptic curve.

Example 3.4. If $A = E$ is an elliptic curve with μ_3-action, then E has a model $E \colon y^2 = x^3 + d$ for some $d \in F^\times$, and the μ_3-action is given by $(x, y) \mapsto (\zeta x, y)$. The line bundles $\mathcal{O}_E(k\infty)$ are preserved by this action since ∞ is sent to ∞. The kernel of $1 - \iota(\zeta)$ is the order three group generated by $(0, \pm\sqrt{d})$.

We now suppose that (A, L) admits a μ_3-action ι. Then for each $n \in F^\times$, we define (A_n, L_n) to be the twist of (A, L) by the cocycle $G_F \to \mathrm{Aut}(A, L)$ sending $g \mapsto \iota(g(\sqrt[3]{n})/\sqrt[3]{n})$. The isomorphism class of (A_n, L_n) depends only on the image of n in $F^\times/F^{\times 3} \simeq H^1(F, \mu_3)$. We have $\dim_F H^0(A, L) = \dim_F H^0(A_n, L_n)$ for all n, and the polarizations $\lambda_n \colon A_n \to \widehat{A}_n$ all have degree $(\dim_F H^0(A, L))^2$.

We assume for the rest of the section that (A, L) has a μ_3-action with isolated fixed points and that $\dim_F H^0(A, L) = 2$. We also assume that L is symmetric, in the sense that $[-1]^* L \simeq L$. The groups $A_n[\lambda_n](\bar{F})$ are then abstractly isomorphic to $(\mathbb{Z}/2\mathbb{Z})^2$, so that $\lambda_n \colon A_n \to \widehat{A}_n$ is a family of $(2, 2)$-isogenies. Let $\lambda = \lambda_1 \colon A \to \widehat{A}$ be the initial polarization.

Example 3.5. If $A = E\colon y^2 = x^3 + d$ is an elliptic curve, we may take $L = \mathcal{O}_E(2\infty)$. After making the identification $E \simeq \widehat{E}$, the polarization $\lambda\colon E \to \widehat{E}$ becomes the multiplication-by-minus-two map $[-2]\colon E \to E$. Indeed, λ sends P to the divisor $2(-P) - 2\infty \sim 2(\infty - P)$. The twist E_n is the elliptic curve $E_{d,n}\colon y^2 = x^3 + dn^2$ from the introduction.

Lemma 3.6. *There exists $d \in F^\times$ such that $A[\lambda] \simeq \mathrm{Stab}_G(y_d)$.*

Proof. The μ_3-action on (A, L) induces a μ_3-action on $A[\lambda]$. The hypothesis that $1-\zeta$ is an isogeny implies that the action of μ_3 on $A[\lambda] \setminus 0$ is simply transitive. Indeed, the degree of $1 - \zeta$ is a power of 3, so $1 - \zeta$ cannot annihilate any point of order 2. We also have $\zeta^g P^g = (\zeta P)^g$ for all $g \in \mathrm{Gal}(\bar{F}/F)$, $\zeta \in \mu_3(\bar{F})$, and $P \in A[\lambda](\bar{F})$. From this we see that $F(\mu_3)$ is contained in $F(A[\lambda])$, which is a Galois extension of F whose Galois group is a subgroup of $S_3 \simeq \mathrm{GL}_2(\mathbb{F}_2)$. It follows from elementary Galois theory (see [12, Lem. 33]) that $F(A[\lambda]) \simeq F(\sqrt[3]{d}, \zeta_3) = F(\sqrt[3]{d^2}, \zeta_3)$ for some $d \in F^\times$. We conclude that $A[\lambda] \simeq E_{2d}[2]$, since $\mathbb{Q}(E_{2d}[2])$ is the splitting field of the cubic polynomial $x^3 + 64d^2$, and $E_{2d}[2]$ is the unique twist of $(\mathbb{Z}/2\mathbb{Z})^2$ with that splitting field.[6] The lemma now follows from Theorem 3.1. □

Example 3.7. If $A\colon y^2 = x^3 + d$ is an elliptic curve then $A[2] \simeq E_{d'}[2] \simeq \mathrm{Stab}_G(y_{d'})$, where $d' = 2d$, so the notation $E_{d,n}$ from the introduction does not quite match the notation for the parameter d whose existence is guaranteed by Lemma 3.6. The choice $d' = 4d^2$ also works. For our purposes, the exact choice of d satisfying Lemma 3.6 will not matter.

We fix once and for all an element $d \in F^\times$ as in Lemma 3.6. Just as in the previous subsection there is an exact sequence defining the theta group $\Theta(L) = \mathrm{Aut}(L)$:

$$0 \to \mathbb{G}_m \to \Theta(L) \to A[\lambda] \to 0.$$

We will sometimes think of points of $\Theta(L)$ as pairs (P, φ), with $P \in A[\lambda]$ and φ an isomorphism $t_P^* L \simeq L$, where $t_P\colon A \to A$ is translation-by-P. If $P = 0$, then φ may be viewed as a scalar in $F^\times$. Taking Galois cohomology, we obtain a map of pointed sets $H^1(F, A[\lambda]) \to H^2(F, \mathbb{G}_m)$. As before, the kernel of this map is $H^1(F, \Theta(L))$.

We can now state our general parametrization result.

Theorem 3.8. *Let $d \in F^\times$ be defined as in Lemma 3.6, and for each $n \in F^\times$, let (A_n, L_n) be the n-th cubic twist of (A, L). Then there is a natural bijection between the $G(F)$-orbits on $Y(F)_{dn}$ and the pointed set $H^1(F, \Theta(L_n))$, sending the reducible orbit v_{dn} to the identity element. For any $v \in Y(F)_{dn}$, the stabilizer $\mathrm{Stab}_G(v)$ is isomorphic to $A_n[\lambda_n]$.*

Proof. Let $E = E_d$ and $L_E = \mathcal{O}_E(2\infty)$. We will show that there is an isomorphism of central extensions $\Theta(L_1) \simeq \Theta(L_E)$. The case $n = 1$ of the Theorem then follows from Corollary 3.2. The general case then follows too, since taking cubic twists we obtain $\Theta(L_n) \simeq \Theta((L_E)_n)$, for all n.

Let $\mathscr{M}$ be the line bundle $p_1^* L \otimes p_2^* L_E$ which gives rise to the product polarization on $A \times E$. The theta group of $\mathscr{M}$ is related to the theta groups of L and L_E in a simple way:

$$\Theta(\mathscr{M}) \simeq (\Theta(L) \times \Theta(L_E))/\Delta,$$

where $\Delta = \{(1, t, 1, t^{-1})\} \subset \mathbb{G}_m \times \mathbb{G}_m \subset \Theta(L) \times \Theta(L_E)$. Thus there is a short exact sequence

$$1 \longrightarrow \mathbb{G}_m \longrightarrow \Theta(\mathscr{M}) \xrightarrow{p} A[\lambda] \times E[2] \longrightarrow 1.$$

[6] Recall that E_n is isomorphic to $y^2 = x^3 + 16n^2$ so E_{2n} is isomorphic to $y^2 = x^3 + 64n^2$.

Now choose an isomorphism $\eta\colon A[\lambda] \simeq E[2]$ and let B be the abelian variety which is the quotient of $A \times E$ by the graph $\Gamma_\eta \subset A[\lambda] \times E[2] \subset A \times E$. Let $\pi\colon A \times E \to B$ be the quotient map. The subgroup $\Gamma_\eta \subset (A \times E)[\lambda_{\mathscr{M}}] = A[\lambda] \times E[2]$ is isotropic with respect to the skew-symmetric Weil pairing induced by $\mathscr{M}$, since

$$\langle (P, \eta(P)), (Q, \eta(Q)) \rangle_{\mathscr{M}} = \langle P, Q \rangle_L \langle \eta(P), \eta(Q) \rangle_{L_E} = \langle P, Q \rangle_L^2 = 1.$$

Here we have used that η is (automatically!) a symplectic isomorphism with respect to the μ_2-valued Weil pairings induced by L and L_E.

Lemma 3.9. *There is a line bundle L_B on B such that $\pi^* L_B \simeq \mathscr{M}$.*

Proof. Define the group scheme $\widehat{\Gamma}_\eta = \ker(\mathrm{Pic}(B) \xrightarrow{\pi^*} \mathrm{Pic}(A \times E))$, abstractly isomorphic to the self-dual group scheme $E[2] \simeq \Gamma_\eta = \ker(\pi)$. Since Γ_η is isotropic, there is a line bundle $\widetilde{L}$ on $B_{\overline{F}}$ such that $\pi^* \widetilde{L} \simeq \mathscr{M}_{\overline{F}}$ [32, §23]. We show that we can choose $\widetilde{L}$ such that it descends to B. Since B has a rational point, this is equivalent to $\widetilde{L}^\sigma \simeq \widetilde{L}$ for all $\sigma \in G_F$. The collection of all such line bundles $\widetilde{L}$ is a $\widehat{\Gamma}_\eta$-torsor over $\mathbb{Q}$, hence gives a class $c \in H^1(F, E[2])$. Moreover, $\widetilde{L}$ descends to B if and only if this torsor is trivial. On the other hand, the μ_3-action on $A \times E$ descends to a μ_3-action on B which interwines the isogeny π. Since L and L_E are fixed by μ_3, we have $\zeta^* \widetilde{L} \otimes \widetilde{L}^{-1} \in \widehat{\Gamma}_\eta$ and hence $c = \zeta(c)$ in $H^1(F(\zeta_3), E[2])$. Let $K = F(\zeta_3, E[2])$ be the cubic étale-algebra over $F(\zeta_3)$ cut out by $E[2] \setminus \{0\}$.[7] Under the isomorphism ([7, Prop. 5.1])

$$H^1(F(\zeta_3), E[2]) \simeq \ker\left(K^\times / K^{\times 2} \xrightarrow{\mathrm{Nm}} F(\zeta_3)^\times / F(\zeta_3)^{\times 2} \right),$$

the action of μ_3 on $H^1(F(\zeta_3), E[2])$ is identified with the action of $\mathrm{Gal}(K/F(\zeta_3))$ on the elements of $K^\times/K^{\times 2}$ of square norm. Since the latter group action is easily seen to have no nontrivial fixed points, it follows that c is trivial and hence $\widetilde{L}$ descends to B. □

Remark 3.10. We remark for later use that Riemann-Roch and the formula $\chi(\mathscr{M}) = \deg(\pi)\chi(L_B)$ together imply that L_B is a principal polarization.

The existence of this line bundle L_B implies, by [32, Thm. 2 §23], that there is a subgroup $H \subset \Theta(\mathscr{M})$ and an isomorphism $\psi\colon \Gamma_\eta \simeq H$ such that $p \circ \psi = \mathrm{id}$. This data determines an isomorphism of theta groups

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{G}_m & \longrightarrow & \Theta(L) & \longrightarrow & A[\lambda] & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \mathrm{id}} & & \downarrow{\scriptstyle \widetilde{\eta}} & & \downarrow{\scriptstyle \eta} & & \\ 0 & \longrightarrow & \mathbb{G}_m & \longrightarrow & \Theta(L_E) & \longrightarrow & E[2] & \longrightarrow & 0. \end{array}$$

Explicitly, on S-valued points, if $\psi(P, \eta(P)) = (P, s_0, \eta(P), r_0) \in H \subset \Theta(\mathscr{M})$, then

$$\widetilde{\eta}(P, s) = (\eta(P), (s_0^{-1} s) r_0),$$

where we view $s_0^{-1} s$ as a scalar in $\mathrm{Aut}(L) \simeq \mathbb{G}_m \simeq \mathrm{Aut}(L_E)$. This proves the claimed isomorphism of central extensions $\Theta(L) \simeq \Theta(L_E)$. □

[7] This is either a $\mathbb{Z}/3\mathbb{Z}$-extension or $F(\zeta_3)^3$. The map $H^1(F, E[2]) \to H^1(F(\zeta_3), E[2])$ is injective in both cases.

Remark 3.11. The bijections of Theorem 3.8 (one for each n) seem to depend on the initial choice of isomorphism $\Theta(L) \simeq \Theta(L_E)$, which itself depends on a choice of isomorphism $A[\lambda] \simeq E_d[2]$. In fact, Jef Laga has pointed out to us that any automorphism of $\Theta(L)$ (as central extensions) that commutes with the μ_3-action and which induces the identity on $A[\lambda]$ is the identity. This uniqueness gives another way to prove the existence of an isomorphism $\Theta(L) \simeq \Theta(L_E)$ over F.

Example 3.12. If $A\colon y^2 = x^3 + d$ is an elliptic curve with $L = \mathcal{O}_A(2\infty)$, then we can make the bijection $H^1(F, \Theta(L)) \simeq H^1(F, \Theta(L_E))$ in the proof of Theorem 3.8 very explicit. Recall that in this case we may take $E = E_{2d^2}\colon y^2 = x^3 + d^4$. Then $H^1(F, \Theta(L_E))$ parametrizes orbits of pairs of binary cubic forms $y = (F_1, F_2)$ with $A_1(y) = 0$ and $A_3(y) = 2d^2$. On the other hand, $H^1(F, \Theta(L))$ parametrizes isomorphism classes of curves of the form $z^2 = f(x, y)$ with Jacobian A. The explicit map between these two sets sends (F_1, F_2) to the curve $dz^2 = \mathrm{Disc}(xF_1 - yF_2)$.

Finally, we make an explicit connection between rational points on $A_n(F)$ and $G(F)$-orbits on $Y(F)_{dn}$. In fact, the more direct connection is with $\widehat{A}_n(F)$ not $A_n(F)$, since the short exact sequence

$$0 \to A_n[\lambda_n] \to A_n \to \widehat{A}_n \to 0$$

induces a map $\delta\colon \widehat{A}_n(F) \to H^1(F, A_n[\lambda_n])$.

Proposition 3.13. *The composition $\widehat{A}_n(F) \xrightarrow{\delta} H^1(F, A_n[\lambda]) \xrightarrow{\mathrm{ob}} H^2(F, \mathbb{G}_m)$ is 0. In particular, the map δ factors through a map $\widehat{A}_n(F) \to H^1(F, \Theta(L_n))$.*

Proof. This is [35, Prop. 6.4] and is where we use the fact that L is symmetric. □

Thus, to each point $P \in \widehat{A}_n(F)$, there is an associated $G(F)$-orbit of pairs of binary cubic forms $v \in Y(F)_{dn}$.

3.3 Parametrization over local fields

We specialize the preceding discussion to the case $F = \mathbb{Q}_p$ for some prime number p or $F = \mathbb{R} = \mathbb{Q}_\infty$. If $p < \infty$, we assume, without loss of generality, that the fixed element $d \in F^\times$ lies in $\mathbb{Z}_p$.

For each $n \in \mathbb{Q}_p^\times$, we say that $v \in Y(\mathbb{Q}_p)_{nd}$ is *locally soluble* if the corresponding element of $H^1(\mathbb{Q}_p, \Theta(L_n))$, via Theorem 3.8, lies in the image of the Kummer map $\widehat{A}_n(\mathbb{Q}_p) \to H^1(\mathbb{Q}_p, \Theta(L_n))$ given by Proposition 3.13. Write $Y(\mathbb{Q}_p)^{\mathrm{ls}}_{dn}$ for the set of locally soluble $v \in Y(\mathbb{Q}_p)$ having A_3-invariant dn. This notion of local solubility of course depends on the pair (A, L).

Example 3.14. If $A\colon y^2 = x^3 + d$ is an elliptic curve, then $v = (F_1, F_2) \in Y(\mathbb{Q}_p)$ is locally soluble if and only if the curve $C\colon dz^2 = \mathrm{Disc}(xF_1 - yF_2)$ has $C(\mathbb{Q}_p) \neq \emptyset$.

For $p < \infty$, we need the following integrality result, which is due to the second author and Ho in the case where A is an elliptic curve E with a 3-torsion point and $L = \mathcal{O}_E(2\infty)$ [9, Thm. 4.2]. Write $\mathfrak{f}_A$ for the conductor of A, so that $p \mid \mathfrak{f}_A$ if and only if A has bad reduction. If $A = E$ is an elliptic curve with minimal Weierstrass model, then $p \mid \mathfrak{f}_E$ if and only if p divides the discriminant of a minimal Weierstrass model of E.

Theorem 3.15. *Assume $p \nmid 6d\mathfrak{f}_A\infty$ and $n \in \mathbb{Z}_p$. If $v \in Y(\mathbb{Q}_p)^{\mathrm{ls}}_{dn}$, then $Y(\mathbb{Z}_p) \cap G(\mathbb{Q}_p)y \neq \emptyset$. In other words, the $G(\mathbb{Q}_p)$-orbit of y contains an integral representative.*

The proof uses the following lemma.

Lemma 3.16. *Suppose $p > 3$ and $m \in p\mathbb{Z}_p$ has valuation 1 or 2. Then $H^1(\mathbb{Q}_p, E_m[2]) = 0$.*

Proof. The cubic field $L = \mathbb{Q}_p(\sqrt[3]{m^2})$ is totally ramified at p. Since, by [7, Prop. 5.1],

$$E_m[2] \simeq \ker\left(\mathrm{Res}^L_{\mathbb{Q}_p}\mu_2 \xrightarrow{\mathrm{Nm}} \mu_2\right),$$

we compute

$$\begin{aligned} H^1(\mathbb{Q}_p, E_m[2]) &\simeq \ker(L^\times/L^{\times 2} \xrightarrow{\mathrm{Nm}} \mathbb{Q}_p^\times/\mathbb{Q}_p^{\times 2}) \\ &\simeq \ker(\mathcal{O}_L^\times/\mathcal{O}_L^{\times 2} \xrightarrow{\mathrm{Nm}} \mathbb{Z}_p^\times/\mathbb{Z}_p^{\times 2}) \\ &\simeq \ker(\mathbb{F}_p^\times/\mathbb{F}_p^{\times 2} \xrightarrow{x\mapsto x^3} \mathbb{F}_p^\times/\mathbb{F}_p^{\times 2}) = 0. \end{aligned}$$

□

Proof of Theorem 3.15. Recall that $A[\lambda] \simeq E_d[2]$. Twisting both sides, we have $A_n[\lambda_n] \simeq E_{dn}[2]$, and hence $H^1(\mathbb{Q}_p, A_n[\lambda_n]) \simeq H^1(\mathbb{Q}_p, E_{dn}[2])$. If $v_p(n) \not\equiv 0 \pmod 3$, then

$$H^1(\mathbb{Q}_p, \Theta(L_n)) \subset H^1(\mathbb{Q}_p, A_n[\lambda_n]) \simeq H^1(\mathbb{Q}_p, E_{dn}) = 0$$

by Lemma 3.16. It follows that there is a unique $G(\mathbb{Q}_p)$-orbit in $Y(\mathbb{Q}_p)_{nd}$, namely the reducible orbit. This orbit has an explicit integral representative, namely, the pair of binary cubic forms $(x^3 + ndy^3, xy^2)$; see [9, §4.6(b)].

If $v_p(n) \equiv 0 \pmod 3$, then because $p \nmid 6d\mathfrak{f}_A\mathrm{Disc}(\mathbb{Q}_p(\sqrt[3]{n}))$, the twist A_n has good reduction at p. The image of the Kummer map $\widehat{A}_n \to H^1(\mathbb{Q}_p, A_n[\lambda_n])$ is therefore the subgroup of unramified classes [15, Prop. 2.7(d)]. By the same reasoning, this is also the image of the Kummer map $E_{dn}(\mathbb{Q}_p) \to H^1(\mathbb{Q}_p, E_{dn}[2]) \simeq H^1(\mathbb{Q}_p, A_n[\lambda_n])$; in this case, the result follows from [9, Theorem 4.2]. □

Finally, we record a near-converse to Theorem 3.15.

Proposition 3.17. *If $p \nmid 6d\mathfrak{f}_A\infty$ and $n \in \mathbb{Z}_p$ has valuation $v_p(n) \le 2$, then $Y(\mathbb{Z}_p)_{dn} \subset Y(\mathbb{Q}_p)^{\mathrm{ls}}_{dn}$.*

Proof. If $1 \le v_p(n) \le 2$, then in the proof of Theorem 3.15 we saw that all $v \in Y(\mathbb{Q}_p)$ having A_3-invariant dn are locally soluble. So it remains to show that if $v_p(n) = 0$ and $v \in Y(\mathbb{Z}_p)_{dn}$, then the class $v' \in H^1(\mathbb{Q}_p, E_{dn}[2]) \simeq H^1(\mathbb{Q}_p, A_n[\lambda_n])$ corresponding to v is in the image of the Kummer map (recall that in this case, the image of the Kummer map is the subgroup of unramified classes, for both A_n and E_{dn}). Now, the pair of binary cubic forms $v = (F_1, F_2)$ determines an explicit genus one curve $C/\mathbb{Q}_p$ with an integral model $\mathcal{C}\colon z^2 = \mathrm{Disc}(xF_1 - yF_2)$. The special fiber $\mathcal{C}_p$ is a hyperelliptic curve over $\mathbb{F}_p$ which comes from a pair of binary cubic forms over $\mathbb{F}_p$ and has discriminant which is nonzero in $\mathbb{F}_p$. It follows that $\mathcal{C}_p$ is smooth and hence $\mathcal{C}/\mathbb{Z}_p$ is smooth. Since $C/\mathbb{Q}_p$ has good reduction, it has a $\mathbb{Q}_p$-point, and hence $C \simeq E_{dn}$. Equivalently, the class v' lies in $\ker(H^1(\mathbb{Q}_p, E_{dn}[2]) \to H^1(\mathbb{Q}_p, E_{dn}))$. Since this kernel is also the image of the Kummer map, we have proven the desired claim. □

3.4 Parametrization over global fields

Now let us specialize the situation of Section 3.2 to the case $F = \mathbb{Q}$. Recall that the Selmer group $\mathrm{Sel}_{\lambda_n}(A_n)$ is the kernel of the natural map

$$H^1(\mathbb{Q}, A_n[\lambda_n]) \longrightarrow \prod_{p\le\infty} H^1(\mathbb{Q}_p, A_n)[\lambda_n].$$

Equivalently, the Selmer group $\mathrm{Sel}_{\lambda_n}(A_n)$ consists of those $\alpha \in H^1(\mathbb{Q}, A_n[\lambda_n])$ whose restrictions $\mathrm{res}_p(\alpha)$ lie in the image of the map $\widehat{A}_n(\mathbb{Q}_p) \to H^1(\mathbb{Q}_p, A_n[\lambda_n])$ for all primes $p \leq \infty$. We have an exact sequence

$$0 \to \widehat{A}_n(\mathbb{Q})/\lambda_n(A_n(\mathbb{Q})) \to \mathrm{Sel}_{\lambda_n}(A_n) \to Ш(A_n)[\lambda_n] \to 0,$$

where $Ш(A_n)[\lambda_n]$ is the λ_n-torsion subgroup of the Tate-Shafarevich group of A_n.

By Proposition 3.13, there are inclusions

$$\mathrm{Sel}_{\lambda_n}(A_n) \subset H^1(\mathbb{Q}, \Theta(L_n)) \subset H^1(\mathbb{Q}, A_n[\lambda_n]).$$

If $v \in Y(\mathbb{Q})_{dn}$ corresponds, under the bijection of Theorem 3.8, to an element of $\mathrm{Sel}_{\lambda_n}(A_n)$, we say that v is *Selmer*. Equivalently, $v \in Y(\mathbb{Q})_{dn}$ is Selmer if and only if it is locally soluble at p for every prime p. Write $Y(\mathbb{Q})_{dn}^{\mathrm{sel}}$ for the set of Selmer elements v of invariant dn, and $Y(\mathbb{Q})^{\mathrm{sel}}$ for the set of all Selmer elements.

Theorem 3.18. *For each $n \in \mathbb{Q}^\times$, there is a natural bijection between the $G(F)$-orbits on $Y(F)_{dn}^{\mathrm{sel}}$ and the group $\mathrm{Sel}_{\lambda_n}(A_n)$. Under this bijection, the identity class in $\mathrm{Sel}_{\lambda_n}(A_n)$ corresponds to the unique reducible orbit in $Y(F)_{dn}$.*

Proof. This follows from Theorem 3.8. □

Theorem 3.19. *There exists a nonzero integer N such that for all nonzero $n \in N\mathbb{Z}$, and for all $v \in Y(\mathbb{Q})_{dn}^{\mathrm{sel}}$, we have $G(\mathbb{Q})v \cap Y(\mathbb{Z}) \neq \emptyset$.*

Proof. If $v \in Y(\mathbb{Q})_{dn}$, then the orbit $G(\mathbb{Q})y$ contains some $v' \in Y(\mathbb{Z})$ if and only if the orbit $G(\mathbb{Q}_p)y$ contains an element of $Y(\mathbb{Z}_p)$, for every prime p. This follows from the fact that the class number of $G = \mathrm{SL}_2^2/\mu_2$ is 1. Thus, by Theorem 3.15, it suffices to show that for any prime p, if $v_p(n)$ is large enough (depending on p), then all of the $G(\mathbb{Q}_p)$-orbits on $Y(\mathbb{Q}_p)_{nd}$ have representatives in $Y(\mathbb{Z}_p)_{nd}$.

Note that for any given prime p, there are only finitely many $G(\mathbb{Q}_p)$-orbits with A_3-invariant nd, since the set $H^1(\mathbb{Q}_p, \Theta(L_n)) \subset H^1(\mathbb{Q}_p, E_{dn}[2])$ is finite. There are also only finitely many cube-classes of $n \in \mathbb{Q}_p^\times/\mathbb{Q}_p^{\times 3}$. Thus, we may scale any $v \in Y(\mathbb{Q}_p)_{dn}$ by an appropriate power of p, to obtain an integral element of $Y(\mathbb{Q}_p)_{dn'}$ with $n' \equiv n \pmod{\mathbb{Q}_p^{\times 3}}$. It follows that if $v_p(n)$ is large enough (depending only on p), then all of the (finitely many) $G(\mathbb{Q}_p)$-orbits with A_3-invariant dn contain representatives in $Y(\mathbb{Z}_p)$. □

By the previous two theorems, in order to estimate $\sum_{|n|<X} \#\mathrm{Sel}_{\lambda_n}(A_n)$, we must determine the number of Selmer $G(\mathbb{Q})$-orbits lying in $Y(\frac{1}{N}\mathbb{Z})$ and having A_3-invariant nd bounded by X in absolute value. Here, the "Selmer" condition is determined by (infinitely many) congruence conditions. Even though we impose infinitely many congruence conditions to sieve down to the Selmer elements, the following result shows that these conditions do not throw out too many orbits.

Proposition 3.20. *There exists a nonzero integer M such that if $v \in Y(\frac{1}{M}\mathbb{Z})_{dn}$ and $dn \in \mathbb{Z}$ is cube-free at all primes $p \nmid M$, then $v \in Y(\mathbb{Q}_p)^{\mathrm{sel}}$.*

Proof. This follows from Proposition 3.17. □

Most results in this subsection generalize easily to the case where F is any number field (as well as global fields of characteristic $p > 3$). However, for the analogue of Proposition 3.20 in this more general setting, modifications are needed to account for nontrivial class group and unit group.

We turn now to the problem of estimating $\sum_{|n|<X} \#\mathrm{Sel}_{\lambda_n}(A_n)$, which we shall carry out via a suitable weighted count of $G(\mathbb{Z})$-orbits on $Y(\mathbb{Z})$ satisfying $|A_3| < X$.

4 The number of integral orbits on an invariant quadric with bounded invariants in the space of pairs of binary cubic forms

Let $V := 2 \otimes \mathrm{Sym}_3(2)$, and let G be the image in $\mathrm{GL}(V)$ of $\{(g,h) \in \mathrm{GL}_2 \times \mathrm{GL}_2 : \det g \cdot (\det h)^3 = 1\}$. Then, as already noted, G acts on V; the first GL_2 acts on the first factor via the standard representation, and the second GL_2 acts on the symmetric cube of the standard representation as usual. Note that $G \cong (\mathrm{SL}_2 \times \mathrm{SL}_2)/\mu_2$.

The action of $G(\mathbb{C})$ on $V(\mathbb{C})$ has two independent polynomial invariants denoted A_1 and A_3, and they have degrees 2 and 6, respectively. The degree 2 invariant has an especially simple formula:

$$A_1((F_1,F_2)) = 3r_1r_8 - r_2r_7 + r_3r_6 - 3r_4r_5,$$

where $F_1(x,y) = \sum_{i=0}^{3} r_{i+1}x^{3-i}y^i$ and $F_2(x,y) = \sum_{i=0}^{3} r_{i+5}x^{3-i}y^i$. (For a geometric characterization of these invariants, see Theorem 2.1, which is [9, Thm. 42].) The quadric $Y \subset V = 2 \otimes \mathrm{Sym}^3(2)$ defined by $A_1 = 0$ is thus preserved by the action of G.

In this section, we extend the methods of [9], [37], and [3], involving geometry of numbers and the circle method, to give an asymptotic formula, with a power-saving error term, for the number of $G(\mathbb{Z})$-orbits on $Y(\mathbb{Z})$ such that $|A_3| < X$ and where A_3 satisfies any acceptable set of congruence conditions. We take the methods of [37] and [3] further by expressing our asymptotic formula in terms of real and p-adic volumes with respect to a certain G-invariant measure, which represent a novel re-interpretation of the singular integral and singular series that arise in the circle method. These expressions will be key in the applications to the average sizes of Selmer groups in Section 5.

For every $n \in \mathbb{Z}$, there is a distinguished $G(\mathbb{Z})$-orbit on $Y(\mathbb{Z})$ having A_3-invariant n, namely the orbit of the pair $v_n := (F_1,F_2) = (xy^2, x^3 + ny^3)$. Note that $v_n \in Y(\mathbb{Z})$ is a *reducible* element with $|A_3|$-invariant n, since the corresponding binary quartic form $f(x_1,x_2) = \mathrm{Disc}(x_1F_1 - x_2F_2)$ has a rational root at $[1:0]$; one can check that this is the unique such orbit up to $G(\mathbb{Q})$-equivalence. Since the reducible orbit is unique, we focus on counting the irreducible ones.

More generally, we wish to count weighted irreducible $G(\mathbb{Z})$-orbits on $Y(\mathbb{Z})$ having bounded A_3-invariant, where the weights are defined by finite or appropriate infinite sets of congruence conditions. A function $\varphi : Y(\mathbb{Z}) \to [0,1] \in \mathbb{R}$ is said to be *defined by congruence conditions* if, for all primes p, there exist functions $\varphi_p : Y(\mathbb{Z}_p) \to [0,1]$ satisfying the following conditions:

(1) For all $y \in Y(\mathbb{Z})$, the product $\prod_p \varphi_p(y)$ converges to $\varphi(y)$.

(2) For each p, the function φ_p is locally constant outside some closed set in $Y(\mathbb{Z}_p)$ of measure zero.

We say that such a function φ is *acceptable* if for sufficiently large primes p, we have $\varphi_p(y) = 1$ whenever $p^2 \nmid A_3(y)$.

Let $N_\varphi(Y(\mathbb{Z}); X)$ denote the weighted number of irreducible $G(\mathbb{Z})$-orbits of elements $y \in Y(\mathbb{Z})$ with $|A_3(y)| < X$, where the orbit of each such y is weighted by $\varphi(y)$. The purpose of this section is to prove the following theorem:

Theorem 4.1. *Let $\varphi : Y(\mathbb{Z}) \to [0,1]$ be an acceptable function that is defined by the functions $\varphi_p : Y(\mathbb{Z}_p) \to [0,1]$. Then*

$$N_\varphi(Y(\mathbb{Z}); X) = X \cdot \int_{\substack{y \in G(\mathbb{Z})\backslash Y(\mathbb{R}) \\ |A_3(y)|<1}} dy \cdot \prod_p \int_{y \in Y(\mathbb{Z}_p)} \varphi_p(y)\, dy \; + \; O_\varphi\left(X^{1-\Omega(1)}\right); \tag{4.1}$$

here dy is the $G(\mathbb{R})$-invariant (resp. $G(\mathbb{Z}_p)$-invariant) measure $dr_2\, dr_3 \cdots dr_8/(\partial A_1/\partial r_1)$ on $Y(\mathbb{R})$ (resp. $Y(\mathbb{Z}_p)$), where $r_1, \ldots, r_8$ are the coordinates on V.

4.1 Counting irreducible elements of bounded height

In this subsection, we prove the following special case of Theorem 4.1 giving the asymptotic number of $G(\mathbb{Z})$-equivalence classes of irreducible elements of $Y(\mathbb{Z})$ having bounded A_3-invariant and satisfying any specified finite set of congruence conditions.

To state the result, for any $G(\mathbb{Z})$-invariant set $S \subset Y(\mathbb{Z})$, let $N(S;X)$ denote the number of $G(\mathbb{Z})$-equivalence classes of irreducible elements $y \in S$ satisfying $0 \neq |A_3(y)| < X$. Let $S_p \subset Y(\mathbb{Z}_p)$ denote the p-adic closure of S in $Y(\mathbb{Z}_p)$. We prove:

Theorem 4.2. *Let $S \subset Y(\mathbb{Z})$ be defined by a finite set of $G(\mathbb{Z})$-invariant congruence conditions modulo M. Then*

$$N(S;X) = X \cdot \int_{\substack{y \in G(\mathbb{Z})\backslash Y(\mathbb{R}) \\ |A_3(y)|<1}} dy \cdot \prod_p \int_{y \in S_p} dy + O\left(M^{O(1)} X^{1-\Omega(1)}\right).$$

4.1.1 Reduction theory

Define the (naive) *height* $H(v)$ of an element $v \in V(\mathbb{R})$ by

$$H(v) := H(A_1(v), A_3(v)) := \max\left\{|A_1(v)|^{12}, |A_3(v)|^4\right\},$$

and the discriminant $\Delta(v)$ of v by

$$\Delta(v) := \Delta(A_1(v), A_3(v)) := 16A_3(v)^3(A_1(v)^3 - 27A_3(v)).$$

Let $v(A_1, A_3) := (xy^2 + A_1y^3, x^3 + A_3y^3)$, so that $v(A_1, A_3)$ has invariants A_1 and A_3. Let

$$L_{\Delta<0} := \{v(A_1, A_3) : \Delta(A_1, A_3) < 0,\ H(A_1, A_3) = 1\}.$$

Then, as shown in [9, Section 5], $L_{\Delta<0}$ is a fundamental domain for the action of $G(\mathbb{R})$ on the set of height 1 elements of $V(\mathbb{R})_{\Delta<0}$.

Let $R = \mathbb{R}^+ \cdot L_{\Delta<0}$. Then R is a fundamental domain for the action of $G(\mathbb{R})$ on $V(\mathbb{R})_{\Delta<0}$. For any $v \in V(\mathbb{R})_{\Delta<0}$, let v_R denote the unique $G(\mathbb{R})$-representative of v in R, and v_L denote the unique $\mathbb{R}^+ \cdot G(\mathbb{R})$-representative of v in $L_{\Delta<0}$. Similarly let $\lambda_v \in \mathbb{R}^+$ be such that $v_R = \lambda_v \cdot v_L$.

Let $R(X) := \{v \in R : H(v) \leq X\}$. Since $H(\lambda v) = \lambda^{24} H(v)$, the coordinates of any $v \in \lambda L_{\Delta<0} \subseteq R(X)$ are all $O(\lambda) = O(X^{1/24})$. Hence for any compact $G_0 \subseteq G(\mathbb{R})$, the coefficients of any $v \in G_0 \cdot \lambda L_{\Delta<0} \subseteq R(X)$ are all $O_{G_0}(\lambda) = O_{G_0}(X^{1/24})$.

Let $v_\pm := (xy^2, x^3 \pm y^3) \in L_{\Delta<0}$ be the two points in $L_{\Delta<0}$ with $A_1 = 0$.

Let $\mathcal{F}$ be a fundamental domain for the action of $G(\mathbb{Z})$ on $G(\mathbb{R})$, as constructed in [9, §5.2]. Thus $\mathcal{F}$ lies inside a Siegel set; explicitly, if we write $t = (t_1, t_2)$ and $u = (u_1, u_2)$, then

$$\mathcal{F} = \{n_u a_t k : n_u \in N'(t), a_t \in A', k \in K\},$$

where

$$N'(t) := \{n_u := (n_{u_1}, n_{u_2}) \in G(\mathbb{R}) : u_i \in I(t_i)\},$$
$$A' := \left\{a_t := (a_{t_1}, a_{t_2}) \in G(\mathbb{R}) : t_i \geq \sqrt{\frac{\sqrt{3}}{2}}\right\},$$
$$K := \mathrm{SO}_2(\mathbb{R}) \times \mathrm{SO}_2(\mathbb{R}) \subseteq G(\mathbb{R});$$

here $n_{u_i} := \left(\begin{smallmatrix} 1 & 0 \\ u_i & 1 \end{smallmatrix}\right)$ and $a_{t_i} := \left(\begin{smallmatrix} t_i^{-1} & 0 \\ 0 & t_i \end{smallmatrix}\right)$, and $I(t_i)$ is a union of one or two subintervals of $[-\frac{1}{2}, \frac{1}{2}]$ depending only on the value of t_i.

In terms of these Iwasawa coordinates, we define a Haar measure dg on $G(\mathbb{R})$ by

$$dg = t_1^{-2}t_2^{-2}du_1\,du_2\,d^\times t_1\,d^\times t_2\,dk.$$

Here, dk is a Haar measure on K such that $\int_{k\in K} dk = 1$.

4.1.2 Averaging over fundamental domains

We now extend the averaging method of [5, 11, 3] to estimate $N(S;X)$. Let

$$V(\mathbb{Z})^{\mathrm{irr}} := \{(F_1,F_2)\in V(\mathbb{Z}) : \mathrm{Disc}(xF_1-yF_2) \text{ has nonzero discriminant and no root in } \mathbb{P}^1(\mathbb{Q})\}.$$

For any subset $S\subset V(\mathbb{Z})$, let $S^{\mathrm{irr}} := S\cap V(\mathbb{Z})^{\mathrm{irr}}$. Let $Y(\mathbb{Z}) := \{v\in V(\mathbb{Z}) : A_1(v)=0\}$. Our goal in this section is to count the number of $G(\mathbb{Z})$-orbits on $Y(\mathbb{Z})^{\mathrm{irr}}$ with $|A_3|<X$.

Let $\delta\in\mathbb{R}^+$ be a small absolute constant that we will choose at the end of the argument. Let $\mu_0\in C^\infty(\mathbb{R})$ be such that $\mu_0\geq 0$ on $\mathbb{R}$, $\mu_0(x)=1$ if $x\leq 0$ and $\mu_0(x)=0$ if $x\geq 1$. Let

$$\mu_+(x) := \mu_0(X^{\delta^2}(x-1))\,\mu_0(X^{\delta^2}(-x-1)),$$

and let

$$\mu_-(x) := \mu_0(X^{\delta^2}(x-1)+1)\,\mu_0(X^{\delta^2}(-x-1)+1).$$

Thus $\mu_+(x)=1$ when $|x|\leq 1$ and $\mu_+(x)=0$ when $|x|\geq 1+X^{-\delta^2}$, and similarly $\mu_-(x)=1$ when $|x|\leq 1-X^{-\delta^2}$ and $\mu_-(x)=0$ when $|x|\geq 1$. Note that $||\mu_\pm^{(k)}||_\infty \ll X^{k\delta^2}$, where $\mu_\pm^{(k)}$ denotes the k-th derivative of $\mu_\pm$.[8] Note that $\mu_\pm$ depends on the parameter X, but our notation suppresses this.

Let $\alpha\in C_c^\infty(G(\mathbb{R}))$ be a smooth, compactly supported, and K-invariant function such that $\int_{G(\mathbb{R})}\alpha = 1$. Let $\beta\in C_c^\infty(L_{\Delta<0})$ be such that $\beta(v_\pm)=\frac{1}{2}$, and such that both $\mathrm{supp}(\beta)$ and $\beta^{-1}(1/2)\subseteq L_{\Delta<0}$ are unions of two small compact intervals containing $v_\pm$. For $v\in V(\mathbb{R})$, define

$$\nu(v) := \sum_{\lambda_v\cdot g\cdot v_L = v}\alpha(g)\,\beta(v_L) = \sum_{g\cdot v_R = v}\alpha(g)\,\beta(v_L)$$

and

$$w_\pm(v;X) := \mu_\pm\left(\frac{A_3(v)}{X}\right)\nu(v).$$

Note that $w_\pm(v;X) = w_\pm(X^{-\frac{1}{6}}v;1)$.

Remark 4.3. We will suppress the dependence of all implicit constants on α, β, and μ_0.

For a $G(\mathbb{Z})$-invariant set $S\subset Y(\mathbb{Z})$, we define $N_\pm(S;X) := \sum_{v\in G(\mathbb{Z})\backslash S^{\mathrm{irr}}}\mu_\pm\left(\frac{A_3(v)}{X}\right)$. Thus

$$N_-(S;X)\leq N(S;X)\leq N_+(S;X).$$

Now because the defining sum over $G(\mathbb{Z})\backslash S^{\mathrm{irr}}$ in the definition of $N_\pm(S;X)$ may be computed using any fundamental domain of $G(\mathbb{Z})\backslash V(\mathbb{R})$, it follows that

$$N_\pm(S;X) = \sum_{v\in\mathcal{F}h\cdot R\cap S^{\mathrm{irr}}}\mu_\pm\left(\frac{A_3(v)}{X}\right)\beta(v_L)$$

for all $h\in G(\mathbb{R})$, where $\mathcal{F}h\cdot R$ records multiplicity (i.e. is a multiset) and we have used that $\beta(v_L)=\frac{1}{2}=\frac{1}{\#|\mathrm{Stab}_{G(\mathbb{R})}(v_+)|}=\frac{1}{\#|\mathrm{Stab}_{G(\mathbb{R})}(v_-)|}$ for all $v\in S^{\mathrm{irr}}\subseteq Y(\mathbb{Z})^{\mathrm{irr}}$.

[8] The point of having a δ^2 in the exponent is to have $X^{o(\delta)}$ lost upon each differentiation (so that in total we save $\gg X^{\delta-o(\delta)}$ each time we integrate by parts) during the repeated integration by parts in the proof of Proposition 4.11.

Averaging this equality over $h \in G(\mathbb{R})$ with respect to $\alpha(h)dh$, we find:

$$\begin{aligned}
N_\pm(S;X) &= \int_{h\in G(\mathbb{R})} dh \sum_{v\in \mathcal{F}h\cdot R\cap S^{\mathrm{irr}}} \mu_\pm\left(\frac{A_3(v)}{X}\right)\alpha(h)\beta(v_L)\\
&= \int_{h\in G(\mathbb{R})} dh \sum_{v\in S^{\mathrm{irr}}} \mu_\pm\left(\frac{A_3(v)}{X}\right)\alpha(h)\beta(v_L)\cdot \#|\{g\in\mathcal{F}: gh\cdot v_R = v\}|\\
&= \sum_{v\in S^{\mathrm{irr}}} \mu_\pm\left(\frac{A_3(v)}{X}\right)\beta(v_L)\int_{h\in G(\mathbb{R})} dh\,\alpha(h)\cdot \#|\{g\in\mathcal{F}: gh\cdot v_R = v\}|\\
&= \sum_{v\in S^{\mathrm{irr}}} \mu_\pm\left(\frac{A_3(v)}{X}\right)\beta(v_L)\sum_{\gamma\cdot v_R=v}\int_{h\in G(\mathbb{R})} dh\,\alpha(h)\cdot \#|\{g\in\mathcal{F}: gh=\gamma\}|\\
&= \sum_{v\in S^{\mathrm{irr}}} \mu_\pm\left(\frac{A_3(v)}{X}\right)\beta(v_L)\sum_{\gamma\cdot v_R=v}\int_{h\in G(\mathbb{R})} dh\,\alpha(h)\cdot \#|\{g\in\mathcal{F}: h=g^{-1}\cdot\gamma\}|\\
&= \sum_{v\in S^{\mathrm{irr}}} \mu_\pm\left(\frac{A_3(v)}{X}\right)\beta(v_L)\sum_{\gamma\cdot v_R=v}\int_{h\in \mathcal{F}^{-1}\gamma} dh\,\alpha(h)\\
&= \sum_{v\in S^{\mathrm{irr}}} \mu_\pm\left(\frac{A_3(v)}{X}\right)\beta(v_L)\sum_{\gamma\cdot v_R=v}\int_{h\in \mathcal{F}} dh\,\alpha(h^{-1}\gamma)\\
&= \int_{h\in\mathcal{F}} dh \sum_{v\in S^{\mathrm{irr}}} \mu_\pm\left(\frac{A_3(v)}{X}\right)\sum_{\gamma\cdot v_R=v}\alpha(h^{-1}\gamma)\beta(v_L)\\
&= \int_{h\in\mathcal{F}} dh \sum_{v\in S^{\mathrm{irr}}} \mu_\pm\left(\frac{A_3(v)}{X}\right)\sum_{\gamma\cdot v_R=h^{-1}v}\alpha(\gamma)\beta(v_L)\\
&= \int_{h\in\mathcal{F}} dh \sum_{v\in S^{\mathrm{irr}}} w_\pm(h^{-1}v;X).
\end{aligned}$$

Thus

$$N_\pm(S;X) = \int_{g\in\mathcal{F}} \sum_{v\in S\cap Y(\mathbb{Z})^{\mathrm{irr}}} w_\pm(g^{-1}v;X)\,dg. \tag{4.2}$$

We use (4.2) as the definition of $N_\pm(S;X)$ even if $S\subset V(\mathbb{Z})$ is not contained in $Y(\mathbb{Z})$ and even if S is not $G(\mathbb{Z})$-invariant. Note that in all cases, $N_\pm(S;X) = N_\pm(S\cap Y(\mathbb{Z});X)$.

For $S\subseteq V(\mathbb{Z})$ and $u=(u_1,u_2)\in I(t)$, let

$$P_\pm(u,t,X;S) := \sum_{v\in S\cap Y(\mathbb{Z})} w_\pm(a_t^{-1}n_u^{-1}\cdot v;X).$$

Then (4.2) may be re-expressed as

$$N_\pm(S;X) = \int_{g=n_u a_t\in N'(t)A'} P_\pm(u,t,X;S^{\mathrm{irr}})\,dg \tag{4.3}$$

$$= \int_{t_1,t_2=\sqrt{\frac{\sqrt{3}}{2}}}^{\infty}\int_{\substack{u_1\in I(t_1)\\ u_2\in I(t_2)}} P_\pm(u,t,X;S^{\mathrm{irr}})\,t_1^{-2}t_2^{-2}du_1\,du_2\,d^\times t_1\,d^\times t_2\,. \tag{4.4}$$

4.1.3 A sufficient condition for reducibility

The following lemma, which is [9, Lemma 6.2], gives sufficient conditions for an element in $V(\mathbb{Z})$ to be reducible.

Lemma 4.4. *Let $v = (F_1, F_2) = (r_1, \dots, r_8) \in V(\mathbb{Q})$ be an element such that either $r_1 = r_2 = 0$ or $r_1 = r_5 = 0$. Then v is reducible.*

Proof. In case (i), we see that x_2 is a factor of $\mathrm{Disc}(x_1F_1 - x_2F_2)$. In case (ii), by replacing the cubic form F_1 by a suitable $\mathbb{Q}$-linear combination of F_1 and F_2, we may transform v by an element of $G(\mathbb{Q})$ so that r_2 is zero. Since r_1 will remain zero, we are then in case (i). Hence v is reducible in either case. □

4.1.4 Cutting off the cusp

Lemma 4.5. *Let $v = (r_1, \dots, r_8) \in V(\mathbb{R})$ be such that $w_+(a_t^{-1}n_u^{-1}v; X) \neq 0$. Then*

$$|r_1| \ll t_1^{-1}t_2^{-3}X^{\frac{1}{6}}, \quad |r_2| \ll t_1^{-1}t_2^{-1}X^{\frac{1}{6}}, \quad |r_3| \ll t_1^{-1}t_2X^{\frac{1}{6}}, \quad |r_4| \ll t_1^{-1}t_2^3X^{\frac{1}{6}}$$
$$|r_5| \ll t_1t_2^{-3}X^{\frac{1}{6}}, \quad |r_6| \ll t_1t_2^{-1}X^{\frac{1}{6}}, \quad |r_7| \ll t_1t_2X^{\frac{1}{6}}, \quad |r_8| \ll t_1t_2^3X^{\frac{1}{6}}.$$

Proof. This follows by computing the action of a_t on each coordinate of $V(\mathbb{R})$ and noting that n_u lies in a compact set. □

For example, we have $P_\pm(u, t, X; S) = \sum_{v \in S \cap Y(\mathbb{Z}): ||v||_\infty \ll t_1t_2^3X^{1/6}} w_\pm(a_t^{-1}n_u^{-1}v; X)$, where $||\cdot||_\infty$ denotes the usual L^∞-norm.

Proposition 4.6. *Let $V_0 := \{(r_1X^3 + \cdots + r_4Y^3, r_5X^3 + \cdots + r_8Y^3) \in V : r_1r_8 = 0\}$. Then*

$$N_\pm(V_0(\mathbb{Z}); X) \ll_\epsilon X^{8/9+\epsilon}.$$

Proof. We first claim that $P_\pm(u, t, X; V_0(\mathbb{Z})^{\mathrm{irr}}) \ll_\epsilon t_1t_2^3X^{5/6+\epsilon}$. By Lemma 4.5 it suffices to show that the number of irreducible $v \in V_0(\mathbb{Z})$ satisfying the inequalities (4.5) in the conclusion of Lemma 4.5 is $\ll_\epsilon t_1t_2^3X^{5/6+\epsilon}$.

We will prove the claim for the subspace where r_1 vanishes—the identical argument produces an even stronger bound for the subspace where r_8 vanishes. So let us assume that $r_1 = 0$.

By Lemma 4.4, if $v \in V(\mathbb{Z})^{\mathrm{irr}}$, then $r_2, r_5 \neq 0$, and so we must have $t_1^{-1}t_2^{-1}X^{1/6} \gg 1$ and $t_1t_2^{-3}X^{1/6} \gg 1$. Hence the number of irreducible integer points satisfying (4.5) will be nonzero only if

$$t_1t_2 \ll X^{1/6} \quad \text{and} \quad t_1^{-1}t_2^3 \ll X^{1/6}. \tag{4.5}$$

Suppose that (4.5) holds and the A_1-invariant vanishes. The number of possibilities for the variables $(r_3, r_4, r_5, r_6, r_8)$ is

$$\ll t_1^{-1}t_2X^{1/6} \cdot t_1^{-1}t_2^3X^{1/6} \cdot t_1t_2^{-3}X^{1/6} \cdot t_1t_2^{-1}X^{1/6} \cdot t_1t_2^3X^{1/6} = t_1t_2^3X^{5/6}.$$

Once these five variables $(r_3, r_4, r_5, r_6, r_8)$ have been fixed, then the condition $A_1 = 0$ also fixes the value of r_2r_7. If this value is nonzero, then we conclude that the number of possibilities for the pair (r_2, r_7) is at most $O_\epsilon(X^\epsilon)$. Hence, the number of irreducible integer points which satisfy (4.5), have vanishing A_1-invariant, and are such that $r_2r_7 \neq 0$ is at most $O_\epsilon(t_1t_2^3X^{5/6+\epsilon})$.

If $r_2r_7 = 0$ (i.e., $r_7 = 0$), then the above estimate on the number of pairs (r_2, r_7) does not apply; but then we could have run the identical argument by fixing all variables except (r_4, r_5), assuming $r_4r_5 \neq 0$. If $r_7 = 0$ and $r_4r_5 = 0$ (i.e., $r_4 = 0$), then the condition $A_1 = 0$ is equivalent to $r_3r_6 = 0$. By Lemma 4.4, $r_4 = 0$ and irreducibility forces $r_3 \neq 0$, and so $r_6 = 0$. Thus the number of possibilities for $(r_1, \dots, r_8)$ (given that $r_1 = r_7 = r_4 = r_6 = 0$) is

$$\ll t_1^{-1}t_2^{-1}X^{1/6} \cdot t_1^{-1}t_2X^{1/6} \cdot t_1t_2^{-3}X^{1/6} \cdot t_1t_2^3X^{1/6} = X^{2/3}.$$

Combining all cases, we see that the number of irreducible integer points satisfying (4.5) with $r_1 r_8 = 0$ and vanishing A_1-invariant is $O_\epsilon\left(t_1 t_2^3 X^{5/6+\epsilon}\right)$. By Lemma 4.5 it follows that $P_\pm(u,t,X;V_0(\mathbb{Z})^{\mathrm{irr}}) \ll_\epsilon t_1 t_2^3 X^{5/6+\epsilon}$ as claimed.

Therefore, by the definition of $N_\pm(V_0(\mathbb{Z});X)$, we have

$$\begin{aligned} N_\pm(V_0(\mathbb{Z});X) &\ll \int_{t_2=\sqrt{\frac{\sqrt{3}}{2}}}^{X^{1/12}} \int_{t_1=\sqrt{\frac{\sqrt{3}}{2}}}^{X^{1/6}/t_2} \int_{\substack{u_1\in I(t_1)\\ u_2\in I(t_2)}} P_\pm(u,t,X;V_0(\mathbb{Z})^{\mathrm{irr}}) du_1\, du_2 \frac{dt_1}{t_1^3}\frac{dt_2}{t_2^3} \\ &\ll_\epsilon \int_{t_2=\sqrt{\frac{\sqrt{3}}{2}}}^{X^{1/12}} \int_{t_1=\max\left(\sqrt{\frac{\sqrt{3}}{2}},\frac{t_2^3}{X^{1/6}}\right)}^{X^{1/6}/t_2} t_1 t_2^3\, X^{5/6+\epsilon} \frac{dt_1}{t_1^3}\frac{dt_2}{t_2^3} \\ &\ll_\epsilon \int_{t_2=\sqrt{\frac{\sqrt{3}}{2}}}^{X^{1/18}} \int_{t_1=\sqrt{\frac{\sqrt{3}}{2}}}^{X^{1/6}/t_2} t_1 t_2^3\, X^{5/6+\epsilon} \frac{dt_1}{t_1^3}\frac{dt_2}{t_2^3} + \int_{t_2=X^{1/18}}^{X^{1/12}} \int_{t_1=\frac{t_2^3}{X^{1/6}}}^{X^{1/6}/t_2} t_1 t_2^3\, X^{5/6+\epsilon} \frac{dt_1}{t_1^3}\frac{dt_2}{t_2^3} \\ &\ll_\epsilon X^{8/9+\epsilon}, \end{aligned} \tag{4.6}$$

as desired. □

Proposition 4.7.

$$\int_{X^\delta \ll ||t||_\infty \ll X^{1/6}} \int_{||u||_\infty \ll 1} P_\pm(u,t,X;Y(\mathbb{Z})^{\mathrm{irr}}) du_1 du_2 \frac{dt_1}{t_1^3}\frac{dt_2}{t_2^3} \ll_\epsilon X^{1-2\delta+\epsilon}.$$

Proof. We follow the proof of Proposition 4.6, except that we may now assume $r_1 r_8 \neq 0$. The number of possibilities for the six coordinates $(r_2,\dots,r_7)$ for a point $(r_1,\dots,r_8) \in Y(\mathbb{Z})^{\mathrm{irr}}$ satisfying (4.5) is

$$\ll t_1^{-1}t_2^{-1}X^{1/6}\cdot t_1^{-1}t_2X^{1/6}\cdot t_1^{-1}t_2^3X^{1/6}\cdot t_1t_2^{-3}X^{1/6}\cdot t_1t_2^{-1}X^{1/6}\cdot t_1t_2X^{1/6} = X.$$

Once these six variables have been fixed, then the condition $A_1 = 0$ fixes the nonzero value of $r_1 r_8$, and thus r_1 and r_8 are determined up to at most $O_\epsilon(X^\epsilon)$ possibilities. Therefore,

$$P_\pm(u,t,X;Y(\mathbb{Z})^{\mathrm{irr}}) \ll_\epsilon X^{1+\epsilon},$$

yielding the desired result upon integration. □

4.1.5 A change-of-variable formula

For each $r \in \mathbb{R}$, let $v_r = (xy^2, x^3 + ry^3) \in Y(\mathbb{R})$.

Proposition 4.8. *There exists a rational constant $\mathcal{J}$ such that, for any $\psi \in L^1(Y(\mathbb{R}),dy)$, we have*

$$\frac{|\mathcal{J}|}{2}\int_{\mathbb{R}}\int_{G(\mathbb{R})} \psi(g\cdot v_{A_3})\, dg\, dA_3 = \int_{Y(\mathbb{R})} \psi(y)dy. \tag{4.7}$$

This can be verified by an explicit Jacobian calculation.

A more conceptual proof can also be given as follows. It was proven in [11, Proposition 3.10] that under the local identification $G(\mathbb{R})\times\mathbb{R}\times\mathbb{R} \to V(\mathbb{R})$ (onto its image, which contains $Y(\mathbb{R})_{\Delta\neq 0}$) given by $(g,A_1,A_3) \mapsto g\cdot s(A_1,A_3)$, where $s:\mathbb{R}^2 \to V(\mathbb{R})$ is a smooth section of the invariants map $V(\mathbb{R}) \to \mathbb{R}^2$, we have an equality of differential forms $dg\wedge dA_1\wedge dA_3 = c\, dv$ on $V(\mathbb{R})$; here c is a rational constant and $dv = dr_1\wedge\cdots\wedge dr_8$ is the Euclidean volume on V. If $D_1 = dg\wedge dA_3$ and $D_2 = c\, dy = c\, dr_1\wedge\cdots\wedge dr_7/(\partial A_1/\partial r_8)$, then $D_1\wedge dA_1 = cD_2\wedge dA_1$, as both are equal to $c\, dv$. We conclude by Lemma 4.9 below that $D_1 = -cD_2$ as differential forms on $Y(\mathbb{R})_{\Delta\neq 0}$. Proposition 4.8 follows.

Lemma 4.9. *Let M be a manifold. Let $k \in \mathbb{Z}^+$. Let $f \in C^\infty(M)$ be such that df is nowhere-vanishing on M. Let $\omega \in \Omega^k(M)$ be a k-form on M. Then: $df \wedge \omega = 0$ if and only if there is an $\widetilde{\omega} \in \Omega^{k-1}(M)$ such that $\omega = df \wedge \widetilde{\omega}$. In particular, if $df \wedge \alpha = df \wedge \beta$, then $\alpha|_{\{f=0\}} = \beta|_{\{f=0\}}$.*

Proof. For the forward direction, let X be a vector field on M such that the contraction $i_X(df) \in C^\infty(M)$ vanishes nowhere (e.g., choose a metric and dualize df). Then, because

$$0 = i_X(df \wedge \omega) = i_X(df) \cdot \omega - df \wedge i_X(\omega),$$

it follows that $\omega = df \wedge \left(\frac{i_X(\omega)}{i_X(df)}\right)$. The reverse direction is evident by antisymmetry since df is a 1-form, and the last sentence is then evident given the equivalence. □

We may use Proposition 4.8 to give a convenient expression for the volume of $\{y \in G(\mathbb{Z})\backslash Y(\mathbb{R}) : |A_3(y)| < 1\}$ with respect to the measure dy:

$$\int_{y \in G(\mathbb{Z})\backslash Y(\mathbb{R}) : |A_3(y)|<1} dy = \frac{|\mathcal{J}|}{2} \int_{-1}^{1} \int_{\mathcal{F}} dg \, dA_3 = |\mathcal{J}| \mathrm{Vol}(G(\mathbb{Z})\backslash G(\mathbb{R})). \tag{4.8}$$

A similar Jacobian calculation / conceptual proof yields the following more general change-of-measure formula:

Proposition 4.10. *Let K be $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{Z}_p$ for some prime p, and let $\psi \in L^1(Y(K), dy)$. Then there exists a rational constant $\mathcal{J}$, independent of K and ψ, such that*

$$\int_{Y(K)} \psi(y) dy = |\mathcal{J}| \int_{0 \neq A_3 \in K} \Big(\sum_{y \in G(K)\backslash Y_{A_3}(K)} \frac{1}{\#\,\mathrm{Stab}_{G(K)}(y)} \int_{g \in G(K)} \psi(g \cdot y) dg \Big) dA_3, \tag{4.9}$$

where $Y_{A_3}(K)$ denotes the set of elements in $Y(K)$ having invariant A_3.

Proposition 4.10 follows from [11, Proposition 3.12] using Lemma 4.9 just as Proposition 4.8 was deduced from [11, Proposition 3.10].

4.1.6 The main body

To count points in $Y(\mathbb{Z})$ in the main bodies of our fundamental domains, we use the circle/smoothed delta symbol method. For our application, unlike Heath-Brown [21] and previous treatments of the circle method, we require an estimate of the weighted number of integer points on the quadric $Y(\mathbb{R})$ in skew boxes, and we require knowledge of the dependence of the error term on the skewness of the box (i.e., on the parameter t). The other key contribution of our treatment is the expression of the singular integral and singular series both in terms of integrals over $\mathbb{R}$ and over $\mathbb{Z}_p$, respectively, with respect to the canonical measure dy.

We prove:

Proposition 4.11. *Let $S \subset Y(\mathbb{Z})$ be defined by congruence conditions modulo M. Then*

$$P_\pm(u,t,X;S) = X \cdot \int_{y \in Y(\mathbb{R})} w_\pm(y;1) dy \cdot \prod_p \int_{y \in S_p} dy + O_\epsilon(||t||_\infty^{16} M^{O(1)} X^{2/3+\epsilon})$$

Proof. We follow Heath-Brown [21], but, to handle weighted counts in skew regions, we keep careful track of the dependence of the error terms on the skewness parameter t.

We first compute $P_{\pm}(u,t,X;v_0+M\cdot V(\mathbb{Z}))$ for $v_0\in S$, and then will sum over a set of representatives of $S \pmod M$ to conclude. So let $v_0\in S$, and let $\sigma(v):=v_0+M\cdot v$. Let $\lambda:=X^{\frac{1}{6}}$. Let $N\in\mathbb{Z}^+$ with $N\asymp\delta^{-2}$.

By Heath-Brown's [21, Theorem 2, (1.2)], i.e., the key identity of the smoothed delta symbol method, we have

$$\begin{aligned}P_{\pm}(u,t,X;v_0+M\cdot V(\mathbb{Z})) &= \sum_{v\in V(\mathbb{Z}):A_1(\sigma(v))=0} w_{\pm}(a_t^{-1}n_u^{-1}\cdot\sigma(v);X)\\ &= (1+O_N(\lambda^{-N}))\lambda^{-2}\sum_{q\geq 1}q^{-8}\sum_{c\in V(\mathbb{Z})^*}\Big(\sum_{u\in(\mathbb{Z}/q)^\times}\sum_{v\in V(\mathbb{Z}/q)}e_q(uA_1(\sigma(v))+c\cdot v)\Big)\\ &\quad\cdot\Big(\int_{V(\mathbb{R})}dv\, w_{\pm}(a_t^{-1}n_u^{-1}\cdot\sigma(v);X)\cdot h\Big(\frac{q}{\lambda},\frac{A_1(\sigma(v))}{\lambda^2}\Big)\cdot e_q(-c\cdot v)\Big),\end{aligned}$$

where $e(x):=\exp(2\pi ix), e_q(x):=e(x/q)$, we have used his Theorem 1 to estimate his c_Q (here his Q is our λ), and we define h as follows: let $w_0(x):=\begin{cases}e^{-\frac{1}{1-x^2}} & |x|<1\\ 0 & |x|\geq 1\end{cases}$, $\widetilde{w}_0(x):=\frac{4w_0(4x-3)}{\int_{\mathbb{R}}w_0(x)dx}$, and $h(x,y):=\sum_{k\geq 1}\frac{1}{k\cdot x}\left(\widetilde{w}_0(k\cdot x)-\widetilde{w}_0\left(\frac{|y|}{k\cdot x}\right)\right)$. Evidently h is smooth and satisfies $h(x,y)\ll x^{-1}$ for all y and $h(x,y)=0$ when $x\gg 1+|y|$.

By [21, Lemma 25],

$$\Big|\sum_{u\in(\mathbb{Z}/q)^\times}\sum_{v\in V(\mathbb{Z}/q)}e_q(uA_1(\sigma(v))+c\cdot v)\Big|\ll q^5 \tag{4.10}$$

(we will not use sharper analysis because this will suffice for our purposes).

As for the integral, via the change of variables

$$v\mapsto\sigma^{-1}(\lambda n_u a_t\cdot\sigma(v))=\lambda n_u a_t\cdot v+\frac{(\lambda n_u a_t-\mathrm{id})\cdot v_0}{M}$$

we find:

$$\begin{aligned}&\int_{V(\mathbb{R})}w_{\pm}(a_t^{-1}n_u^{-1}\cdot\sigma(v);X)h\left(\frac{q}{\lambda},\frac{A_1(\sigma(v))}{\lambda^2}\right)e_q(-c\cdot v)\,dv\\ &=\lambda^8 e_{Mq}\left((\lambda n_u a_t-\mathrm{id})\cdot v_0\right)\int_{V(\mathbb{R})}w_{\pm}(\sigma(v);1)h\left(\frac{q}{\lambda},A_1(\sigma(v))\right)e_q(-\lambda((n_u a_t)^\dagger\cdot c)\cdot v)\,dv\end{aligned}\tag{4.11}$$

where $g^\dagger$ denotes the transpose of g.

Because $w_{\pm}(v;1)=0$ when $||v||_\infty\gg 1$, it follows that $w_{\pm}(\sigma(v);1)h\left(\frac{q}{\lambda},A_1(\sigma(v))\right)=0$ for all $v\in V(\mathbb{R})$ when $q\gg M\lambda$. Thus we assume without loss of generality that $q\ll M\lambda$.

Applying the identity $\int_{\mathbb{R}}f(x)e(\xi x)dx=-\frac{1}{2\pi i\xi}\int_{\mathbb{R}}f'(x)e(\xi x)dx$ for $f\in C_c^\infty(\mathbb{R})$ (proven via integration by parts) N times, it follows that if $||(n_u a_t)^\dagger\cdot c||_\infty\gg\frac{q}{\lambda}||t||_\infty\lambda^\delta$, then

$$\int_{V(\mathbb{R})}dv\, w_{\pm}(\sigma(v);1)h\left(\frac{q}{\lambda},A_1(\sigma(v))\right)e_q(-\lambda((n_u a_t)^\dagger\cdot c)\cdot v)\ll_N M^N X^{\delta^2 N}(||t||_\infty\lambda^\delta)^{-N}\left(\frac{\lambda}{q}\right),$$

where we have used the bound $(\partial_x^k h)(x,y)\ll_k x^{-1-k}$ as in [21, Lemma 5] as well as $||\mu_{\pm}^{(k)}||_\infty\ll X^{k\delta^2}$.

Combining (4.10) and (4.11), the sum over $c \neq 0$—which we will see contributes only to the error term—is:

$$\begin{aligned}
&\lambda^{-2}\sum_{q\geq 1} q^{-8} \sum_{0\neq c\in V(\mathbb{Z})^*} \Big(\sum_{u\in(\mathbb{Z}/q)^\times} \sum_{v\in V(\mathbb{Z}/q)} e_q(uA_1(\sigma(v)) + c\cdot v)\Big) \\
&\qquad \cdot \int_{V(\mathbb{R})} dv\, w_\pm(a_t^{-1}n_u^{-1}\sigma(v);1) h\left(\frac{q}{\lambda}, A_1(\sigma(v))\right) e_q(-c\cdot v) \\
&\ll O_N\left(M^N\lambda^{O(1)-N\delta}\right) \\
&\quad + \lambda^{6+o(1)} \sum_{\frac{\lambda^{1-\delta}}{||t||_\infty^4}\ll q\ll M\lambda} q^{-3} \sum_{0<||(n_ua_t)^\dagger\cdot c||_\infty \ll \frac{q||t||_\infty}{\lambda^{1-\delta}}} \left| \int_{v\in V(\mathbb{R})} dv\, w_\pm(\sigma(v);1) h\left(\frac{q}{\lambda}, A_1(\sigma(v))\right) e_q(-\lambda((n_ua_t)^\dagger\cdot c)\cdot v)\right|,
\end{aligned}$$

the lower bound on q in the first sum of the second term arising from the fact that

$$||(n_u\cdot a_t)^\dagger\cdot c||_\infty \gg \frac{||c||_\infty}{||t||_\infty^4},$$

and the upper bound arising from the fact that $h(x,y) = 0$ when $x \gg 1 + |y|$. Inserting absolute values into the integral, and using $h(x,y) \ll x^{-1}$, the sum over $c \neq 0$ is:

$$\begin{aligned}
&\ll O_N\big(M^N\lambda^{O(1)-N\delta}\big) + \lambda^{7+o(1)} \sum_{\frac{\lambda^{1-\delta}}{||t||_\infty^4}\ll q\ll M\lambda} q^{-4} \#\left|\left\{c\in\mathbb{Z}^8 : 0<||(n_ua_t)^\dagger\cdot c||_\infty \ll \frac{q}{\lambda^{1-\delta}}||t||_\infty\right\}\right|. \\
&\ll O_N\big(M^N\lambda^{O(1)-N\delta}\big) + \lambda^{7+o(1)} \sum_{\frac{\lambda^{1-\delta}}{||t||_\infty^4}\ll q\ll M\lambda} q^{-4} \prod_{i=0}^{1}\prod_{j=0}^{3}\left(1 + \frac{qt_1^{(-1)^i}\cdot t_2^{-3+2j}}{\lambda^{1-\delta}}||t||_\infty\right) \\
&\ll O_N\big(M^N\lambda^{O(1)-N\delta}\big) + M^{O(1)}t_1^4t_2^8 \max\left(1,\frac{t_1^2}{t_2}\right)\max\left(1,\frac{t_1^2}{t_2^3}\right)\lambda^{4+O(\delta)} \\
&\ll O_N\big(M^N\lambda^{O(1)-N\delta}\big) + M^{O(1)}||t||_\infty^{16}\lambda^{4+O(\delta)} \\
&\ll M^{O(1)}||t||_\infty^{16}\lambda^{4+O(\delta)}
\end{aligned}$$

since $N \asymp \delta^{-2}$ and $\delta \asymp 1$.

Therefore,

$$\begin{aligned}
&(1+O_N(\lambda^{-N})) \sum_{v\in V(\mathbb{Z}):A_1(\sigma(v))=0} w_\pm(a_t^{-1}n_u^{-1}\sigma(v);X) \\
&= \lambda^{-2}\sum_{q\ll M\lambda} q^{-8}\Big(\sum_{u\in(\mathbb{Z}/q)^\times}\sum_{v\in V(\mathbb{Z}/q)} e_q(uA_1(\sigma(v)))\Big)\left(\int_{V(\mathbb{R})} |dv|\, w_\pm(a_t^{-1}n_u^{-1}\sigma(v);X) h\Big(\frac{q}{\lambda}, \frac{A_1(\sigma(v))}{\lambda^2}\Big)\right) \\
&\qquad + O\big(M^{O(1)}||t||_\infty^{16}\lambda^{4+O(\delta)}\big) \\
&= M^{-8}\lambda^6\sum_{q\ll M\lambda} q^{-8}\Big(\sum_{u\in(\mathbb{Z}/q)^\times}\sum_{v\in V(\mathbb{Z}/q)} e_q(uA_1(\sigma(v)))\Big)\left(\int_{V(\mathbb{R})} |dv|\, w_\pm(v;1) h\Big(\frac{q}{\lambda}, A_1(v)\Big)\right) \\
&\qquad + O\big(M^{O(1)}||t||_\infty^{16}\lambda^{4+O(\delta)}\big),
\end{aligned}$$

where we have performed the change of variables $v \mapsto \lambda n_u a_t\cdot\sigma^{-1}(v)$.

Now we apply [21, Lemma 13]. Since $q \ll M\lambda$, we obtain

$$\int_{V(\mathbb{R})} |dv|\, w_{\pm}(v;1) h\left(\frac{q}{\lambda}, A_1(v)\right) = \int_{Y(\mathbb{R})} |dy|\, w_{\pm}(y;1) + O_N\Big(\Big(\frac{q}{\lambda}\Big)^N\Big),$$

which extracts the singular integral.

Thus

$$\begin{aligned}
&\sum_{q \ll M\lambda} q^{-8}\Big(\sum_{u\in(\mathbb{Z}/q)^\times}\sum_{v\in V(\mathbb{Z}/q)} e_q(u\cdot A_1(\sigma(v)))\Big)\Big(\int_{V(\mathbb{R})} |dv|\, w_{\pm}(v;1)\cdot h\Big(\frac{q}{\lambda}, A_1(v)\Big)\Big)\\
&= \Big(\int_{Y(\mathbb{R})} |dy|\, w_{\pm}(y;1)\Big) \sum_{q\ll M\lambda^{1-\delta}} q^{-8} \sum_{u\in(\mathbb{Z}/q)^\times}\sum_{v\in V(\mathbb{Z}/q)} e_q(uA_1(\sigma(v))) + O\big(M^{O(1)}\lambda^{4+O(\delta)}\big)\\
&= \Big(\int_{Y(\mathbb{R})} |dy|\, w_{\pm}(y;1)\Big) \sum_{q\geq 1} q^{-8} \sum_{u\in(\mathbb{Z}/q)^\times}\sum_{v\in V(\mathbb{Z}/q)} e_q(uA_1(\sigma(v))) + O\big(M^{O(1)}\lambda^{4+O(\delta)}\big),
\end{aligned}$$

where we have used [21, Lemma 25] twice.

We now turn towards the singular series, which we will re-express as an analogous p-adic integral. The function $q \mapsto q^{-8}\sum_{u\in(\mathbb{Z}/q)^\times}\sum_{v\in V(\mathbb{Z}/q)} e_q(u\cdot A_1(\sigma(v)))$ is multiplicative, whence:

$$\sum_{q\geq 1} q^{-8} \sum_{u\in(\mathbb{Z}/q)^\times}\sum_{v\in V(\mathbb{Z}/q)} e_q(u\cdot A_1(\sigma(v))) = \prod_p \sum_{k\geq 1} p^{-8k} \sum_{u\in(\mathbb{Z}/p^k)^\times}\sum_{v\in V(\mathbb{Z}/p^k)} e_q(u\cdot A_1(\sigma(v))).$$

Now

$$\begin{aligned}
&\sum_{u\in(\mathbb{Z}/p^k)^\times}\sum_{v\in V(\mathbb{Z}/p^k)} e_q(uA_1(\sigma(v)))\\
&= \varphi(p^k)\#|\{v\in V(\mathbb{Z}/p^k) : p^k \mid A_1(\sigma(v))\}| - p^{k-1}\#|\{v\in V(\mathbb{Z}/p^k) : p^{k-1} \mid\mid A_1(\sigma(v))\}|\\
&= p^k\#|\{v\in V(\mathbb{Z}/p^k) : p^k \mid A_1(\sigma(v))\}| - p^{k-1}\#|\{v\in V(\mathbb{Z}/p^k) : p^{k-1} \mid A_1(\sigma(v))\}|.
\end{aligned}$$

But $A_1(\sigma(v)) \pmod{p^{k-1}}$ only depends on $v \pmod{p^{k-1}}$. Therefore,

$$\begin{aligned}
&\sum_{u\in(\mathbb{Z}/p^k)^\times}\sum_{v\in V(\mathbb{Z}/p^k)} e_q(u\cdot A_1(\sigma(v)))\\
&= p^k\cdot\#|\{v\in V(\mathbb{Z}/p^k) : p^k \mid A_1(\sigma(v))\}| - p^{k+7}\cdot\#|\{v\in V(\mathbb{Z}/p^{k-1}) : p^{k-1} \mid A_1(\sigma(v))\}|.
\end{aligned}$$

Since the sum telescopes,

$$\begin{aligned}
\sum_{k\geq 1} p^{-8k} \sum_{u\in(\mathbb{Z}/p^k)^\times}\sum_{v\in V(\mathbb{Z}/p^k)} e_q(u\cdot A_1(\sigma(v))) &= \lim_{k\to\infty} p^{-7k}\cdot\#|\{v\in V(\mathbb{Z}/p^k) : A_1(\sigma(v)) \equiv 0 \pmod{p^k}\}|\\
&= \lim_{k\to\infty}\frac{1}{p^{-k}}\int_{v\in V(\mathbb{Z}_p):|A_1(\sigma(v))|_p\leq p^{-k}} |dv|_p\\
&= |M|_p^{-8p}\cdot\lim_{k\to\infty}\frac{1}{p^{-k}}\int_{\substack{v\in V(\mathbb{Z}_p): v\equiv v_0 \pmod M,\\ |A_1(v)|_p\leq p^{-k}}} |dv|_p,
\end{aligned}$$

where in the last equality we have made the change of variables $v \mapsto \sigma^{-1}(v)$.

Because A_1 is nondegenerate for all $v \in V(\mathbb{Z}_p)$, we have

$$||(\nabla A_1)(v)||_\infty \gg ||v||_\infty$$

(proven, e.g., via the adjugate). It follows that either $||v||_\infty \le \delta^{-1}p^{-\frac{k}{2}}$, i.e. v lies in a set of measure $\ll \big(\delta^{-1}p^{-\frac{k}{2}}\big)^8 = \delta^{-8}p^{-4k}$, or else $||(\nabla A_1)(v)||_\infty \gg \delta^{-1}p^{-\frac{k}{2}}$.

Suppose that $||(\nabla A_1)(v)||_\infty \gg \delta^{-1}p^{-\frac{k}{2}}$ and furthermore that $|A_1(v)|_p \le p^{-k}$. Let j be minimal such that $|(\partial_{v_j}A_1)(v)|_p \gg \delta^{-1}p^{-\frac{k}{2}}$. Then $A_1(v) \equiv 0 \pmod{p(\partial_{v_j}A_1)(v)^2}$, whence by Hensel's lemma there is a unique v' with $A_1(v') = 0$ and such that $v'_i = v_i$ for $i \ne j$ and $v'_j \equiv v_j \pmod{p(\partial_{v_j}A_1)(v)}$. Thus $||v' - v||_\infty \le p^{-1}|(\partial_{v_j}A_1)(v)|_p$ and hence $|(\partial_{v_j}A_1)(v')|_p \gg \delta^{-1}p^{-\frac{k}{2}}$ as well.

In reverse, given v' such that $A_1(v') = 0$, $|(\partial_{v_i}A_1)(v')|_p \ll \delta^{-1}p^{-\frac{k}{2}}$ for $i < j$, and such that $|(\partial_{v_j}A_1)(v')|_p \gg \delta^{-1}p^{-\frac{k}{2}}$, the measure of the set of v such that $A_1(v) \equiv 0 \pmod{p^k}$, $v_i = v'_i$ for $i \ne j$, and $v'_j \equiv v_j \pmod{p(\partial_{v_j}A_1)(v)}$ is $p^{-k}|(\partial_{v_j}A_1)(v')|_p^{-1}$, by Taylor expansion.

Thus, for each $1 \le j \le 8$, via the map $v = (v_1, v_2, \dots, v_8) \mapsto v' = (v_1, v_2, \dots, v_{j-1}, v'_j, v_{j+1}, \dots, v_8)$, we have

$$\begin{aligned}
&\frac{1}{p^{-k}} \int_{\substack{v \in V(\mathbb{Z}_p): v \equiv v_0 \pmod M, \\ |A_1(v)|_p \le p^{-k}, |(\partial_{v_j}A_1)(v)|_p \gg \delta^{-1}p^{-\frac{k}{2}}, \\ |(\partial_{v_i}A_1)(v)|_p \ll \delta^{-1}p^{-\frac{k}{2}}\ \forall i<j}} |dv|_p \\
&= \frac{1}{p^{-k}} \int_{v_1,\dots,\hat{v}_j,\dots,v_8 \in \mathbb{Z}_p} |dv_1|_p \cdots \widehat{|dv_j|_p} \cdots |dv_8|_p \int_{\substack{v_j \in \mathbb{Z}_p: v \equiv v_0 \pmod M, \\ |A_1(v)|_p \le p^{-k}, |(\partial_{v_j}A_1)(v)|_p \gg \delta^{-1}p^{-\frac{k}{2}}, \\ |(\partial_{v_i}A_1)(v)|_p \ll \delta^{-1}p^{-\frac{k}{2}}\ \forall i<j}} |dv_j|_p \\
&= \int_{\substack{y \in Y(\mathbb{Z}_p): y \equiv v_0 \pmod M, \\ |(\partial_{v_j}A_1)(y)|_p \gg \delta^{-1}p^{-\frac{k}{2}}, \\ |(\partial_{v_i}A_1)(y)|_p \ll \delta^{-1}p^{-\frac{k}{2}}\ \forall i<j}} \frac{|dv_1|_p \cdots \widehat{|dv_j|_p} \cdots |dv_8|_p}{|(\partial_{v_j}A_1)(y)|_p} = \int_{\substack{y \in Y(\mathbb{Z}_p): y \equiv v_0 \pmod M, \\ |(\partial_{v_j}A_1)(y)|_p \gg \delta^{-1}p^{-\frac{k}{2}}, \\ |(\partial_{v_i}A_1)(y)|_p \ll \delta^{-1}p^{-\frac{k}{2}}\ \forall i<j}} |dy|_p,
\end{aligned}$$

where in the last equality we have implicitly used Lemma 4.9.

Since, as mentioned,

$$\frac{1}{p^{-k}} \int_{\substack{v \in V(\mathbb{Z}_p): v \equiv v_0 \pmod M, \\ |A_1(v)|_p \le p^{-k}}} |dv|_p = \sum_{j=1}^{8} \frac{1}{p^{-k}} \int_{\substack{v \in V(\mathbb{Z}_p): v \equiv v_0 \pmod M, \\ |A_1(v)|_p \le p^{-k}, |(\partial_{v_j}A_1)(v)|_p \gg \delta^{-1}p^{-\frac{k}{2}}, \\ |(\partial_{v_i}A_1)(v)|_p \ll \delta^{-1}p^{-\frac{k}{2}}\ \forall i<j}} |dv|_p + O(\delta^{-8}p^{-3k}),$$

it follows that

$$\lim_{k\to\infty} \frac{1}{p^{-k}} \int_{\substack{v \in V(\mathbb{Z}_p): v \equiv v_0 \pmod M, \\ |A_1(v)|_p \le p^{-k}}} |dv|_p = |M|_p^{-8} \int_{y \in Y(\mathbb{Z}_p): y \equiv v_0 \pmod M} |dy|_p.$$

Therefore, on collecting everything together, we have found that:

$$\begin{aligned}
&\sum_{v \in V(\mathbb{Z}): A_1(\sigma(v))=0} w_\pm(a_t^{-1}n_u^{-1} \cdot \sigma(v); X) \\
&= M^{-8}\lambda^6\Big(\int_{Y(\mathbb{R})} |dy|\, w_\pm(y;1)\Big) \prod_p |M|_p^{-8} \int_{y \in Y(\mathbb{Z}_p): y \equiv v_0 \pmod M} |dy|_p + O\big(M^{O(1)}||t||_\infty^{16}\lambda^{4+O(\delta)}\big) \\
&= \lambda^6\Big(\int_{Y(\mathbb{R})} |dy|\, w_\pm(y;1)\Big) \prod_p \int_{y \in Y(\mathbb{Z}_p): y \equiv v_0 \pmod M} |dy|_p + O\big(M^{O(1)}||t||_\infty^{16}\lambda^{4+O(\delta)}\big).
\end{aligned}$$

Finally, we sum over representatives v_0 of $S \pmod M \subseteq Y(\mathbb{Z}/M)$ to conclude that

$$\sum_{v\in S} w_\pm(a_t^{-1}n_u^{-1}\cdot v;X) = \lambda^6 \int_{y\in Y(\mathbb{R})} |dy|\, w_\pm(y;1) \prod_p \int_{y\in S_p} |dy|_p + O\big(M^{O(1)}||t||_\infty^{16}\lambda^{4+O(\delta)}\big).$$

Choosing $\delta \asymp 1$ sufficiently small gives the claim. □

4.1.7 Estimates on reducibility in the main body

Let g be the completely multiplicative function vanishing outside the squarefree integers such that

$$g(p) = \frac{\int_{y\in Y(\mathbb{Z}_p)\cap V(\mathbb{Z}_p)^{\mathrm{irr}}} dy}{\int_{y\in Y(\mathbb{Z}_p)} dy}$$

for all primes p. Note that $0 \le g(p) \le 1$. Let h be the completely multiplicative function vanishing outside the squarefree integers such that $h(p) := \frac{g(p)}{1-g(p)}$ if $g(p) < 1$ and $h(p) := 2$ otherwise.

Lemma 4.12. *Let $p > 3$ be a prime. Then*

$$g(p) \ge \begin{cases} \frac14 + O\left(\frac1p\right) & \text{if } p \equiv 1 \pmod 3, \\ \frac12 + O\left(\frac1p\right) & \text{if } p \equiv 2 \pmod 3. \end{cases}$$

Proof. By Proposition 4.10, we have

$$1 - g(p) = \frac{\int_{0\neq A_3\in\mathbb{Z}_p} \left(\sum_{y\in G(\mathbb{Z}_p)\backslash Y_{A_3}(\mathbb{Z}_p)^{\mathrm{red}}} \frac{1}{\#|\operatorname{Stab}_{G(\mathbb{Z}_p)}(y)|}\right) dA_3}{\int_{0\neq A_3\in\mathbb{Z}_p} \left(\sum_{y\in G(\mathbb{Z}_p)\backslash Y_{A_3}(\mathbb{Z}_p)} \frac{1}{\#|\operatorname{Stab}_{G(\mathbb{Z}_p)}(y)|}\right) dA_3},$$

where $Y_{A_3}(\mathbb{Z}_p)^{\mathrm{red}} := Y_{A_3}(\mathbb{Z}_p) \cap V(\mathbb{Z}_p)^{\mathrm{red}}$.

For $n \in \mathbb{Z}_p^\times$, let E_n be the elliptic curve $E_n\colon y^2 = x^3 + 16n^2$. For any $v \in Y_n(\mathbb{Z}_p)$, we have

$$\operatorname{Stab}_{G(\mathbb{Z}_p)}(y) \simeq E_n[2](\mathbb{Q}_p),$$

and there is a bijection between $G(\mathbb{Z}_p)\backslash Y_n(\mathbb{Z}_p)$ and $E_n(\mathbb{Q}_p)/2E_n(\mathbb{Q}_p)$, by Proposition 3.17. Under this bijection, the unique reducible orbit corresponds to the identity element.

If $p \equiv 1 \pmod 3$ and $16n^2$ is a cube in $\mathbb{Z}_p^\times$, then

$$\#E_n(\mathbb{Q}_p)/2E_n(\mathbb{Q}_p) = \#E_n[2](\mathbb{Q}_p) = 4,$$

so for such n there are three irreducible orbits for every reducible one. If $16n^2$ is not a cube, then $\#E_n(\mathbb{Q}_p)/2E_n(\mathbb{Q}_p) = 1$ and there is only the reducible orbit. Since $16n^2$ is a cube for one third of all $n \in \mathbb{Z}_p^\times$, we have

$$1 - g(p) = \frac13 \cdot \frac14 + \frac23 \cdot 1 + O\left(\frac1p\right)$$

and hence $g(p) \ge \frac14 + O\big(\frac1p\big)$.

If $p \equiv 2 \pmod 3$ and $p > 2$, then

$$\#E_n(\mathbb{Q}_p)/2E_n(\mathbb{Q}_p) = \#E_n[2](\mathbb{Q}_p) = 2,$$

so for each n, there are two orbits and one of them is reducible. Thus $g(p) \ge \frac12 + O\big(\frac1p\big)$. □

We now define the usual Selberg sieve weights. Let $\eta \in \mathbb{R}^+$ with $\eta \asymp 1$ (certainly $\eta = 10^{-10}$ will suffice). Let $R := X^\eta$. Let $D := R^2$. Let $Q := \prod_{p \le R} p$. Let $J := \sum_{m \le R} h(m)$. Let $\rho_e := \frac{1}{J} \cdot \frac{\mu(e)}{g(e)} \cdot \sum_{e|m|P : m \le R} h(m)$; thus, e.g., $\rho_1 = 1$, $\rho_e = 0$ when $e > R$, and $|\rho_e| \le 1$.[9] Let $\lambda_d := \sum_{[e,e']=d} \rho_e \cdot \rho_{e'}$; thus e.g. $|\lambda_m| \le d_3(m)$, $\lambda_m = 0$ when $m > D$, and $\sum_d \lambda_d \cdot g(d) = \frac{1}{J}$.[10]

Proposition 4.13.

$$\int_{1 \ll ||t||_\infty \ll X^\delta} \int_{||u||_\infty \ll 1} P_\pm(u, t, X; Y(\mathbb{Z})^{\mathrm{red}})\, t_1^{-2} t_2^{-2} du_1\, du_2\, d^\times t_1\, d^\times t_2 \ll_\delta X^{1-\Omega(1)+O(\delta)}. \tag{4.12}$$

Proof. Let $\lambda := X^{\frac{1}{6}}$. Evidently

$$\Big(\sum_{e|Q} \rho_e \cdot \mathbb{1}_{\cap_{p|e} V(\mathbb{Z}_p)^{\mathrm{irr}}}(y) \Big)^2 \ge \mathbb{1}_{\cap_{p \le R} V(\mathbb{Z}_p)^{\mathrm{red}}}(y)$$

(if $y \notin \bigcap_{p \le R} V(\mathbb{Z}_p)^{\mathrm{red}}$ there is nothing to prove, and if $y \in \bigcap_{p \le R} V(\mathbb{Z}_p)^{\mathrm{red}}$ then both sides are 1). Summing over $y \in Y(\mathbb{Z})$, we obtain

$$\begin{aligned} P_\pm(u, t, X; \bigcap_{p \le R} V(\mathbb{Z}_p)^{\mathrm{red}}) &= \sum_{y \in Y(\mathbb{Z}) \cap \bigcap_{p \le R} V(\mathbb{Z}_p)^{\mathrm{red}}} w_\pm(a_t^{-1} n_u^{-1} \cdot y; X) \\ &\le \sum_{m|Q} \lambda_m \sum_{y \in Y(\mathbb{Z}) \cap \bigcap_{p|m} V(\mathbb{Z}_p)^{\mathrm{irr}}} w_\pm(a_t^{-1} n_u^{-1} \cdot y; X). \end{aligned}$$

By Theorem 4.11,

$$P_\pm(u, t, X; \bigcap_{p|m} V(\mathbb{Z}_p)^{\mathrm{irr}}) = X \int_{y \in Y(\mathbb{R})} w_\pm(y; 1) dy \prod_{p \nmid m} \int_{y \in Y(\mathbb{Z}_p)} dy \prod_{p|m} \int_{y \in Y(\mathbb{Z}_p) \cap V(\mathbb{Z}_p)^{\mathrm{irr}}} dy + O\big(||t||_\infty^{16} m^{O(1)} X^{\frac{2}{3}} \big).$$

Thus, on using $|\lambda_m| \le d_3(m)$ and $\lambda_m = 0$ when $m > D$, we find:

$$\begin{aligned} &P_\pm(u, t, X; \bigcap_{p \le R} V(\mathbb{Z}_p)^{\mathrm{red}}) \\ &\le X \cdot \int_{y \in Y(\mathbb{R})} w_\pm(y; 1) dy \cdot \prod_p \int_{y \in Y(\mathbb{Z}_p)} dy \cdot \sum_{m|Q} \lambda_m g(m) + O\Big(X^{\frac{2}{3}} ||t||_\infty^{16} \sum_{m \le D} d_3(m) m^{O(1)} \Big) \\ &= \frac{X}{J} \cdot \int_{y \in Y(\mathbb{R})} w_\pm(y; 1) dy \cdot \prod_p \int_{y \in Y(\mathbb{Z}_p)} dy + O(X^{\frac{2}{3} + O(\eta)} ||t||_\infty^{16}). \end{aligned}$$

[9] This is because

$$|\rho_e| = \frac{1}{J} \sum_{e|m|P : m \le R} \sum_{e'|e} \frac{h(m)}{h(e')} = \frac{1}{J} \sum_{m'|Q : m' \le \frac{R}{e}, (m,e)=1} \sum_{e''|e} h(m' \cdot e'') \le 1.$$

[10] Indeed

$$\begin{aligned} \sum_d \lambda_d \cdot g(d) &= \frac{1}{J^2} \sum_{e,e'} \frac{\mu(e)\mu(e')}{g((e,e'))} \Big(\sum_{e|m|Q : m \le R} h(m) \Big) \Big(\sum_{e'|m'|Q : m' \le R} h(m') \Big) \\ &= \frac{1}{J^2} \sum_{m,m'|Q : m,m' \le R} h(m) h(m') \sum_{e|m} \mu(e) \sum_{e''|(e,m')} \frac{\mu(e'')}{g(e'')} \sum_{e'''|\frac{m'}{(e,m')}} \mu(e'''), \end{aligned}$$

and $\sum_{e''|m'} \frac{\mu(e'')}{g(e'')} = \frac{\mu(m')}{h(m')}$.

Finally, because $g(p) \gg 1$ when $p \gg 1$ (Lemma 4.12) it follows that

$$J = \sum_{m \leq R} h(m) \geq \sum_{m \leq R : (m, O(1))=1} O(1)^{-\#|\{p|m : p \gg 1\}|} \gg \frac{R}{\log R},$$

for example, whence we conclude from

$$P_{\pm}(u, t, X; Y(\mathbb{Z})^{\mathrm{red}}) \leq P_{\pm}(u, t, X; Y(\mathbb{Z}) \cap \bigcap_{p \leq R} V(\mathbb{Z}_p)^{\mathrm{red}})$$

that

$$P_{\pm}(u, t, X; Y(\mathbb{Z})^{\mathrm{red}}) \ll X^{1-\eta+o(1)} + X^{\frac{2}{3}+O(\eta)} ||t||_{\infty}^{16}.$$

Integrating over u and t then gives the claim. □

4.1.8 Putting together the cusp and main body

We now prove Theorem 4.2.

Proof of Theorem 4.2. Since $N_{-}(S; X) \leq N(S; X) \leq N_{+}(S; X)$ it suffices to prove the same asymptotic for $N_{\pm}(S; X)$ instead.

By (4.3),

$$\begin{aligned}
&N_{\pm}(S; X) \\
&= \int_{t_1, t_2 = \sqrt{\frac{\sqrt{3}}{2}}}^{\infty} \int_{\substack{u_1 \in I(t_1) \\ u_2 \in I(t_2)}} P_{\pm}(u, t, X; S^{\mathrm{irr}}) t_1^{-2} t_2^{-2} du_1\, du_2\, d^{\times} t_1\, d^{\times} t_2 \\
&= \int_{t_1, t_2 = \sqrt{\frac{\sqrt{3}}{2}}}^{X^{O(1)}} \int_{\substack{u_1 \in I(t_1) \\ u_2 \in I(t_2)}} P_{\pm}(u, t, X; S^{\mathrm{irr}})\, t_1^{-2} t_2^{-2} du_1\, du_2\, d^{\times} t_1\, d^{\times} t_2 \\
&= \left(\int_{1 \ll ||t||_{\infty} < X^{\delta}} \int_{\substack{u_1 \in I(t_1) \\ u_2 \in I(t_2)}} + \int_{X^{\delta} \leq ||t||_{\infty} \ll X^{O(1)}} \int_{\substack{u_1 \in I(t_1) \\ u_2 \in I(t_2)}} \right) P_{\pm}(u, t, X; S^{\mathrm{irr}})\, t_1^{-2} t_2^{-2} du_1\, du_2\, d^{\times} t_1\, d^{\times} t_2,
\end{aligned}$$

where in the second equality we used that if $||t||_{\infty} \gg X^{\frac{1}{6}}$ then by (4.5) there are no irreducible integer points $y \in Y(\mathbb{Z})$ with $w_{\pm}(a_t^{-1} n_u^{-1} y; X) \neq 0$.

By Proposition 4.7 the second integral is $\ll X^{1-\Omega(1)}$.

Let us further write the first integral as:

$$\begin{aligned}
&\int_{1 \ll ||t||_{\infty} < X^{\delta}} \int_{\substack{u_1 \in I(t_1) \\ u_2 \in I(t_2)}} P_{\pm}(u, t, X; S^{\mathrm{irr}}) t_1^{-2} t_2^{-2} du_1\, du_2\, d^{\times} t_1\, d^{\times} t_2 \\
&= \int_{1 \ll ||t||_{\infty} < X^{\delta}} \int_{\substack{u_1 \in I(t_1) \\ u_2 \in I(t_2)}} P_{\pm}(u, t, X; S) t_1^{-2} t_2^{-2} du_1\, du_2\, d^{\times} t_1\, d^{\times} t_2 \\
&\quad - \int_{1 \ll ||t||_{\infty} < X^{\delta}} \int_{\substack{u_1 \in I(t_1) \\ u_2 \in I(t_2)}} P_{\pm}(u, t, X; S^{\mathrm{red}}) t_1^{-2} t_2^{-2} du_1\, du_2\, d^{\times} t_1\, d^{\times} t_2.
\end{aligned}$$

By Proposition 4.13 the second integral is $\ll X^{1-\Omega(1)}$. As for the first integral, applying Proposition 4.11 and observing that the resulting error terms also integrate to $\ll M^{O(1)} X^{1-\Omega(1)}$ and that

the main term is independent of t and u, we need only observe that

$$\begin{aligned}
&\int_{1\ll||t||_\infty<X^\delta}\int_{\substack{u_1\in I(t_1)\\ u_2\in I(t_2)}} t_1^{-2}t_2^{-2}dy\,du_1\,du_2\,d^\times t_1\,d^\times t_2\\
&=\left(\int_{t_1,t_2=\sqrt{\frac{\sqrt{3}}{2}}}^{\infty}\int_{\substack{u_1\in I(t_1)\\ u_2\in I(t_2)}} - \int_{||t||_\infty\geq X^\delta}\int_{\substack{u_1\in I(t_1)\\ u_2\in I(t_2)}}\right)t_1^{-2}t_2^{-2}dy\,du_1\,du_2\,d^\times t_1\,d^\times t_2\\
&=\mathrm{Vol}(G(\mathbb{Z})\backslash G(\mathbb{R}))+O(X^{1-\Omega(1)}).
\end{aligned}$$

and that (by Proposition 4.8)

$$\begin{aligned}
&\int_{y\in Y(\mathbb{R})} w_\pm(y;1)dy\\
&=\frac{|\mathcal{J}|}{2}\int_{g\in G(\mathbb{R})}\int_{A_3\in\mathbb{R}} w_\pm(g\cdot v(0,A_3);1)dA_3\,dg\\
&=\frac{|\mathcal{J}|}{2}\int_{g\in G(\mathbb{R})}\int_{A_3\in\mathbb{R}} \mu_\pm(A_3)\sum_{hg\cdot v(0,A_3)=v(0,A_3)}\alpha(h)\beta(v_{\mathrm{sgn}(A_3)})dA_3\,dg\\
&=\frac{|\mathcal{J}|}{4}\int_{g\in G(\mathbb{R})}\left(\int_{A_3\in\mathbb{R}^+}\mu_\pm(A_3)\sum_{h\in\mathrm{Stab}_{G(\mathbb{R})}(v_+)}\alpha(hg^{-1})dA_3+\int_{A_3\in\mathbb{R}^-}\mu_\pm(A_3)\sum_{h\in\mathrm{Stab}_{G(\mathbb{R})}(v_-)}\alpha(hg^{-1})dA_3\right)dg\\
&=\frac{|\mathcal{J}|}{4}\left(\int_{A_3\in\mathbb{R}^+}\mu_\pm(A_3)\sum_{h\in\mathrm{Stab}_{G(\mathbb{R})}(v_+)}+\int_{A_3\in\mathbb{R}^-}\mu_\pm(A_3)\sum_{h\in\mathrm{Stab}_{G(\mathbb{R})}(v_-)}\right)\int_{g\in G(\mathbb{R})}\alpha(hg^{-1})dg\,dA_3\\
&=\frac{|\mathcal{J}|}{2}\int_{A_3\in\mathbb{R}}\mu_\pm(A_3)\\
&=|\mathcal{J}|+O(X^{-\delta^2}),
\end{aligned}$$

where we have used that $\int_{g\in G(\mathbb{R})}\alpha(hg^{-1})dg=1$ via the change of variables $g\mapsto g^{-1}h$.

It follows that

$$\begin{aligned}
\int_{1\ll||t||_\infty<X^\delta}\int_{\substack{u_1\in I(t_1)\\ u_2\in I(t_2)}}\int_{y\in Y(\mathbb{R})} w_\pm(y;1)t_1^{-2}t_2^{-2}dy\,du_1\,du_2\,d^\times t_1\,d^\times t_2&=|\mathcal{J}|\cdot\mathrm{Vol}(G(\mathbb{Z})\backslash G(\mathbb{R}))+O(X^{1-\Omega(1)})\\
&=\int_{\substack{y\in G(\mathbb{Z})\backslash Y(\mathbb{R}),\\ |A_3(y)|<1}}dy+O(X^{1-\Omega(1)}),
\end{aligned}$$

and so by Proposition 4.11 we are done. □

4.2 A uniformity estimate

We will make use of the following theorem of Browning–Heath-Brown [14, Theorem 1.1].

Theorem 4.14 (Browning–Heath-Brown). *Let $X\subseteq\mathbb{P}^m$ be a hypersurface defined over $\mathbb{Q}$ by a quadratic form of rank at least 5. Let $Z\subseteq X$ be a codimension 2 subvariety defined over $\mathbb{Q}$, let $\mathscr{Z}$ be its scheme-theoretic closure in $\mathbb{P}^m_{\mathbb{Z}}$, and let $Z_p:=\mathscr{Z}\otimes_{\mathbb{Z}}\mathbb{F}_p$. Then for any $\epsilon>0$ there exists a constant $c_{\epsilon,X,Z}>0$ depending only on X, Z, and ϵ, such that the number of $x\in X(\mathbb{Q})$ of height $H_{\mathbb{P}^m}(x)<B$ which specialize to a point in $Z_p(\mathbb{F}_p)$ for some $p>M$ is at most*

$$c_{\varepsilon,X,Z}B^\varepsilon\left(\frac{B^{m-1}}{M\log M}+B^{m-1-1/m}\right).$$

Here $H_{\mathbb{P}^m}([x_0,\dots,x_m]):=\prod_v\max_i|x_i|_v$ as usual.

We now prove the desired uniformity estimate. Let $W_p(Y)$ denote the set of $y \in Y(\mathbb{Z})$ such that $p^2 \mid A_3(y)$.

Proposition 4.15. *Suppose $M \in \mathbb{Z}^+$ with $M \gg 1$. Then*

$$N(\bigcup_{p>M} W_p(Y); X) \ll \frac{X^{1+O(\delta^3)}}{M \log M} + X^{1-\Omega(1)}.$$

Proof. By Proposition 4.7,

$$\begin{aligned}
&N(\bigcup_{p>M} W_p(Y); X) \\
&\leq N_+(\bigcup_{p>M} W_p(Y); X) \\
&= \int_{t_1,t_2=\sqrt{\frac{\sqrt{3}}{2}}}^{\infty} \int_{\substack{u_1 \in I(t_1) \\ u_2 \in I(t_2)}} P_+(u,t,X; \bigcup_{p>M} W_p(Y)^{\mathrm{irr}})\, t_1^{-2} t_2^{-2} du_1\, du_2\, d^\times t_1\, d^\times t_2 \\
&= \left(\int_{\sqrt{\frac{\sqrt{3}}{2}} \leq ||t||_\infty < X^{\delta^3}} + \int_{||t||_\infty > X^{\delta^3}} \right) \int_{\substack{u_1 \in I(t_1) \\ u_2 \in I(t_2)}} P_+(u,t,X; \bigcup_{p>M} W_p(Y)^{\mathrm{irr}})\, t_1^{-2} t_2^{-2} du_1\, du_2\, d^\times t_1\, d^\times t_2 \\
&= \int_{\sqrt{\frac{\sqrt{3}}{2}} \leq ||t||_\infty < X^{\delta^3}} \int_{\substack{u_1 \in I(t_1) \\ u_2 \in I(t_2)}} P_+(u,t,X; \bigcup_{p>M} W_p(Y)^{\mathrm{irr}})\, t_1^{-2} t_2^{-2} du_1\, du_2\, d^\times t_1\, d^\times t_2 + O(X^{1-\Omega(1)}).
\end{aligned}$$

Thus it suffices to prove that, for $1 \ll t_i \ll X^{\delta^3}$,

$$P_+(u,t,X; \bigcup_{p>M} W_p(Y)^{\mathrm{irr}}) \ll \frac{X^{1+O(\delta^3)}}{M \log M} + X^{1-\Omega(1)},$$

because inserting this bound into the above integral yields the claim. Note also that for such $t = (t_1, t_2)$, by Lemma 4.5 if $w_+(a_t^{-1} n_u^{-1} v; X) \neq 0$, then, writing $v = (r_1, \ldots, r_8)$, we have $H_{\mathbb{P}^7}([r_1 : \ldots : r_8]) \ll X^{1/6+O(\delta^3)}$.

Let now $v = (F_1, F_2) = (r_1, \ldots, r_8) \in V(\mathbb{Z})$ be such that $w_+(a_t^{-1} n_u^{-1} v; X) \neq 0$ and $p > M$ be such that $p^2 \mid A_3(v)$. Let also $f = \mathrm{Disc}_{X,Y}(xF_1(X,Y) - yF_2(X,Y)) \in \mathbb{Z}[x,y]$. Since f is a binary quartic form, we may refer to its invariants $I(f)$ and $J(f)$ [11], which we will write $I(v)$ and $J(v)$. From a computation, we have $A_1(v) \mid I(f)$. If $v \in Y(\mathbb{Z})$, then $A_1(v) = 0$ and hence $I(f) = 0$. It then follows from degree considerations that $J(f)$ is a nonzero rational constant times $A_3(v)^2$. Since $p^2 \mid A_3(v)$, it follows that $p^4 \mid J(v)$.

Since $I(f) = 0$, the discriminant of $f(x,y)$ is a nonzero rational constant times $J(f)^2$, so is divisible by p^8. In particular, $f(x,y)$ has a repeated root over $\mathbb{F}_p$. However, because $I(ax^4 + bx^3y + cx^2y^2) = c^2$, we see that any repeated root of $f(x,y)$ must actually be at least a triple root since $I(f) = 0$. On the other hand, if $f(x,y)$ has a quadruple root modulo p or is a multiple of p, then v reduces to an $\mathbb{F}_p$-point on a fixed codimension 2 subvariety $Z \subseteq Y$ defined over $\mathbb{Z}$ and we may apply Theorem 4.14 (with $B = X^{1/6+O(\delta^3)}$) to bound the number of such v.

So it remains to consider the case where $f(x,y)$ has splitting type $(1^3 1)$. By a change of basis over $\mathbb{Z}_p$ which is uniquely determined up to $(x,y) \mapsto (tx, t^{-1}y)$ mod p, we may assume that $f(x,y) = bx^3y + dxy^3 + ey^4$ with $p \nmid b$, $p \mid d$, $p \mid e$. Since $I(f) = -3bd = 0$, we see that $d = 0$. We then see that $J(f) = -27b^2e$, whence, because the discriminant of f is a multiple of $J(f)^2$, we find

that $p^4 \mid e = \mathrm{Disc}(F_2)$. Now, if $p \mid F_2$, then we may again use Theorem 4.14 as before. So we may assume that F_2 is primitive over $\mathbb{Z}_p$, and hence splits as (1^21) or (1^3) over $\mathbb{F}_p$.

If F_2 has splitting type (1^21), then, after a second (independent) change of basis over $\mathbb{Z}_p$, we have $F_2(X,Y) = r_6X^2Y + r_8Y^3$ with $p \nmid r_6$ and $p \mid r_8$. Since $0 = A_1(v) \equiv -r_6r_3 \pmod{p^4}$, it follows that $p^4 \mid r_3$. We also compute $0 = d \equiv 4r_4r_6^3 \pmod{p^4}$, so that $p^4 \mid r_4$. Hence

$$xF_1(X,Y) - yF_2(X,Y) \equiv X^2(r_1xX + (r_2x - r_6y)Y) \pmod{p}$$

and so $f(x,y) = \mathrm{Disc}(xF_1(X,Y) - yF_2(X,Y)) \equiv 0 \pmod{p}$, which is not of type (1^31), a contradiction.

The final case is when $F_2(X,Y) = r_5X^3 + r_6X^2Y + r_7XY^2 + r_8Y^3$ is (1^3) at p. After a $\mathbb{Z}_p$-change of basis, we may assume $p \nmid r_5$, $r_6 = 0$, and $p \mid r_7, r_8$. That $p^4 \mid \mathrm{Disc}(F_2)$ means that $p^2 \mid r_8$ and thus $p^2 \mid r_7$. Then $A_1(v) = 0$ implies that $p^2 \mid r_4$. Since f is primitive and $F_2(X,Y) \equiv r_5X^3 \pmod{p}$, we have $p \nmid r_3$. We may again change basis over $\mathbb{Z}_p$ to ensure that moreover $r_2 = 0$. (This determines the mod p reduction of our basis over $\mathbb{Z}_p$ up to transformations of the form $(X,Y) \mapsto (tX, t^{-1}Y)$.) Since $0 = a = \mathrm{Disc}(F_1) \equiv 4r_1r_3^3 \pmod{p^4}$, it follows that $p^4 \mid r_1$. Now let $\tau(v) = \gamma v$, where $\gamma = (\mathrm{diag}(\sqrt{p}, 1/\sqrt{p}), \mathrm{diag}(\sqrt{p}, 1/\sqrt{p}))$. Explicitly, the action of γ is given by

$$(F_1, F_2) \mapsto (\sqrt{p}F_1(\sqrt{p}X, Y/\sqrt{p}), F_2(\sqrt{p}X, Y/\sqrt{p})/\sqrt{p}),$$

or in terms of coefficients: $(r_1, \dots, r_8) \mapsto (p^2r_1, pr_2, r_3, r_4/p, pr_5, r_6, r_7/p, r_8/p^2)$. The pair $\tau(v)$ has the same invariants, but the associated binary quartic has been replaced by $(p^2a, pb, c, d/p, e/p^2)$, which is divisible by p, and indeed is p times a (1^31) binary quartic. Note also that

$$(p^2r_1, pr_2, r_3, r_4/p, pr_5, r_6, r_7/p, r_8/p^2) \equiv (0, 0, r_3, 0, 0, 0, 0, r_8/p^2) \pmod{p},$$

so that the first binary cubic in this pair is also of type (1^21).

Now note that the above argument did not rely on the exact vanishing of a, c, d, r_6, etc., but rather needed only that they be divisible by, e.g., p^{10}. Thus (by mod-p^{10} weak approximation), to produce this transformation, we may have first chosen changes of basis over $\mathbb{Z}$ rather than over $\mathbb{Z}_p$.

We claim that the association above $v \mapsto \tau(v)$ gives a well-defined A_3-preserving injection from $G(\mathbb{Z})$-orbits of $v \in Y(\mathbb{Z})$ with f of splitting type (1^31) and F_2 of type (1^3) to a set of $G(\mathbb{Z})$-orbits $Y(\mathbb{Z})$ with binary quartic equal to p times a (1^31) form and F_1 a (1^21) form. To see that τ is well-defined at the level of $G(\mathbb{Z})$-orbits, note that all choices made above in changing basis over $\mathbb{Z}$ differ by elements of $G(\mathbb{Z})$ that are congruent to a diagonal matrix mod p. Since, for $t \in \mathbb{Z}_p^\times$,

$$\mathrm{diag}(\sqrt{p}, 1/\sqrt{p})^{-1} \cdot (\mathrm{diag}(t, t^{-1}) + p \cdot M_2(\mathbb{Z}_p)) \cdot \mathrm{diag}(\sqrt{p}, 1/\sqrt{p}) \subseteq \mathrm{GL}_2(\mathbb{Z}_p), \tag{4.13}$$

these give the same $G(\mathbb{Z})$-orbit (integrality at p is the only thing to be checked). To see that τ is injective, suppose we have v and v' such that $\tau(v') = g \cdot \tau(v)$ with $g = (g_1, g_2) \in G(\mathbb{Z})$. Because of the uniqueness up to $(x,y) \mapsto (tx, t^{-1}y)$ of the change of basis over $\mathbb{F}_p$ taking a (1^31) mod-p binary quartic (namely the binary quartic of $\tau(v)$ divided by p) to a multiple of $x^3y \pmod{p}$, it follows that $g_2 \pmod{p}$ is diagonal, whence by (4.13) we may assume without loss of generality (by changing v' by an element of $\{\mathrm{id}\} \times \mathrm{SL}_2(\mathbb{Z})$) that $g_2 = \mathrm{id}$ and hence $f = \widetilde{f}$. Similarly, because of (4.13) and the uniqueness up to $(X,Y) \mapsto (tX, t^{-1}Y)$ of the change of basis over $\mathbb{F}_p$ taking a (1^21) mod-p binary cubic (namely the first binary cubic of $\tau(v)$) to a multiple of $XY^2 \pmod{p}$, it follows that $g_1 \pmod{p}$ is diagonal, whence by (4.13) we conclude that $v' \in G(\mathbb{Z}) \cdot v$, and so τ is indeed injective at the level of $G(\mathbb{Z})$-orbits.

The desired result now follows, since we have already bounded the set of $v \in Y(\mathbb{Z})$ for which f is a multiple of p, $p^2 \mid A_3(v)$, and $|A_3(v)| < X$. $\square$

4.3 Proof of the main counting theorem

We first state the following weighted version of Theorem 4.2.

Theorem 4.16. *Let $\varphi_p : Y(\mathbb{Z}_p) \to \mathbb{R}$ be locally constant, $G(\mathbb{Z})$-invariant, and such that $\varphi_p = 1$ for all but finitely many p. Let $N_\varphi(Y(\mathbb{Z}); X)$ denote the weighted number of irreducible $G(\mathbb{Z})$-orbits in $Y(\mathbb{Z})$ having $|A_3|$ bounded by X, where each orbit $G(\mathbb{Z}) \cdot y$ is counted with weight $\varphi(y) := \prod_p \varphi_p(y)$. Then*

$$N_\varphi(Y(\mathbb{Z}); X) = X \cdot \int_{\substack{y \in G(\mathbb{Z}) \backslash Y(\mathbb{R}) \\ |A_3(y)| < 1}} dy \cdot \prod_p \int_{y \in Y(\mathbb{Z}_p)} \varphi_p(y)\, dy + O_\varphi\big(X^{1-\Omega(1)}\big). \tag{4.14}$$

Proof. We apply Theorem 4.2 to the level sets of φ. □

To prove Theorem 4.1, we must extend Theorem 4.16 to weight functions $\varphi = \prod \varphi_p$ that are acceptable but nontrivial at infinitely many primes. It is for this extension that the uniformity estimate in Proposition 4.15 is needed.

Proof of Theorem 4.1. We may repeat the argument in [11, §2.7], but with the uniformity estimate [11, Theorem 2.13] there replaced by Proposition 4.15. □

5 The average size of the 2-Selmer group in a cubic twist family

Let (A, L) be a polarized abelian variety over $\mathbb{Q}$ with a μ_3-action, as in Section 3.2. For each integer $n \geq 1$, let (A_n, L_n) be the corresponding cubic twist and let $\lambda_n \colon A_n \to \widehat{A}_n$ be the corresponding polarization. Recall that we assume

(1) The μ_3-action has isolated fixed points (which is automatic if A is simple), and

(2) $[-1]^* L \simeq L$, and

(3) $\dim_{\mathbb{Q}} H^0(A, L) = 2$, so that each λ_n is a $(2, 2)$-isogeny.

Definition 5.1. A subset $\Sigma \subset \mathbb{Z}$ is *defined by congruence conditions* if there are open subsets $\Sigma_p \subset \mathbb{Z}_p$ such that $\Sigma = \bigcap_p \Sigma_p$.[11]

Definition 5.2. A subset $\Sigma = \bigcap_p \Sigma_p \subset \mathbb{Z}$ defined by congruence conditions is *acceptable* if

(1) each Σ_p is nonempty and open, and

(2) for all but finitely many primes p, the set Σ_p contains all $n \in \mathbb{Z}_p$ with $v_p(n) \leq 1$.

The following theorem gives Theorem 1.3 as a special case.

Theorem 5.3. *Let $\Sigma \subset \mathbb{Z}$ be acceptable. Then* $\operatorname{avg}_{n \in \Sigma} \#\mathrm{Sel}_{\lambda_n}(A_n) = 3$.

Proof. We fix some $d \in \mathbb{Z}$ such that $\mathrm{Stab}_G(v_d) \simeq A[\lambda]$, as in Lemma 3.6. Note that we are free to replace d by dt^3, for any nonzero $t \in \mathbb{Z}$. Recall from Section 3.3 the notion of a locally soluble $v \in Y(\mathbb{Q}_p)$, relative to (A, L). By Theorems 3.18 and 3.19, we may choose d so that for every nonzero $n \in \mathbb{Z}$, there is a bijection between the Selmer group $\mathrm{Sel}_{\lambda_n}(A_n)$ and the $G(\mathbb{Q})$-orbits of locally soluble elements $v \in Y(\mathbb{Z})$ with $A_3(v) = dn$.

[11] More precisely, $\Sigma = \mathbb{Z} \cap \bigcap_p \Sigma_p \subseteq \widehat{\mathbb{Z}}$.

To compute $\mathrm{avg}_{n\in\Sigma}\#\mathrm{Sel}_{\lambda_n}(A_n)$, it is therefore enough to estimate the function $N_{\widetilde{\varphi}}(Y(\mathbb{Z});X)$, where $\widetilde{\varphi}\colon Y(\mathbb{Z})\to[0,1]$ is defined as follows. For $v\in Y(\mathbb{Z})$, let $\widetilde{m}(v)$ be the number of $G(\mathbb{Z})$-orbits in the $G(\mathbb{Q})$-orbit of v. Then we define $\widetilde{\varphi}(v)=1/\widetilde{m}(v)$ if v is everywhere locally soluble and $A_3(v)\in d\Sigma$, otherwise $\widetilde{\varphi}(v)=0$.

As is usual in these types of arguments, it is more convenient to replace $\widetilde{\varphi}$ with the slightly different function $\varphi(v)$ which is defined in the same way except we replace $\widetilde{m}(v)$ with

$$m(v)=\sum_{v'\in O(v)}\frac{\#\,\mathrm{Stab}_{G(\mathbb{Q})}(v)}{\#\,\mathrm{Stab}_{G(\mathbb{Z})}(v')},$$

where $O(v)$ is a set of representatives for the $G(\mathbb{Z})$-orbits in the same $G(\mathbb{Q})$-orbit as y. Notice that $m(v)=\widetilde{m}(v)$ whenever $\mathrm{Stab}_{G(\mathbb{Q})}(v)$ is trivial. This switch is therefore justified by the fact that the number of everywhere locally soluble $G(\mathbb{Q})$-orbits on $Y(\mathbb{Z})$ with $|A_3(v)|<X$ and nontrivial stabilizer $\mathrm{Stab}_{G(\mathbb{Q})}(v)$ is $O(X^{1/3})$. Indeed, if $A_3(v)=dn$, then $\mathrm{Stab}_{G(\mathbb{Q})}(v)=\mathrm{Stab}_{G(\mathbb{Q})}(v_{dn})\simeq E_{dn}[2](\mathbb{Q})$ which is nontrivial if and only if d^2n^2 is a cube in $\mathbb{Q}^\times$. There are $O(X^{1/3})$ such values of n and the number of locally soluble $G(\mathbb{Q})$-orbits is the same for each one (since the corresponding elliptic curves are isomorphic). Thus the total number of such orbits is $O(X^{1/3})$ and will be negligible when we average over $n\in\Sigma$ with $|n|<X$.

To invoke our general counting result Theorem 4.1, we must first show that φ is an acceptable function defined by congruence conditions. We have $m(v)=\prod_p m_p(v)$ where

$$m_p(v)=\sum_{v'\in O_p(v)}\frac{\#\,\mathrm{Stab}_{G(\mathbb{Q}_p)}(v)}{\#\,\mathrm{Stab}_{G(\mathbb{Z}_p)}(v')},$$

where $O_p(v)$ is a set of representatives for the $G(\mathbb{Z}_p)$-orbits in the same $G(\mathbb{Q}_p)$-orbit as y. The proof is as in [11, Prop. 3.6], using the fact that G has class number 1. From this expression we see that φ is defined by congruence conditions. The acceptability of φ follows from Proposition 3.20 and the acceptability of the set Σ.

Thus, by Theorem 4.1, we have

$$N_\varphi(Y(\mathbb{Z});X)=X\cdot\frac{1}{2}\int_{\substack{y\in G(\mathbb{Z})\backslash Y(\mathbb{R})\\ |A_3(y)|<1}}dy\prod_p\int_{y\in Y(\mathbb{Z}_p)}\varphi_p(y)\,dy\;+\;O\big(X^{1-\Omega(1)}\big).\tag{5.1}$$

By Proposition 4.10, we have

$$\int_{y\in Y(\mathbb{Z}_p)}\varphi_p(y)dy=|\mathcal{J}|_p\cdot\mathrm{Vol}(G(\mathbb{Z}_p))\int_{n\in\Sigma_p}\sum_{\sigma\in\widehat{A}_n(\mathbb{Q}_p)/\lambda_n(A_n(\mathbb{Q}_p))}\frac{1}{\#A_n[\lambda_n](\mathbb{Q}_p)}dn,\tag{5.2}$$

using the bijection between locally soluble orbits with $A_3(y)=dn$ and the group $\widehat{A}_n(\mathbb{Q}_p)/\lambda_n(A_n(\mathbb{Q}_p))$, as well as the isomorphism $\mathrm{Stab}_{G(\mathbb{Q}_p)}(v_m)\simeq A_n[\lambda_n](\mathbb{Q}_p)$ of Theorem 3.8.

Combining (5.1) and (5.2), we obtain

$$N_\varphi(Y(\mathbb{Z});X)=|\mathcal{J}|\mathrm{Vol}(G(\mathbb{Z})\backslash G(\mathbb{R}))\prod_p\left(|\mathcal{J}|_p\mathrm{Vol}(G(\mathbb{Z}_p))\int_{n\in\Sigma_p}c_p(\lambda_n)dn\right)X+O\big(X^{1-\Omega(1)}\big)$$

where

$$c_p(\lambda_n):=\frac{\#\widehat{A}_n(\mathbb{Q}_p)/\lambda_n(A_n(\mathbb{Q}_p))}{\#A_n[\lambda_n](\mathbb{Q}_p)}.$$

For finite $p \neq 2$ we have $c_p(\lambda_n) = c_p(\widehat{A}_n)/c_p(A_n) = 1$ by [40, Prop. 3.1] and [29, Prop. 4.3]. In fact, since $\widehat{\lambda}_n = \lambda_n$, we have the more general formula

$$c_p(\lambda_n) = |2|_p^{-1} \tag{5.3}$$

for all $p \leq \infty$ [28, Lem. 7.1].

It follows that

$$\begin{aligned} N_\varphi(Y(\mathbb{Z}); X) &= |\mathcal{J}|\mathrm{Vol}(G(\mathbb{Z})\backslash G(\mathbb{R}))X \prod_p \left(|\mathcal{J}/2|_p \cdot \mathrm{Vol}(G(\mathbb{Z}_p)) \cdot \mathrm{Vol}(\Sigma_p)\right) X + O\big(X^{1-\Omega(1)}\big) \\ &= 2 \cdot \mathrm{Vol}(\Sigma) \cdot X \cdot \mathrm{Vol}(G(\mathbb{Z})\backslash G(\mathbb{R})) \prod_p \mathrm{Vol}(G(\mathbb{Z}_p)) + O\big(X^{1-\Omega(1)}\big) \\ &= 4 \cdot \mathrm{Vol}(\Sigma) \cdot X + O\big(X^{1-\Omega(1)}\big), \end{aligned}$$

where $\mathrm{Vol}(\Sigma)$ is the natural density of $\Sigma \subset \mathbb{Z}$, and we have used that the Tamagawa number of G is 2. We conclude that

$$\mathrm{avg}_{n\in\Sigma} \#\mathrm{Sel}_{\lambda_n}(A_n) = 1 + \lim_{X\to\infty} \frac{N_\varphi(Y(\mathbb{Z}); X)}{\mathrm{Vol}(\Sigma \cap [-X, X])} = 1 + 2 = 3,$$

as desired. □

We will also require the following variant of Theorem 5.3, where we impose additional local conditions on the Selmer elements.

Theorem 5.4. *Fix a prime p and suppose $\Sigma \subset \mathbb{Z}$ is an acceptable subset such that $\widehat{A}_n(\mathbb{Q}_p)/\lambda_n A_n(\mathbb{Q}_p)$ has constant size 2^k for all $n \in \Sigma$. Let $\mathrm{Sel}_s(A_n) \subset \mathrm{Sel}_{\lambda_n}(A_n)$ be the subgroup of Selmer elements which are locally trivial at p. Then $\mathrm{avg}_{n\in\Sigma} \#\mathrm{Sel}_s(A_n) = 1 + 2^{1-k}$.*

Proof. The proof is the same as in Theorem 5.3, except we tweak the local weight function φ_p so that $\varphi_p(y) = 0$ unless y is in the reducible $G(\mathbb{Q}_p)$-orbit with A_3-invariant $A_3(y)$, which corresponds to the identity element of $\widehat{A}_n(\mathbb{Q}_p)/\lambda_n A_n(\mathbb{Q}_p)$ under the bijection of Theorem 3.8. Since $\widehat{A}_n(\mathbb{Q}_p)/\lambda_n A_n(\mathbb{Q}_p)$ has size 2^k for all $n \in \Sigma$, this has the effect of multiplying the Euler factor at p by 2^{-k}, and leaving all other Euler factors the same. So the proof gives $\mathrm{avg}_{n\in\Sigma} \#\mathrm{Sel}_s(A_n) = 1 + 2 \cdot 2^{-k} = 1 + 2^{1-k}$. □

Using similar arguments, one can prove a more general equidistribution theorem as in [13, Thm. 9], but for the applications in this paper, Theorems 5.3 and 5.4 will suffice.

6 The root number in any cubic twist family is equidistributed

Let d and n be nonzero integers, and let $E_{d,n}\colon y^2 = x^3 + dn^2$. Write $w_{d,n} \in \{\pm 1\}$ for the root number of $E_{d,n}$; thus the functional equation for the completed L-function of $E_{d,n}$ reads

$$L(E_{d,n}, s) = w_{d,n} \cdot L(E_{d,n}, 2 - s).$$

The purpose of this section is to prove Theorems 2.6 and 2.7.

To prove these theorems we make use of known explicit formulas for the roots numbers $w_{d,n}$. For each $0 \neq d \in \mathbb{Z}$, let $f_d : \mathbb{Z}^+ \to \{\pm 1\}$ be the multiplicative function such that $f_d(p^k) = 1$ for all primes $p \mid 6d$ and $k \in \mathbb{N}$, such that $f_d(p^k) = f_d(p^{k-3})$ for all p and $k \geq 3$, and such that $f(p^2) = f(p) = \chi_{-3}(p)$, for all primes $p \nmid 6d$, where $\chi_{-3}(p) = \left(\frac{-3}{p}\right)$.

The following is a corollary of Várilly-Alvarado's [47, Prop. 4.4], drawing on formulas of Rohrlich.

Proposition 6.1. *Let $0 \neq d \in \mathbb{Z}$. Then there is a function $g_d : (\mathbb{Z}/9)^\times \times (\mathbb{Z}/3)^{\omega(6d)} \to \{\pm 1\}$ such that*

$$w_{d,n} = g_d\left(\left(\frac{n}{3^{v_3(n)}}\right)^2 \bmod 9, (v_p(n) \bmod 3)_{p|6d}\right) \cdot f_d(n)$$

for all $n \in \mathbb{Z}^+$.

Proposition 6.1 shows that if we ignore a factor coming from primes dividing $6d$, the root number of $E_{d,n}$ agrees with the evaluation of the multiplicative function f_d at n. As a corollary, we deduce the following result, which is a more precise version of Theorem 2.7.

Corollary 6.2. *Let $\Sigma \subseteq \mathbb{Z}^+$, and let $\gamma \in \{\pm 1\}$. For each $(s, \nu, a) \in \mathbb{Z}^+ \times \mathbb{Z}^+ \times (\mathbb{Z}/9)^\times$ with $s \mid (6d)^\infty$ and ν squarefull and coprime to $6d$, there is an $\varepsilon_{(s,\nu,a)} \in \{\pm 1\}$ such that*

$$\{n \in \Sigma \mid w_{d,n} = \gamma\} = \bigsqcup_{\substack{(s,\nu,a) \in \mathbb{Z}^+ \times \mathbb{Z}^+ \times (\mathbb{Z}/9)^\times: \\ s \mid (6d)^\infty, \\ (\nu, 6d) = 1, \\ \nu \text{ squarefull}}} \left\{ s \cdot \nu \cdot t \;\middle|\; \begin{array}{c} s\cdot\nu\cdot t \in \Sigma, \\ t \text{ squarefree}, \\ t^2 \equiv a \pmod 9, \\ (t, 6d\cdot\nu) = 1, \\ t \equiv \gamma \cdot \varepsilon_{(s,\nu,a)} \pmod 3 \end{array} \right\}.$$

Proof. For each such (s, ν, a), let

$$\varepsilon_{(s,\nu,a)} := g_d\left(\left(\frac{s}{3^{v_3(s)}}\right)^2 \cdot \nu^2 \cdot a, (v_p(s) \bmod 3)_{p|6d}\right) \cdot f_d(\nu).$$

Now, each $n \in \Sigma$ can be written uniquely as $n = s\nu t$ where νt is prime to $6d$, and t (resp. ν) is the "squarefree part" (resp. "squarefull part") of νt. In particular $(t, 6d\nu) = 1$. Setting $a := t^2 \pmod 9$, Proposition 6.1 gives

$$w_{d,n} = g_d\left(\left(\frac{s}{3^{v_3(s)}}\right)^2 \nu^2 a \bmod 9, (v_p(s) \bmod 3)_{p|6d}\right) f_d(\nu) f_d(t) = \varepsilon_{(s,\nu,a)} f_d(t).$$

This completes the proof, since $f_d(t) = \chi_{-3}(t) \equiv t \pmod 3$ for t squarefree and prime to $6d$. □

We may now deduce the following special case of Theorem 2.6 (with much better error term).

Remark 6.3. For the remainder of this section, we restrict without loss of generality to $n \in \mathbb{Z}^+$.

Theorem 6.4. *Let $m \in \mathbb{Z}^+$, and let $r \in \mathbb{Z}/m$. Then there is a constant $c_{d,r} \in \mathbb{R}$ such that*

$$\sum_{n \leq X : n \equiv r \pmod m} w_{d,n} = c_{d,r} \cdot X + O(d^{O(1)} m^{O(1)} X^{1-\Omega(1)}).$$

Moreover, if $3 \nmid m$, then $c_{d,r} = 0$. In particular, for d fixed and $n \to \infty$ over all of $\mathbb{Z}^+$ (or any arithmetic progression with common difference not divisible by 3), the root numbers of $E_{d,n}$ uniformly distribute in $\{\pm 1\}$.

The constant $c_{d,r}$ is easy enough to determine but we will not specify it here.

Proof. For notational convenience we will only treat the case of $3 \nmid m$—from the argument it will be clear how to proceed when $3 \mid m$.[12]

Of course

$$\sum_{n \leq X : n \equiv r \pmod m} w_{d,n} = \sum_{\gamma \in \{\pm 1\}} \gamma \sum_{n \leq X : n \equiv r \pmod m, w_{d,n} = \gamma} 1.$$

[12]The point is that when $3 \mid m$ then exactly one of the conditions $t \equiv \pm\varepsilon_{(s,\nu,a)} \pmod 3$ may contradict the congruence $s\nu t \equiv r \pmod 3$, producing a term of order X because the sum over t corresponding to $\gamma = +1$ no longer cancels the sum corresponding to $\gamma = -1$ to leading order.

By Corollary 6.2, it follows that

$$\begin{aligned}
&\sum_{\gamma\in\{\pm1\}}\gamma\sum_{n\le X:n\equiv r\,(\mathrm{mod}\ m),w_{d,n}=\gamma}1\\
&=\sum_{a\in(\mathbb{Z}/9)^\times}\sum_{s\le X:s|(6d)^\infty}\sum_{\substack{\nu\le\frac{X}{s}:\\(\nu,6d)=1,\\(6d\cdot\nu,r,\frac{m}{(m,s\nu)})=1,\\\nu\text{ squarefull}}}\sum_{\gamma\in\{\pm1\}}\gamma\sum_{\substack{t\le\frac{X}{s\cdot\nu}:\\(t,6d\cdot\nu)=1,\\t^2\equiv a\,(\mathrm{mod}\ 9),\\t\equiv\gamma\cdot\varepsilon_{(s,\nu,a)}\,(\mathrm{mod}\ 3),\\s\nu t\equiv r\,(\mathrm{mod}\ m)}}\mu^2(t)\\
&=\sum_{a\in(\mathbb{Z}/9)^\times}\sum_{s\le X^\delta:s|(6d)^\infty}\sum_{\substack{\nu\le X^\delta:\\(\nu,6d)=1,\\(6d\cdot\nu,r,\frac{m}{(m,s\nu)})=1,\\\nu\text{ squarefull}}}\sum_{\gamma\in\{\pm1\}}\gamma\sum_{\substack{t\le\frac{X}{s\cdot\nu}:\\(t,6d\cdot\nu)=1,\\t^2\equiv a\,(\mathrm{mod}\ 9),\\t\equiv\gamma\cdot\varepsilon_{(s,\nu,a)}\,(\mathrm{mod}\ 3),\\s\nu t\equiv r\,(\mathrm{mod}\ m)}}\mu^2(t)+O(X^{1-\Omega(1)})\\
&=\sum_{a\in(\mathbb{Z}/9)^\times}\sum_{s\le X^\delta:s|(6d)^\infty}\sum_{\substack{\nu\le X^\delta:\\(\nu,6d)=1,\\(6d\cdot\nu,r,\frac{m}{(m,s\nu)})=1,\\\nu\text{ squarefull}}}\sum_{\gamma\in\{\pm1\}}\gamma\sum_{\substack{t\le\frac{X}{s\cdot\nu}:\\(t,6d\cdot\nu)=1,\\t^2\equiv a\,(\mathrm{mod}\ 9),\\t\equiv\gamma\cdot\varepsilon_{(s,\nu,a)}\,(\mathrm{mod}\ 3),\\t\equiv r\cdot\left(\frac{s\nu}{(m,s\nu)}\right)^{-1}\,(\mathrm{mod}\ \frac{m}{(m,s\nu)})}}\mu^2(t)+O(X^{1-\Omega(1)}).
\end{aligned}$$

Note that the conditions $(t,6d\cdot\nu)=1, t^2\equiv a\,(\mathrm{mod}\ 9), t\equiv\gamma\cdot\varepsilon_{(s,\nu,a)}\,(\mathrm{mod}\ 3), t\equiv r\cdot\left(\frac{s\nu}{(m,s\nu)}\right)^{-1}\,(\mathrm{mod}\ \frac{m}{(m,s\nu)})$ can be written as at most $\ll dX^\delta$ congruences modulo $[9,6d\cdot\nu,m]\ll dmX^\delta$.

Now, for $x\in\mathbb{Z}/y$, if (x,y) is not squarefree, then there are no squarefree $n\equiv x\,(\mathrm{mod}\ y)$. Otherwise,

$$\begin{aligned}
\sum_{n\le X:n\equiv x\,(\mathrm{mod}\ y)}\mu^2(n)&=\sum_{n\le X:n\equiv x\,(\mathrm{mod}\ y)}\sum_{d^2|n}\mu(d)\\
&=\sum_{d\le\sqrt{X}:(y,d^2)|(x,y)}\mu(d)\sum_{e\le\frac{X}{d^2}:\frac{d^2}{(y,d^2)}\cdot e\equiv x\,(\mathrm{mod}\ \frac{y}{(y,d^2)})}1\\
&=\sum_{f|(x,y):(f,\frac{y}{f})=1}\mu(f)\sum_{g\le\sqrt{\frac{X}{f}}:(g,y)=1}\mu(g)\sum_{e\le\frac{X}{f^2g^2}:e\equiv\frac{x}{f}\cdot f^{-1}g^{-2}\,(\mathrm{mod}\ \frac{y}{f})}1\\
&=\sum_{f|(x,y):(f,\frac{y}{f})=1}\mu(f)\sum_{g\le\sqrt{\frac{X}{f}}:(g,y)=1}\mu(g)\left(\frac{X}{fg^2y}+O(1)\right)\\
&=\frac{X}{y}\sum_{f|(x,y):(f,\frac{y}{f})=1}\frac{\mu(f)}{f}\sum_{g\le\sqrt{\frac{X}{f}}:(g,y)=1}\frac{\mu(g)}{g^2}+O\left(O(1)^{\frac{\sqrt{\log y}}{\log\log y}}\sqrt{X}+O(1)^{\frac{\log y}{\log\log y}}\right)\\
&=\frac{X}{y}\sum_{f|(x,y):(f,\frac{y}{f})=1}\frac{\mu(f)}{f}\sum_{g\ge1:(g,y)=1}\frac{\mu(g)}{g^2}+O\left(O(1)^{\frac{\sqrt{\log y}}{\log\log y}}\sqrt{X}+O(1)^{\frac{\log y}{\log\log y}}\right)\\
&=\frac{6}{\pi^2}\cdot\frac{X}{y}\cdot\prod_{p|(x,y):p^2\nmid y}\left(1-\frac{1}{p}\right)\cdot\prod_{p|y}\left(1-\frac{1}{p^2}\right)^{-1}+O\left(O(1)^{\frac{\sqrt{\log y}}{\log\log y}}\sqrt{X}+O(1)^{\frac{\log y}{\log\log y}}\right)
\end{aligned}$$

(of course one can be more precise). Consequently (since the leading terms match—this is where we use that $3 \nmid m$),

$$\sum_{\gamma\in\{\pm1\}} \gamma \sum_{\substack{t\le \frac{X}{s\nu}:\\ (t,6d\nu)=1,\\ t^2\equiv a \pmod 9,\\ t\equiv \gamma\varepsilon_{(s,\nu,a)} \pmod 3,\\ t\equiv r\left(\frac{s\nu}{(m,s\nu)}\right)^{-1} \pmod{\frac{m}{(m,s\nu)}}}} \mu^2(t) = \sum_{\substack{t\le \frac{X}{s\nu}:\\ (t,6d\nu)=1,\\ t^2\equiv a \pmod 9,\\ t\equiv \varepsilon_{(s,\nu,a)} \pmod 3,\\ t\equiv r\left(\frac{s\nu}{(m,s\nu)}\right)^{-1} \pmod{\frac{m}{(m,s\nu)}}}} \mu^2(t) - \sum_{\substack{t\le \frac{X}{s\nu}:\\ (t,6d\nu)=1,\\ t^2\equiv a \pmod 9,\\ t\equiv -\varepsilon_{(s,\nu,a)} \pmod 3,\\ t\equiv r\left(\frac{s\nu}{(m,s\nu)}\right)^{-1} \pmod{\frac{m}{(m,s\nu)}}}} \mu^2(t) \ll d^{O(1)}m^{O(1)}X^{1/2+O(\delta)}.$$

Summing this over $s,\nu \le X^\delta$ and $a \in (\mathbb{Z}/9)^\times$, we conclude that

$$\sum_{n\le X: n\equiv r \pmod m} w_{d,n} \ll d^{O(1)}m^{O(1)}X^{1/2+O(\delta)} + X^{1-\Omega(1)},$$

as desired. □

We may now prove Theorem 2.6.

Proof of Theorem 2.6. Let $\varepsilon > 0$. Let $m \in \mathbb{Z}^+$ and $A \subseteq \mathbb{Z}/m$ be such that $3 \nmid m$ and the symmetric difference $\{n : n \pmod m \in A\}\Delta\Sigma$ has density $\le \varepsilon$ (we may use e.g. $m := \prod_{\substack{p\le T,\\ p\ne 3}} p^T$ with $T \gg_{\Sigma,\varepsilon} 1$). Thus

$$\sum_{n\in\Sigma: n\le X} w_{d,n} = \sum_{r\in A} \sum_{n\le X: n\equiv r \pmod m} w_{d,n} + O_S(\varepsilon\cdot X).$$

We conclude by applying Theorem 6.4 and taking $\varepsilon \to 0$ sufficiently slowly with X. □

7 Cubic twists having ranks 0 and 1

Fix a nonzero $d \in \mathbb{Z}$, and let $E_{d,n}\colon y^2 = x^3 + dn^2$ be the corresponding cubic twist family of elliptic curves, with $n \in \mathbb{Z}$ varying. Theorem 1.3 does not quite imply that a positive proportion of twists $E_{d,n}$ have 2-Selmer rank 0 (and hence Mordell–Weil rank 0); for example, it is consistent with the possibility that asymptotically half of the curves $E_{d,n}$ satisfy $\#\mathrm{Sel}_2(E_{d,n}) = 2$ and half satisfy $\#\mathrm{Sel}_2(E_{d,n}) = 4$. This hypothetical distribution is also consistent with the fact that the parity of $\dim_{\mathbb{F}_2} \mathrm{Sel}_2(E_{d,n})$ is equidistributed in these families.

In this section, we apply the results on root numbers from the previous section and the p-parity theorem to prove the existence of twists having ranks 0 and 1, respectively.

Theorem 7.1. *Fix a nonzero integer d. Then the average size of the 2-Selmer group of those elliptic curves in $E_{d,n}$ having root number $+1$ (resp. -1) is 3.*

Proof. This follows from Theorem 2.7, or rather, its more precise version Corollary 6.2, which expresses the set of elliptic curves of root number $+1$ (resp. -1) as the union of acceptable families. The average size of $\mathrm{Sel}_2(E_{d,n})$ is 3 on each such acceptable family. Using the uniformity estimate Proposition 4.15, one shows as in [6, §6.4] that the average is still 3 when we average over the union of all of these families as well. □

Theorems 2.6 and 7.1 together give Theorem 1.4. We now prove the existence of many curves in any cubic twist family $E_{d,n}$ having 2-Selmer rank 0 and 2-Selmer rank 1.

Theorem 7.2. *Fix an integer $d \neq 0$ and let $E_{d,n}\colon y^2 = x^3 + dn^2$ be the corresponding family of cubic twists. The proportion of n such that $\mathrm{Sel}_2(E_{d,n}) = 0$ is at least $1/6$, and the proportion of n such that $\dim_{\mathbb{F}_2} \mathrm{Sel}_2(E_{d,n}) = 1$ is at least $5/12$.*

Proof. For $w \in \{\pm 1\}$, let $\Sigma(w)$ be the set of integers n such that $E_{d,n}$ has root number w. By the p-parity Theorem [17], the parity of the $\mathbb{F}_2$-rank of $\mathrm{Sel}_2(E_{d,n})$ is constant on $\Sigma(w)$ and is even if and only if $w = 1$. If the parity is even, then at least $\frac{1}{3}$ of $m \in \Sigma(w)$ have $\mathbb{F}_2$-rank 0, as otherwise the average size of $\mathrm{Sel}_2(E_{d,n})$ would be larger than $\frac{2}{3} \cdot 4 + \frac{1}{3} \cdot 1 = 3$. Similarly, if the parity is odd, then at least $\frac{5}{6}$ of $m \in \Sigma(w)$ have $\mathbb{F}_2$-rank equal to 1, as otherwise the average size of $\mathrm{Sel}_2(E_{d,n})$ would be larger than $\frac{1}{6} \cdot 8 + \frac{5}{6} \cdot 2 = 3$. □

For any elliptic curve $E/\mathbb{Q}$, if $\mathrm{Sel}_2(E) = 0$, then $\mathrm{rk}\, E(\mathbb{Q}) = 0$. This follows from the usual exact sequence

$$0 \to E(\mathbb{Q})/2E(\mathbb{Q}) \to \mathrm{Sel}_2(E) \to Ш(E)[2] \to 0.$$

It is conjectured that $Ш(E)[2]$ has even $\mathbb{F}_2$-dimension; this is a consequence of the conjectural finiteness of $Ш(E)$ and properties of the Cassels-Tate pairing $Ш(E) \times Ш(E) \to \mathbb{Q}/\mathbb{Z}$. If this is the case, and if $E(\mathbb{Q})[2] = 0$, we see that $\dim_{\mathbb{F}_2} \mathrm{Sel}_2(E) = 1$ implies $\mathrm{rk}\, E(\mathbb{Q}) = 1$. Thus, the following is a corollary of Theorem 7.2.

Corollary 7.3. *At least $1/6$ of the elliptic curves $E_{d,n}$ have algebraic rank 0. If $Ш(E_{d,n})$ is finite for 100% of n, then at least $5/12$ of the elliptic curves $E_{d,n}$ have algebraic rank 1.*

We can prove an unconditional but weaker version of the second assertion of Corollary 7.3 in certain circumstances, using the p-converse theorem of Burungale and Skinner, which is Corollary A.2 of the Appendix. This result requires elliptic curves with good reduction at 2.

Remark 7.4. For a fixed d, there may not exist any cubic twists $E_{d,n}$ with good reduction at 2. By Tate's algorithm, if dn^2 is sixth-power-free, then $E_{d,n}\colon y^2 = x^3 + dn^2$ has good reduction at 2 if and only if there exists an odd integer D such that $dn^2 \equiv 16D^2 \pmod{64}$. It follows that there exist n such that $E_{d,n}$ has good reduction at 2 if and only if the 2-adic valuation $v_2(d)$ of d is even and $d = 2^{v_2(d)}D$ with $D \equiv 1 \pmod 4$. In particular, if $d = -432$, then there are many cubic twists with good reduction at 2.[13]

Theorem 7.5. *Fix a nonzero integer d. Among the elliptic curves in the cubic twist family $E_{d,n}$ that have good reduction at 2 and root number -1, a proportion of at least $1/3$ have rank 1.*

Proof. By the remark, we may assume that $\mathbb{Q}(\sqrt{d})$ is unramified at 2, otherwise there are no twists of good reduction. Let $\alpha_i(X)$ (resp. $\beta_i(X)$) denote the proportion of curves in the family $E_{d,n}$ with $|n| \leq X$ having good reduction at 2 and root number -1 such that $\#\mathrm{Sel}_2(E_{d,n}) = 2^i$ and such that $\mathrm{Sel}_2(E_{d,n})$ maps trivially (resp. nontrivially) to $E(\mathbb{Q}_2)/2E(\mathbb{Q}_2)$. Then we have already seen that $\alpha_1(X) + \beta_1(X) \geq 5/6$. Since the average size of the 2-Selmer group in this family is 3, we have

$$\sum_i 2^i \alpha_i(X) + \sum_i 2^i \beta_i(X) = 3 + o_{X\to\infty}(1). \tag{7.1}$$

Now, the elliptic curve $E_{d,n}$ has complex multiplication by the field $K = \mathbb{Q}(\sqrt{-3})$, and the prime 2 is inert in $\mathcal{O}_K$. It follows that $E_{d,n}$ has supersingular reduction at 2 and hence $E_{d,n}[2](\mathbb{Q}_2) = 0$.[14] Hence

$$\#E_{d,n}(\mathbb{Q}_2)/2E_{d,n}(\mathbb{Q}_2) = 2\#E_{d,n}[2](\mathbb{Q}_2) = 2,$$

[13] Indeed, the model $E_{-432,n}\colon x^3 + y^3 = n$ visibly has good reduction at 2 when n is an odd integer.

[14] More directly: the polynomial $x^3 + dn^2$ has no roots over $\mathbb{Q}_2$ if $dn^2 \equiv 16D^2 \pmod{64}$ with D odd.

by (5.3). If $E_{d,1}$ has good reduction at 2, and $2 \mid n$ but $v_2(n) \not\equiv 0 \pmod 3$, then $E_{d,n}$ has bad reduction at 2 since $\mathbb{Q}(\sqrt[3]{n})$ is ramified at 2. Thus, as $E_{d,n}$ varies within the subfamily of good reduction curves, all the curves are isomorphic over $\mathbb{Q}_2$ (using that $\mathbb{Z}_2^\times = \mathbb{Z}_2^{\times 3}$). By Theorem 5.4, half of all nontrivial 2-Selmer elements in this family remain nontrivial over $\mathbb{Q}_2$. Thus

$$\frac{1}{2}\sum_i 2^i \beta_i(X) \le 1 + o_{X\to\infty}(1). \tag{7.2}$$

Subtracting twice (7.2) from (7.1), we conclude that

$$\sum_i 2^i \alpha_i(X) \le 1 + o_{X\to\infty}(1). \tag{7.3}$$

In particular, $2\alpha_1(X) \le 1 + o_{X\to\infty}(1)$. Therefore, $\alpha_1(X) \le 1/2 + o_{X\to\infty}(1)$ and so $\beta_1(X) \ge 1/3 + o_{X\to\infty}(1)$. The elliptic curves whose density is given by $\beta_1(X)$ all have algebraic rank 1 by Burungale and Skinner's Corollary A.2 in the Appendix. □

Proof of Theorems 1.1 and 1.2. When $d = -432 = -2^4 \cdot 3^3$, the curve $E_{d,n}$ has good reduction at 2 whenever n has 2-adic valuation that is a multiple of 3; this set has density 4/7 (among all integers, and also among all cubefree integers). Imposing the root number -1 condition is again a density $\frac{1}{2}$ condition by Theorem 2.6. Thus, Theorem 7.5 guarantees that at least $\frac{1}{3}\cdot\frac{4}{7}\cdot\frac{1}{2} = \frac{2}{21}$ of cubic twists $E_{16,n}$ have algebraic rank 1. Together with Corollary 7.3, this gives Theorem 1.1. The proof of Theorem 1.2 is similar, using the fact that $E_{d,n}$ has good reduction at 2 if and only if $dn^2 = 2^{6k+4}D$ for integers $D \equiv 1 \pmod 4$ and $k \ge 0$. □

Proof of Theorem 1.5. Since the average size of the 2-Selmer group across the curves $E_{d,n}$ of positive root number is 3, and for even integers $r \ge 0$ we have the inequality

$$\frac{3}{2}r + 1 \le 2^r,$$

it follows that

$$\operatorname{avg}_n \operatorname{rk}(E_{d,n}) \le \operatorname{avg}_n \frac{2}{3}(2^{\dim_{\mathbb{F}_2} \operatorname{Sel}_2(E_{d,n})} - 1) = \operatorname{avg}_n \frac{2}{3}(\#\operatorname{Sel}_2(E_{d,n}) - 1) = \frac{2}{3}(3-1) = \frac{4}{3}$$

across the curves $E_{d,n}$ having root number $+1$; here we have used the fact that, by the p-parity theorem, curves with root number $+1$ have even 2-Selmer rank.

Similarly, since the average size of the 2-Selmer group across the curves $E_{d,n}$ of negative root number is 3, and for odd integers $r \ge 0$ we have the inequality

$$3r - 1 \le 2^r,$$

it follows that

$$\operatorname{avg}_n \operatorname{rk}(E_{d,n}) \le \operatorname{avg}_n \frac{1}{3}(2^{\dim_{\mathbb{F}_2} \operatorname{Sel}_2(E_{d,n})} + 1) \le \operatorname{avg}_n \frac{1}{3}(\#\operatorname{Sel}_2(E_{d,n}) + 1) = \frac{1}{3}(3+1) = \frac{4}{3}$$

across the curves $E_{d,n}$ having root number -1; here we have used the fact that curves -1 have odd 2-Selmer rank. Thus, across all curves $E_{d,n}$, the average rank is bounded above by 4/3.

For the lower bound, we simply observe that when the 2-adic valuation of d is even, then at least $\frac{1}{21}2^{r-1}$ of the curves in the family $E_{d,n}$ have rank 1. □

8 A higher-dimensional example: cubic twists of Prym surfaces

We give examples of cubic twist families of geometrically simple abelian surfaces. We then combine Theorem 5.3 with Pantazis' bigonal construction to prove Theorem 1.13.

In the proof of Theorem 3.8, we observed that if (A, L) is an abelian variety with μ_3-action, then there is an elliptic curve E and an isomorphism $\eta\colon A[\lambda] \simeq E[2]$ such that the abelian variety $B = (A \times E)/\Gamma_\eta$ is principally polarized. If A is an abelian surface, then B is a principally polarized abelian threefold. Since the Torelli map $\mathcal{M}_3 \to \mathcal{A}_3$ is birational, we might hope to realize B as a Jacobian of a curve C. To find such curves, it is natural to insist that the curve itself has an automorphism of order 3 and is bielliptic.

This leads us to consider plane quartic curves over $\mathbb{Q}$ with affine model $C\colon y^3 = x^4 + ax^2 + b$, for some $a, b \in \mathbb{Q}$. We suppose that $b(a^2 - 4b) \neq 0$, so that C is smooth. Such a curve admits a μ_3-action over $\mathbb{Q}$, generated by the order 3 automorphism $(x, y) \mapsto (x, \zeta_3 y)$. We consider the cubic twist family

$$C_n\colon ny^3 = x^4 + ax^2 + b.$$

Another model for C_n is $y^3 = x^4 + an^4x^2 + bn^8$.

Note the double cover $\pi\colon C \to E$ to the elliptic curve $E\colon y^3 = x^2 + ax + b$. The evident automorphism of order 3 means E has j-invariant 0, and indeed the short Weierstrass model for E is $y^2 = x^3 + 16(a^2 - 4b)$. More generally, the curve C_n is a double cover of $E_n\colon y^2 = x^3 + 16m^2(a^2 - 4b)$, and the latter is a cubic twist family of elliptic curves $E_{d,n}$, with $d = 16(a^2 - 4b)$ in our notation.

The involution generating $\mathrm{Aut}(C/E)$ is $\tau(x, y) = (-x, y)$, so π is ramified at the three points where $x = 0$ and the unique point ∞ at infinity. Let $J = \mathrm{Jac}(C)$ and let $A := \ker(J \to E)$.

Lemma 8.1. *The map $\pi^*\colon \mathrm{Pic}^0(E) \to \mathrm{Pic}^0(C) \simeq J$ is injective and A is an abelian variety.*

Proof. The kernel of π^* is 2-torsion, and it is nontrivial if and only if π is unramified. Since π is ramified, the map π^* is injective, and it follows by duality that the kernel of $\pi_*\colon J \to E$ is connected and hence an abelian variety. □

The abelian surface A is an example of a *Prym variety*. It may alternatively be described as the subgroup of degree zero divisor classes in J on which τ acts as -1. For generic parameters $a, b \in \mathbb{Q}$, the surface A is absolutely simple, so there is no obvious way to reduce the study of the Mordell-Weil group $A(\mathbb{Q})$ to rational points on elliptic curves.

By the Lemma, we may view $E \simeq \mathrm{Pic}^0(E)$ inside J, and it follows immediately that $E \cap A = E[2]$. This subgroup plays a role in the geometry of A, as we explain. First, let the theta divisor $\theta \in \mathrm{Div}(J)$ be the image of the map $C^{(2)} \to J$ sending an effective divisor D of degree two to $D - 2\infty$. The line bundle $\mathcal{O}_J(\theta)$ determines a principal polarization on J, and its restriction to A is an ample line bundle which we denote L_A. The corresponding polarization $\lambda\colon A \to \widehat{A}$ has degree 4 and its kernel is precisely $E[2] \subset A$; for more details on Prym varieties, see [31].

The order 3 automorphism on J preserves A and preserves the theta divisor θ as well. It follows that (A, L_A) is a polarized abelian surface with μ_3-action. In particular, the abelian variety A has cubic twists, which are simply the Prym varieties A_n attached to the curves C_n. Let L_n be the line bundle on A_n giving the degree 4 polarization $A_n \to \widehat{A}_n$ described above.

Corollary 8.2. *We have* $\mathrm{avg}_n \#\mathrm{Sel}_\lambda(A_n) = 3$ *and* $\mathrm{avg}_n \dim_{\mathbb{F}_2} \mathrm{Sel}_\lambda(A_n) \leq 1.5$.

Proof. This follows from Theorem 5.3, once we observe that $\dim_{\mathbb{Q}} H^0(A, L_A) = \sqrt{\deg(\lambda)} = 2$. □

Next, we leverage our understanding of the Selmer groups $\mathrm{Sel}_{\lambda_n}(A_n)$ to deduce information about the $\mathrm{Sel}_2(A_n)$, and hence the ranks of $A_n(\mathbb{Q})$. Let $\widetilde{\lambda}_n\colon \widehat{A}_n \to A_n$ be the isogeny (over $\mathbb{Q}$) such that $\widetilde{\lambda}_n \circ \lambda_n = [2]$. Beware that $\widetilde{\lambda}_n$ is not the dual of λ_n, as λ_n is self-dual whereas A_m is generally not. To study $\widetilde{\lambda}_n$ we use a beautiful special case of Pantazis' bigonal construction:

Proposition 8.3. *Recall $d = 16(a^2 - 4b)$. The surface $\widehat{A}$ is the Prym attached to the genus three curve $\widehat{C}\colon y^3 = x^4 + 8ax^2 + d$. Moreover, the map $\widetilde{\lambda}\colon \widehat{A} \to A$ is the natural polarization on $\widehat{A}$.*

Proof. This is a special case of [28, Thm. 3.14]. □

We may now prove Theorem 1.13.

Proof of Theorem 1.13. The average $\mathbb{F}_2$-rank of $\mathrm{Sel}_\lambda(A_n)$ is at most 1.5 by Corollary 8.2. By Proposition 8.3 and Corollary 8.2, the average $\mathbb{F}_2$-rank of $\mathrm{Sel}_{\widetilde{\lambda}}(\widehat{A}_n)$ is also at most 1.5. Since $\widetilde{\lambda} \circ \lambda = [2]$, it follows that the average $\mathbb{F}_2$-dimension of $\mathrm{Sel}_2(A_n)$ is at most $1.5 + 1.5 = 3$, and hence the average rank of A_n is at most 3. □

9 The average size of the 3-Selmer group in a cubic twist family is infinite

Proof of Theorem 1.10. Since d is fixed, we write $E_n = E_{d,n}$, and let $E'_n = E_{-3d,3n}$. There is a 3-isogeny $\varphi_n\colon E_n \to E'_n$, whose base change to $\mathbb{Q}(\sqrt{-3})$ becomes multiplication by $\sqrt{-3}$ [6]. The kernel of the natural map $\mathrm{Sel}_{\varphi_n}(E_n) \to \mathrm{Sel}_3(E_n)$ is $E'_n[\hat{\varphi}_n](\mathbb{Q})/\varphi(E_n[3](\mathbb{Q}))$, whose size is at most 3. Thus it suffices to show that the average size of $\mathrm{Sel}_{\varphi_n}(E_n)$ is unbounded as $n \to \infty$.

Combining the Greenberg–Wiles formula [34, 8.7.9] and [40, Prop. 3.1], we have

$$\#\mathrm{Sel}_{\varphi_n}(E_n) \gg_d c(E'_n)/c(E_n),$$

where $c(E) = \prod_p c_p(E)$ is the product of all the Tamagawa numbers of E. The ratios

$$c_p(E'_n)/c_p(E_n)$$

are uniformly bounded (from above and below), independent of both m and p (and even d). This is a general fact about ℓ-isogenies of abelian varieties of a given dimension, but it follows easily from Tate's algorithm in this case, especially since E has everywhere potentially good reduction. Thus, we can safely ignore finitely many primes, and we have

$$\#\mathrm{Sel}_{\varphi_n}(E_n) \gg_d \prod_{p>3d} \frac{c_p(E'_n)}{c_p(E_n)}.$$

Let $\chi = \left(\frac{d}{\cdot}\right)$ be the quadratic character cutting out the field $\mathbb{Q}(\sqrt{d})$. For $p > 3d$, we have [6, Prop. 34]:

$$\frac{c_p(E'_n)}{c_p(E_n)} = \begin{cases} 3^{-\chi(n)} & \text{if } p \equiv 2 \pmod 3 \text{ and } p \mid n, \\ 1 & \text{otherwise.} \end{cases}$$

Now let $\alpha(n)$ (resp. $\beta(n)$) be the number of primes $p \equiv 2 \pmod 3$ dividing n such that $\chi(n) = -1$ (resp. $\chi(n) = 1$). Then

$$\#\mathrm{Sel}_{\varphi_n}(E_n) \gg_d 3^{\alpha(n)-\beta(n)}. \tag{9.1}$$

To estimate the sum $\sum_{n<X} \#\mathrm{Sel}_{\varphi_n}(E_n)$, we use the following result of Selberg–Delange type.

Theorem 9.1 ([18, Prop. 4]). *Let f be a multiplicative real valued function on the natural numbers. Suppose that there exist constants u and v such that $0 \leq f(p^r) \leq ur^v$ for all primes p and all positive integers r. Suppose also that there exist real numbers $\xi > 0$ and $0 < \beta < 1$ such that*

$$\sum_{p<X} f(p) = \xi \cdot \frac{X}{\log X} + O\left(\frac{X}{(\log X)^{1+\beta}}\right)$$

as $X \to \infty$. Then there is an explicit constant C_f such that

$$\sum_{n \leq X} f(n) = C_f \cdot X(\log X)^{\xi-1} + O_f\left(X(\log X)^{\xi-1-\beta}\right)$$

as $X \to \infty$.

To apply Theorem 9.1, first consider the case that $-3d$ is not a square. Then half of primes are congruent to 2 (mod 3) and among those, half of them satisfy $\chi(m) = 1$. Thus

$$\begin{aligned}\sum_{p \leq X} 3^{\alpha(p)-\beta(p)} &= \left(\frac{1}{2} + 3 \cdot \frac{1}{4} + \frac{1}{3} \cdot \frac{1}{4}\right) X \left(\log X\right)^{-1} + O\big(X \left(\log X\right)^{-2}\big) \\ &= \frac{4}{3} \cdot X(\log X)^{-1} + O\big(X(\log X)^{-2}\big).\end{aligned}$$

It then follows from (9.1) and Theorem 9.1 that there is an explicit constant $C_d \in \mathbb{R}^+$ such that

$$\sum_{n<X} \#\mathrm{Sel}_{\varphi_n}(E_n) \gg_d \sum_{n<X} 3^{\alpha(n)-\beta(n)} = (C_d + o(1)) \cdot X \left(\log X\right)^{1/3},$$

which proves Theorem 1.10 in this case. If $-3d$ is a square, then a similar argument shows that

$$\sum_{n<X} \#\mathrm{Sel}_{\varphi_n}(E_n) \gg_d (1 + o(1)) \cdot X \log X,$$

which completes the proof. □

Remark 9.2. There is a single cubic twist family which is not covered by Theorems 1.10, namely, $E_{1,n}\colon y^2 = x^3 + n^2$. In this case, we suspect that the average size of $\mathrm{Sel}_{\varphi_n}(E_n)$ is bounded.

Remark 9.3. It seems likely that our lower bound is close to sharp. That is, aside from the two exceptional cubic twist families $E_{1,n}$ and $E_{-3,n}$, we expect that the main term will be on the order of $X(\log X)^{1/3}$.

Acknowledgements

We thank Asher Auel, Ashay Burungale, Wei Ho, Jef Laga, Arul Shankar, and Christopher Skinner for helpful conversations. We also thank Ashay Burungale and Christopher Skinner for the beautiful results proven in the Appendix below on p-converse theorems for primes p of supersingular reduction for elliptic curves with complex multiplication. The first author was supported by the National Science Foundation (grant DMS-2002109) and the Society of Fellows. The second author was supported by a Simons Investigator Grant and NSF grant DMS-1001828. The third author was supported by the Israel Science Foundation (grant No. 2301/20).

References

[1] L. Alpöge, M. Bhargava, and A. Shnidman. A positive proportion of cubic fields are not monogenic yet have no local obstruction to being so. *arXiv:2011.01186*, 2020.

[2] L. Alpöge. Quadrics in arithmetic statistics. *arXiv:2110.03947*, 2021.

[3] L. H. A. Alpöge. *Points on Curves*. ProQuest LLC, Ann Arbor, MI, 2020. Thesis (Ph.D.)–Princeton University.

[4] B. Bektemirov, B. Mazur, W. Stein, and M. Watkins. Average ranks of elliptic curves: tension between data and conjecture. *Bull. Amer. Math. Soc. (N.S.)*, 44(2):233–254, 2007.

[5] M. Bhargava. The density of discriminants of quintic rings and fields. *Ann. of Math. (2)*, 172(3):1559–1591, 2010.

[6] M. Bhargava, N. Elkies, and A. Shnidman. The average size of the 3-isogeny Selmer groups of elliptic curves $y^2 = x^3 + k$. *J. Lond. Math. Soc. (2)*, 101(1):299–327, 2020.

[7] M. Bhargava and B. H. Gross. Arithmetic invariant theory. In *Symmetry: representation theory and its applications*, volume 257 of *Progr. Math.*, pages 33–54. Birkhäuser/Springer, New York, 2014.

[8] M. Bhargava and W. Ho. Coregular spaces and genus one curves. *Camb. J. Math.*, 4(1):1–119, 2016.

[9] M. Bhargava and W. Ho. On average sizes of Selmer groups and ranks in families of elliptic curves having marked points. *arXiv:2207.03309*, 2022.

[10] M. Bhargava, Z. Klagsbrun, R. J. Lemke Oliver, and A. Shnidman. 3-isogeny Selmer groups and ranks of abelian varieties in quadratic twist families over a number field. *Duke Math. J.*, 168(15):2951–2989, 2019.

[11] M. Bhargava and A. Shankar. Binary quartic forms having bounded invariants, and the boundedness of the average rank of elliptic curves. *Ann. of Math. (2)*, 181(1):191–242, 2015.

[12] M. Bhargava and A. Shnidman. On the number of cubic orders of bounded discriminant having automorphism group C_3, and related problems. *Algebra Number Theory*, 8(1):53–88, 2014.

[13] M. Bhargava and C. Skinner. A positive proportion of elliptic curves over $\mathbb{Q}$ have rank one. *J. Ramanujan Math. Soc.*, 29(2):221–242, 2014.

[14] T. Browning and R. Heath-Brown. The geometric sieve for quadrics. *arXiv:2003.09593*, 2020.

[15] K. Česnavičius. Selmer groups as flat cohomology groups. *J. Ramanujan Math. Soc.*, 31(1):31–61, 2016.

[16] L. E. Dickson. *History of the theory of numbers. Vol. II: Diophantine analysis*. Chelsea Publishing Co., New York, 1966.

[17] T. Dokchitser and V. Dokchitser. Regulator constants and the parity conjecture. *Invent. Math.*, 178(1):23–71, 2009.

[18] S. Finch, G. Martin, and P. Sebah. Roots of unity and nullity modulo n. *Proc. Amer. Math. Soc.*, 138(8):2729–2743, 2010.

[19] D. Goldfeld. Conjectures on elliptic curves over quadratic fields. In *Number theory, Carbondale 1979 (Proc. Southern Illinois Conf., Southern Illinois Univ., Carbondale, Ill., 1979)*, volume 751 of *Lecture Notes in Math.*, pages 108–118. Springer, Berlin, 1979.

[20] D. R. Heath-Brown. The size of Selmer groups for the congruent number problem. II. *Invent. Math.*, 118(2):331–370, 1994. With an appendix by P. Monsky.

[21] D. R. Heath-Brown. A new form of the circle method, and its application to quadratic forms. *J. Reine Angew. Math.*, 481:149–206, 1996.

[22] D. Kane. On the ranks of the 2-Selmer groups of twists of a given elliptic curve. *Algebra Number Theory*, 7(5):1253–1279, 2013.

[23] D. M. Kane and J. A. Thorne. On the φ-Selmer groups of the elliptic curves $y^2 = x^3 - Dx$. *Math. Proc. Cambridge Philos. Soc.*, 163(1):71–93, 2017.

[24] N. M. Katz and P. Sarnak. Zeroes of zeta functions and symmetry. *Bull. Amer. Math. Soc. (N.S.)*, 36(1):1–26, 1999.

[25] B. D. Kim. The parity conjecture for elliptic curves at supersingular reduction primes. *Compos. Math.*, 143(1):47–72, 2007.

[26] Z. Klagsbrun, B. Mazur, and K. Rubin. A Markov model for Selmer ranks in families of twists. *Compos. Math.*, 150(7):1077–1106, 2014.

[27] D. Kriz. Supersingular main conjectures, Sylvester's conjecture and Goldfeld's conjecture. *arXiv:2002.04767*, 2020.

[28] J. Laga. Arithmetic statistics of Prym surfaces. *arXiv:2107.06803*, 2021.

[29] D. Lorenzini. Torsion and Tamagawa numbers. *Ann. Inst. Fourier (Grenoble)*, 61(5):1995–2037 (2012), 2011.

[30] P. Monsky. Generalizing the Birch-Stephens theorem. I. Modular curves. *Math. Z.*, 221(3):415–420, 1996.

[31] D. Mumford. Prym varieties. I. In *Contributions to analysis (a collection of papers dedicated to Lipman Bers)*, pages 325–350. 1974.

[32] D. Mumford. *Abelian varieties*, volume 5 of *Tata Institute of Fundamental Research Studies in Mathematics*. Published for the Tata Institute of Fundamental Research, Bombay; by Hindustan Book Agency, New Delhi, 2008. With appendices by C. P. Ramanujam and Yuri Manin, Corrected reprint of the second (1974) edition.

[33] J. Nekovář. On the parity of ranks of Selmer groups. IV. *Compos. Math.*, 145(6):1351–1359, 2009. With an appendix by Jean-Pierre Wintenberger.

[34] J. Neukirch, A. Schmidt, and K. Wingberg. *Cohomology of number fields*, volume 323 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2008.

[35] B. Poonen and E. Rains. Random maximal isotropic subspaces and Selmer groups. *J. Amer. Math. Soc.*, 25(1):245–269, 2012.

[36] K. Rubin and A. Silverberg. Ranks of elliptic curves in families of quadratic twists. *Experiment. Math.*, 9(4):583–590, 2000.

[37] S. Ruth. A bound on the average rank of j-invariant 0 elliptic curves. *Princeton University Ph.D. thesis*, 2013.

[38] E. S. Selmer. The Diophantine equation $ax^3+by^3+cz^3=0$. *Acta Math.*, 85:203–362 (1 plate), 1951.

[39] A. Shankar and J. Tsimerman. Counting S_5-fields with a power saving error term. *Forum Math. Sigma*, 2:Paper No. e13, 8, 2014.

[40] A. Shnidman. Quadratic twists of Abelian varieties with real multiplication. *Int. Math. Res. Not. IMRN*, (5):3267–3298, 2021.

[41] A. Shnidman and A. Weiss. Ranks of abelian varieties in cyclotomic twist families. *arXiv:2107.06803*, 2021.

[42] A. Silverberg. The distribution of ranks in families of quadratic twists of elliptic curves. In *Ranks of elliptic curves and random matrix theory*, volume 341 of *London Math. Soc. Lecture Note Ser.*, pages 171–176. Cambridge Univ. Press, Cambridge, 2007.

[43] J. H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, Dordrecht, second edition, 2009.

[44] A. Smith. *ℓ^∞-Selmer Groups in Degree ℓ Twist Families*. PhD thesis, Harvard University, Graduate School of Arts & Sciences, 2020.

[45] P. Swinnerton-Dyer. The effect of twisting on the 2-Selmer group. *Math. Proc. Cambridge Philos. Soc.*, 145(3):513–526, 2008.

[46] J. J. Sylvester. On Certain Ternary Cubic-Form Equations. *Amer. J. Math.*, 2(4):357–393, 1879.

[47] A. Várilly-Alvarado. Density of rational points on isotrivial rational elliptic surfaces. *Algebra Number Theory*, 5(5):659–690, 2011.

A A p-converse theorem for CM elliptic curves (by Ashay Burungale and Christopher Skinner)

In this appendix we explain a proof of:

Theorem A.1. *Let E be a CM elliptic curve over $\mathbb{Q}$ and let p be a prime of supersingular reduction for E. If*

(a) $\mathrm{corank}_{\mathbb{Z}_p}\mathrm{Sel}_{p^\infty}(E) = 1$ *and*

(b) *the localisation map* $\mathrm{Sel}_{p^\infty}(E) \xrightarrow{\sim} E(\mathbb{Q}_p) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p/\mathbb{Z}_p$ *is surjective,*

then

$$\mathrm{ord}_{s=1}L(E,s) = 1 = \mathrm{rank}_{\mathbb{Z}}E(\mathbb{Q}).$$

As a consequence we deduce:

Corollary A.2. *Let E be a CM elliptic curve over $\mathbb{Q}$ and let p be a prime of supersingular reduction for E. If*

(a) $\mathrm{Sel}_p(E) \simeq \mathbb{Z}/p\mathbb{Z}$ *and*

(b) *the localisation map* $\mathrm{Sel}_p(E) \to E(\mathbb{Q}_p)/pE(\mathbb{Q}_p)$ *is nonzero,*

then $\mathrm{ord}_{s=1}L(E,s) = 1 = \mathrm{rank}_{\mathbb{Z}}\, E(\mathbb{Q})$.

In the case of good ordinary reduction we actually have a stronger result:

Theorem A.3. *Let E be a CM elliptic curve over $\mathbb{Q}$ and let p be a prime of good ordinary reduction for E. Then*

$$\mathrm{corank}_{\mathbb{Z}_p}\mathrm{Sel}_{p^\infty}(E) = 1 \implies \mathrm{ord}_{s=1}L(E,s) = 1 = \mathrm{rank}_{\mathbb{Z}}\, E(\mathbb{Q}).$$

This theorem follows immediately from the main results of [5], [2], and [13].

As a consequence of Theorem A.3 we have:

Corollary A.4. *Let E be a CM elliptic curve over $\mathbb{Q}$ and let p be a prime of good ordinary reduction for E. If* $\mathrm{Sel}_p(E)/\mathrm{im}(E[p](\mathbb{Q})) \simeq \mathbb{Z}/p\mathbb{Z}$, *then* $\mathrm{ord}_{s=1}L(E,s) = 1 = \mathrm{rank}_{\mathbb{Z}}\, E(\mathbb{Q})$.

In the above corollary, $\mathrm{im}(E[p](\mathbb{Q}))$ is the image of $E[p](\mathbb{Q})$ under the Kummer map.

Before embarking on the proof of Theorem A.1, we make a few remarks about these results.

Remark A.5.

(i) We emphasize that all these results allow for $p = 2$. This is, of course, crucial for the application of Corollary A.2 in the main body of this paper.

(ii) In both the theorems and corollaries the finiteness of Ш(E) (that is, $\#$Ш$(E) < \infty$) can be added to the final conclusion.

(iii) Theorem A.1 is the culmination of a number of prior results, especially [10], [3], and [5].

(iv) A similar converse for non-CM curves was proved in [12], with similar application (cf. [1]).

A.1 Proof of Theorem A.1

The proof of Theorem A.1 ties together a number of results on the Iwasawa theory of elliptic curves, especially curves with CM, which we now recall.

A.1.1 Kato's main conjecture

Let E be an elliptic curve over $\mathbb{Q}$. For a prime p, let T denote the p-adic Tate module of E and $V = T \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$.

Let $\mathbb{Q}_\infty$ be the cyclotomic $\mathbb{Z}_p$-extension of $\mathbb{Q}$, $\Gamma = \mathrm{Gal}(\mathbb{Q}_\infty/\mathbb{Q})$ and $\Lambda = \mathbb{Z}_p[\![\Gamma]\!]$. Fix a topological generator $\gamma \in \Gamma$. For a finitely-generated Λ- or $\Lambda \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$-module M let $\xi(M)$ denote its characteristic ideal (which should be clear from context).

Let $S_{\mathrm{st}}(E) \subset H^1(\mathbb{Z}[\frac{1}{p}], T \otimes_{\mathbb{Z}_p} \Lambda^*)$ be the strict Selmer group of E over $\mathbb{Q}_\infty$ (the subgroup of classes that are trivial at p). Here Λ^* is the Pontryagin dual of Λ with $G_{\mathbb{Q}}$-acting by the inverse of the canonical character $G_{\mathbb{Q}} \twoheadrightarrow \Gamma \subset \Lambda^\times$. Let $X_{\mathrm{st}}(E)$ be the Pontryagin dual of $S_{\mathrm{st}}(E)$. It is one of the main results of Kato [7, Thm. 12.4] that $X_{\mathrm{st}}(E)$ (in the guise of $H^2(\mathbb{Z}[\frac{1}{p}], T \otimes_{\mathbb{Z}_p} \Lambda)$) is a finitely-generated torsion Λ-module. Let $\mathbf{z}_E \in H^1(\mathbb{Z}[\frac{1}{p}], T \otimes_{\mathbb{Z}_p} \Lambda) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$ denote the Beilinson–Kato element [7] and let $z_E \in H^1(\mathbb{Q}, V)$ be its image under the specialisation $\gamma \mapsto 1$. The following special case of [7, Conj. 12.10] is proved in [6]:

Theorem A.6. *Let E be a CM elliptic curve over $\mathbb{Q}$ and p any prime. Then*

$$\xi((H^1(\mathbb{Z}[\tfrac{1}{p}], T \otimes_{\mathbb{Z}_p} \Lambda) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)/(\Lambda \otimes \mathbb{Q}_p) \cdot \mathbf{z}_E) = \xi(X_{\mathrm{st}}(E) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p).$$

Remark A.7. For primes of ordinary reduction the same result is due to Kato and Rubin, at least if $p \nmid \#\mathcal{O}_K^\times$.

A.1.2 Perrin-Riou's Conjecture

Let E be an elliptic curve over $\mathbb{Q}$ and p a prime. Let $H^1_{\mathrm{f}}(\mathbb{Q}_p, V) \subset H^1(\mathbb{Q}_p, V)$ denote the subgroup arising from the Kummer image of $E(\mathbb{Q}_p)$. By Kato's explicit reciprocity law [7, Thm. 12.5],

$$\mathrm{loc}_p(z_E) \in H^1_{\mathrm{f}}(\mathbb{Q}_p, V) \iff L(E, 1) = 0.$$

If $L(E,1) = 0$, then Perrin-Riou [9] conjectured z_E to be closely linked with the arithmetic of E. The following theorem, proved in [3] and [8], is evidence for this in the supersingular case:

Theorem A.8. *Let E be an elliptic curve over $\mathbb{Q}$ and p a prime of supersingular reduction. If $L(E,1) = 0$, then there exist $P \in E(\mathbb{Q})$ and $c_P \in \mathbb{Q}^\times$ with the following properties.*

(a) We have

$$\log(\mathrm{loc}_p(z_E)) = c_P \left(\frac{p + 1 - a_p(E)}{p}\right) \cdot \log(P)^2$$

for $\log : H^1_{\mathrm{f}}(\mathbb{Q}_p, V) \to \mathbb{Q}_p$ the logarithm map associated to the Néron differential.

(b) The point P is non-torsion if and only if $\mathrm{ord}_{s=1} L(E, s) = 1$.

Remark A.9. For p a prime of good ordinary reduction the same result is known (cf. [4] for $p \geq 5$ and even for $p \geq 3$ under an assumption that the conductor of E is suitably minimal at all primes ramified in the CM field—a condition that can be relaxed). This allows for a uniform proof of Theorem A.1 for all primes of good reduction. Of course, the stronger result Theorem A.3 is known, but the existing proofs run along different lines than our proof of Theorem A.1.

A.1.3 Putting the pieces together

Let E be as in Theorem A.1. Let $\mathrm{Sel}_{\mathrm{st}}(E) \subset \mathrm{Sel}_{p^\infty}(E)$ be the strict Selmer group, consisting of those classes that vanish under loc_p. By the assumption, $\mathrm{Sel}_{\mathrm{st}}(E)$ is finite. The same is then true of $X_{\mathrm{st}}(E)/(\gamma-1)X_{\mathrm{st}}(E)$ (which naturally surjects onto the Pontryagin dual of $\mathrm{Sel}_{\mathrm{st}}(E)$ with finite kernel). Kato proved that $H^1(\mathbb{Z}[\frac{1}{p}], T \otimes_{\mathbb{Z}_p} \Lambda) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$ is a free $\Lambda \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$-module of rank one [7, Thm. 12.4]. So it follows from Theorem A.6 and the finiteness of $X_{\mathrm{st}}(E)/(\gamma-1)X_{\mathrm{st}}(E)$ that

$$0 \neq z_E \in H^1(\mathbb{Q}, V).$$

Since $\mathrm{Sel}_{\mathrm{st}}(E)$ is finite, it further follows that

$$0 \neq \mathrm{loc}_p(z_E) \in H^1(\mathbb{Q}_p, V).$$

The hypothesis that $\mathrm{corank}_{\mathbb{Z}_p}\mathrm{Sel}_{p^\infty}(E) = 1$ implies that $L(E,1) = 0$ (this just follows from the Gross–Zagier and Kolyvagin theorem or from the parity conjecture). It then follows from Theorem A.8 that $\mathrm{rank}_{\mathbb{Z}} E(\mathbb{Q}) \geq 1$ and $\mathrm{ord}_{s=1} L(E,s) = 1$. That $\mathrm{rank}_{\mathbb{Z}} E(\mathbb{Q}) = 1$ now follows from $\mathrm{corank}_{\mathbb{Z}_p}\mathrm{Sel}_{p^\infty}(E) = 1$. This completes the proof of Theorem A.1.

We now briefly explain how Corollary A.2 follows from Theorem A.1. We first note that since E has good supersingular reduction at p, $E[p](\mathbb{Q}_p) = 0$ by [11, Prop. 12(d)]. Hence $E[p](\mathbb{Q}) = 0$. The Cassels–Tate pairing implies that $\mathrm{Sel}_{p^\infty}(E) \cong (\mathbb{Q}_p/\mathbb{Z}_p)^r \oplus M \oplus M$, for some integer $r \geq 0$ and some finitely-generated torsion $\mathbb{Z}_p$-module M. As $\mathrm{Sel}_p(E)/\mathrm{im}(E[p](\mathbb{Q})) = \mathrm{Sel}_{p^\infty}(E)[p]$, condition (a) of the corollary then implies that $\mathrm{Sel}_{p^\infty}(E) \cong \mathbb{Q}_p/\mathbb{Z}_p$, so condition (a) of the theorem holds. As $E(\mathbb{Q}_p) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p/\mathbb{Z}_p \cong \mathbb{Q}_p/\mathbb{Z}_p$ and the natural map $E(\mathbb{Q}_p)/pE(\mathbb{Q}_p) \twoheadrightarrow (E(\mathbb{Q}_p) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p/\mathbb{Z}_p)[p]$ is an isomorphism (since $E[p](\mathbb{Q}_p) = 0$), condition (b) of the corollary then implies condition (b) of the theorem.

Acknowledgements. We thank Levent Alpöge, Manjul Bhargava and Ari Shnidman for helpful discussions. The work of C.S. was partially supported by the Simons Investigator Grant #376203 from the Simons Foundation and the National Science Foundation Grant DMS-1901985, and that of A.B. by the National Science Foundation Grant DMS-2001409.

References

[1] M. Bhargava and C. Skinner, *A positive proportion of elliptic curves over $\mathbb{Q}$ have rank one*, J. Ramanujan Math. Soc. 29 (2014), no. 2, 221–242.

[2] A. Burungale, F. Castella, C. Skinner, and Y. Tian, *p^∞-Selmer groups and rational points on CM elliptic curves*, Annales Math, Quebec, Special issue in honor of Bernadette Perrin-Riou (to appear).

[3] A. Burungale, S. Kobayashi and K. Ota, *p-adic L-functions and rational points on CM elliptic curves at inert primes*, preprint.

[4] A. Burungale, C. Skinner and Y. Tian, *Elliptic curves and Beilinson–Kato elements: rank one aspects*, preprint.

[5] A. Burungale and Y. Tian, *p-converse to a theorem of Gross-Zagier, Kolyvagin and Rubin*, Invent. Math. 220 (2020), no. 1, 211–253.

[6] A. Burungale and Y. Tian, *A rank zero p-converse to a theorem of Gross–Zagier, Kolyvagin and Rubin*, preprint.

[7] K. Kato, *p-adic Hodge theory and values of zeta functions of modular forms*, Cohomologies p-adiques et applications arithmétiques. III. Astérisque No. 295 (2004), ix, 117–290.

[8] S. Kobayashi, *The p-adic Gross–Zagier formula for elliptic curves at supersingular primes*, Invent. Math. 191 (2013), no. 3, 527–629.

[9] B. Perrin-Riou, *Fonctions L p-adiques d'une courbe elliptique et points rationnels*, Ann. Inst. Fourier (Grenoble) 43 (1993), no. 4, 945–995.

[10] K. Rubin, *p-adic variants of the Birch and Swinnerton-Dyer conjecture for elliptic curves with complex multiplication*, p-adic monodromy and the Birch and Swinnerton-Dyer conjecture (Boston, MA, 1991), 71–80, Contemp. Math., 165, Amer. Math. Soc., Providence, RI, 1994.

[11] J.-P. Serre, *Propriétés galoisiennes des points d'ordre fini des courbes elliptiques*, Invent. Math. 15 (1972), no. 4, 259–331.

[12] C. Skinner, *A converse to a theorem of Gross, Zagier and Kolyvagin*, Ann. of Math. (2) 191 (2020), no. 2, 329–354.

[13] Q. Yu, *p-Converse to a Theorem of Gross-Zagier, Kolyvagin, and Rubin for Small Primes*, 2021 Caltech Thesis, `https://resolver.caltech.edu/CaltechTHESIS:06022021-000654935`.