3-isogeny Selmer Groups and Binary Cubic Forms

Eliot Hodges

Math 223ar: Elliptic Curves
December 2024
Noam Elkies

1. Introduction

One of the driving questions in arithmetic statistics is that of the distribution of ranks of elliptic curves over global fields. This problem, along with the Cohen–Lenstra heuristics, which conjecture the distribution of class groups of number fields, has motivated the lion's share of research in arithmetic statistics over the past two decades. Conjectures of Goldfeld and Katz–Sarnak dating back to as early as 1979 state that 50% of elliptic curves over $\mathbb{Q}$ have rank 0 and that the other 50% have rank 1 [13, 14]. These densities are conjectured to be insensitive to whether the curves have been ordered by height, discriminant, or conductor.[1] The Birch–Swinnerton-Dyer conjecture, in conjunction with the generalized Riemann hypothesis, tell us that the average rank of elliptic curves, when ordered by height, is finite and at most 2.3. However, until 2015, it was not previously known that the average rank of elliptic curves is even finite.

For an elliptic curve over a global field K, the theory of descent tells us that, given generators for the weak Mordell–Weil group $E(K)/mE(K)$, it is possible to compute generators for $E(K)$. The difficulty lies with the fact that there is no known, general method for computing generators of the weak Mordell–Weil group. Enter the theory of Selmer groups. Selmer groups, which will be defined momentarily, are finite abelian groups associated to an isogeny of elliptic curves whose sizes give upper bounds on the size of the weak Mordell–Weil group and thus the rank of an elliptic curve. Despite their complexity, the advantage of working with Selmer groups is that, often, their elements are in natural correspondence with orbits of certain families of homogeneous polynomials. These *parametrizations*, one of which is the main concern of this paper, allow for the explicit computation of the average sizes of Selmer groups. In turn, these averages give upper bounds for the average ranks of elliptic curves.

Because we did not define Selmer groups of isogenies in class, we give a brief discussion of the requisite theory, adapted from Chapter X of Silverman's book [18]. Let E and E' be elliptic curves with an isogeny $\phi : E \to E'$ between them, all defined over a global field K. Take G_K to denote the absolute Galois group of K. We have an exact sequence of G_K-modules

$$0 \longrightarrow E[\phi] \longrightarrow E \xrightarrow{\phi} E' \longrightarrow 0$$

to which applying Galois cohomology yields the long exact sequence

$$\begin{tikzcd}
0 \arrow[r] & E(K)[\phi] \arrow[r] & E(K) \arrow[r, "\phi"] & E'(K) \arrow[dlll, "\partial"'] \\
H^1(G_K, E([\phi])) \arrow[r] & H^1(G_K, E) \arrow[r, "\phi"] & H^1(G_K, E') \arrow[r] & \cdots
\end{tikzcd}$$

From this, we extract the following short exact sequence

$$0 \longrightarrow E'(K)/\phi(E(K)) \xrightarrow{\partial} H^1(G_K, E[\phi]) \longrightarrow H^1(G_K, E)[\phi] \longrightarrow 0. \tag{1}$$

[1] When we reference probabilities and averages, we mean the following. Let h be a function from the isomorphism classes of elliptic curves over $\mathbb{Q}$ (or any global field) to the integers (h is called an *ordering*). The statements about the densities of ranks of elliptic curves can be restated as follows: consider the isomorphism classes of elliptic curves E over $\mathbb{Q}$ with $|h(E)| < X$ for some positive real number X. For E_X chosen uniformly at random from this set,

$$\lim_{X\to\infty} \mathbb{P}(\mathrm{rank}(E_X(\mathbb{Q})) = 0) = \frac{1}{2} = \lim_{X\to\infty} \mathbb{P}(\mathrm{rank}(E_X(\mathbb{Q})) = 1).$$

Similar remarks apply for the statements about average ranks of elliptic curves.

For each nonarchimedean place v of K, choose an extension of v to $\overline{K}$. Let $\overline{K}_v$ denote the completion of $\overline{K}$ with respect to this extension, and let $G_v \subset G_K$ be the corresponding decomposition group. We know that G_v acts on $E(\overline{K}_v)$ and $E'(\overline{K}_v)$. Repeating the process outlined above gives an exact sequence as in (1) where the Galois groups and fields have been replaced by their local counterparts. Assembling all of these local considerations, we see that the restriction maps on cohomology given by $G_v \subset G_K$ and $E(\overline{K}) \subset E(\overline{K}_v)$ yield the following commutative diagram with exact rows, where the products run over all nonarchimedean places of K:

$$\begin{tikzcd}
0 \arrow[r] & E'(K)/\phi(E(K)) \arrow[r, "\partial"] \arrow[d] & H^1(G_K, E[\phi]) \arrow[r] \arrow[d] & H^1(G_K, E)[\phi] \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \prod_v E'(K_v)/\phi(E(K_v)) \arrow[r, "\partial"] & \prod_v H^1(G_v, E[\phi]) \arrow[r] & \prod_v H^1(G_v, E)[\phi] \arrow[r] & 0.
\end{tikzcd}$$

Our interest lies in computing the image of $E'(K)/\phi(E(K))$ under the map ∂, or, equivalently, the kernel of the map $H^1(G_K, E[\phi]) \to H^1(G_K, E)[\phi]$. By the theory of twists, the problem of computing this kernel can be rephrased as the problem of determining whether certain principal homogeneous spaces (smooth curves C/K with a simply transitive group action of E on C defined over K) have a K-rational point. Unfortunately, this is not always tractable. However, determining the (local) kernel of $H^1(G_v, E[\phi]) \to H^1(G_v, E)[\phi]$ is made possible by Hensel's lemma, since it requires only finite computation to check whether a curve has a point in the finite ring $O_{K_v}/\pi_v^e O_{K_v}$. Thus, we define the *ϕ-Selmer group* of E/K to be the subgroup of $H^1(G_K, E[\phi])$ given by

$$\mathrm{Sel}_\phi(E) = \ker\left(H^1(G_K, E[\phi]) \to \prod_v H^1(G_v, E)[\phi]\right).$$

That is, the ϕ-Selmer group is the subgroup of $H^1(G_K, E[\phi])$ consisting all "locally soluble" cohomology classes, i.e., the cohomology classes of $H^1(G_K, E[\phi])$ that are locally in the image of the connecting map

$$\partial_v : E'(K_v) \to H^1(G_v, E[\phi])$$

for all archimedean places v. The *Tate–Shafarevich* group of E over K is defined as

$$\text{Ш}(E) = \ker\left(H^1(G_K, E) \to \prod_v H^1(G_v, E)\right).$$

It is well-known that $\mathrm{Sel}_\phi(E)$ is finite and that $\mathrm{Sel}_\phi(E)$ and $\text{Ш}(E)[\phi]$ fit into an exact sequence

$$\begin{tikzcd}
0 \arrow[r] & E'(K)/\phi(E(K)) \arrow[r] & \mathrm{Sel}_\phi(E) \arrow[r] & \text{Ш}(E)[\phi] \arrow[r] & 0.
\end{tikzcd}$$

From this exact sequence, we see that knowing the average size of $\mathrm{Sel}_\phi(E)$ implies upper bounds on the size of $E'(K)/\phi(E(K))$. Setting $E' = E$ and $\phi = [2]$, for example, we see that $\mathrm{Sel}_2(E)$ is an abelian 2-group of size 2^s for some integer $s \geq 0$. Knowing the average size of s then implies immediately implies an upper bound on $\mathrm{rank}(E(K))$.

There have been significant advances in the last thirty years towards understanding the average sizes of Selmer groups of elliptic curves over a global field K. Brumer showed that the average rank of elliptic curves over $\mathbb{F}_q(t)$ is finite for $q \geq 5$, and this result was strengthened by de Jong, who lowered the bound in question and extended these results for all q [5, 11]. Later, Bhargava and Shankar showed that the average size of 2-Selmer groups of elliptic curves over $\mathbb{Q}$ ordered by height is 3, implying that their average rank is at most 1.5 [3], giving the first ever bound on ranks of elliptic curves over number fields. The goal of this paper is to give an expository account of a paper of Bhargava, Elkies, and Shnidman, in which the authors prove results of a similar flavor for 3-isogeny Selmer groups of elliptic curves with j-invariant 0 [2]. Our main focus will be the parametrization that makes these results possible.

2. 3-isogenies of $y^2 = x^3 + k$

In the sequel, F will always denote a field with $\mathrm{char}(F) \notin \{2, 3\}$. For $k \in F^*$, the elliptic curve modeled by

$$E_k : y^2 = x^3 + k$$

has j-invariant 0, and, conversely an elliptic curve over F with j-invariant zero is isomorphic to E_k for some $k \in F^*$. The isomorphism classes of the E_k's are parametrized by F^*/F^{*6}, as two curves E_k and E_ℓ are isomorphic if and only if $\ell = m^6 k$ for some $m \in F^*$. As usual, let O denote the point at infinity on E_k. We see immediately (either from the group law or from computing the flexes of E_k) that $T = (0, \sqrt{k})$ and $-T = (0, -\sqrt{k})$ are two distinct 3-torsion points on E_k. The elliptic curve E_k comes equipped with an isogeny $\phi_k : E_k \longrightarrow E_{-27k}$, defined over F, given by

$$\phi_k(x, y) = \left(\frac{x^3 + 4k}{x^2}, \frac{y(x^3 - 8k)}{x^3} \right);$$

there is an isogeny $\hat{\phi}_k : E_{-27k} \longrightarrow E_{3^6 k}$ given by

$$\hat{\phi}_k(x, y) = \left(\frac{x^3 - 108k}{9x^2}, \frac{y(x^3 + 216k)}{27x^3} \right).$$

Some algebra verifies that the multiplication-by-3 map $[3] : E_k \longrightarrow E_k$ is equal to $\hat{\phi}_k \circ \phi_k$, so ϕ_k is a 3-isogeny with dual $\hat{\phi}_k$. Identifying $E_{3^6 k}$ with E_k via the map $(x, y) \mapsto (3^{-2}x, 3^{-3}y)$, we see that $\hat{\phi}_k$ is simply ϕ_{-27k}. When k is understood, we suppress the subscript from our notation and denote ϕ_k by ϕ. Recall that 3-isogenies are degree-3 maps and therefore have kernel of size 3. The ϕ-torsion points of E_k (respectively, $\hat{\phi}$-torsion) are O, T, and $-T$ (respectively, O, $(0, 3\sqrt{-3k})$, and $(0, -3\sqrt{-3k})$).

3. Binary Cubic Forms Á la Bhargava

Recall from Section 1 that the goal of this paper is to lay out a parametrization of the cohomology group $H^1(G_F, E_{-27k}[\hat{\phi}])$ by orbits of certain homogeneous polynomials called *triply symmetric binary cubic forms*, which will be defined imminently. Let R be a ring, and consider the lattice $\mathrm{Sym}^3(R^2)$ of binary cubic forms with integer coefficients, i.e., binary cubic forms

$$ax^3 + bx^2y + cxy^2 + dy^3$$

with $a, b, c, d \in R$. For our intents and purposes, it will be more convenient to study the dual lattice of $\mathrm{Sym}^3(R^2)$: the lattice $\mathrm{Sym}_3(R^2)$ of *triply symmetric* binary cubic forms

$$ax^3 + 3bx^2y + 3cxy^2 + dy^3$$

with $a, b, c, d \in R$. Denote this lattice by $V(R)$, and note that there is a natural $\mathrm{GL}_2(R)$-action on $V(R)$ given by

$$g \cdot f(x, y) = \det(g)^{-1} f((x, y)g).$$

We may associate to each form a natural invariant (that is, a polynomial in the coefficients in the form that is invariant under the action of $\mathrm{SL}_2(R)$) called its *(reduced) discriminant*: if $f(x, y) = ax^3 + 3bx^2y + 3cxy^2 + dy^3$, then the discriminant is given by

$$\mathrm{disc}(f) = -3b^2c^2 + 4ac^3 + 4b^3d + a^2d^2 - 6abcd.^2$$

Given $g \in \mathrm{GL}_2(R)$, we have that

$$\mathrm{disc}(g \cdot f) = \det(g)^6 \mathrm{disc}(f),$$

so the discriminant is indeed an invariant of f. For a subset U of $V(R)$, let U_d denote the set of forms in U with discriminant d.

Having established the requisite notation, we now state the parametrization in question:

Theorem 3.1 ([2] Theorem 27)**.** *There is a canonical bijection between $H^1(G_F, E_{-27k}[\hat{\phi}])$ and the $\mathrm{SL}_2(F)$-orbits of $V_{4k}(F)$. The $\mathrm{SL}_2(F)$-stabilizer of any f in $V_{4k}(F)$ is isomorphic to $E_{-27k}[\hat{\phi}](F)$.*

At the heart of Theorem 3.1 is a parametrization of what are essentially 3-torsion ideals in certain quadratic extensions of D by $\mathrm{SL}_2(D)$-orbits of binary cubic forms, where D is a Dedekind domain with characteristic not 2 or 3. In his seminal thesis, Bhargava astounded the mathematical community by offering a series of "higher composition laws," parametrizations of several families of arithmetic objects by simpler and more familiar objects, such as $2 \times 2 \times 2$ cubes of integers [1]. Among these results is a parametrization of 3-torsion ideal classes in quadratic extensions of $\mathbb{Z}$, which is a special case of the following.

Let D be a Dedekind domain with $\mathrm{char}(D) \notin \{2, 3\}$ and field of fractions F. Given $k \in F^*$, let $S_k = D[z]/(z^2 - k) \simeq D[\sqrt{k}]$ and $K_k = F[z]/(z^2 - k) \simeq F(\sqrt{k})$. When k is understood, we suppress it from our notation and denote S_k and K_k by S and K, respectively. A *fractional ideal* of S is a finitely generated S-submodule of K that spans K over F. To each fractional ideal, we may associate a fractional ideal of D via the following procedure. Consider I and S as D-modules in K. Because D is Dedekind, for each prime $\mathfrak{p}$ of D we have that the localization at $\mathfrak{p}$—denoted $D_{\backslash\mathfrak{p}}$ so as to avoid confusion with the completion at $\mathfrak{p}$—is a principal ideal domain. Since I is fractional, we have that $I_{\backslash\mathfrak{p}} \simeq K_{\backslash\mathfrak{p}} \simeq S_{\backslash\mathfrak{p}}$ as $D_{\backslash\mathfrak{p}}$-modules. Choose an isomorphism $\psi_\mathfrak{p} : S_{\backslash\mathfrak{p}} \longrightarrow I_{\backslash\mathfrak{p}}$ of $D_{\backslash\mathfrak{p}}$-modules, and let $\hat{\psi}_\mathfrak{p}$ denote its extension to $F_{\backslash\mathfrak{p}}$. Set

$$[S_{\backslash\mathfrak{p}} : I_{\backslash\mathfrak{p}}] = \det(\psi_\mathfrak{p}) D_{\backslash\mathfrak{p}}.$$

Note that the ideal $[S_{\backslash\mathfrak{p}} : I_{\backslash\mathfrak{p}}]$ does not depend on choice of isomorphism $\psi_\mathfrak{p}$, since choosing any other isomorphism simply changes the determinant by a unit. Define the *norm of I relative to S* to be

$$[S : I] = \bigcap_{\mathfrak{p}} [S_{\backslash\mathfrak{p}} : I_{\backslash\mathfrak{p}}],$$

where the intersection runs over all primes of D. It is immediate from the definition of the ideal norm that the ideal norm of a principal ideal generated by an element α in K is simply the ideal of F generated by the element norm of α, which we denote using $N(\alpha)$.

We are now properly algebraically equipped to define the arithmetic objects parametrized by binary cubic forms:

[2] Note that the reduced discriminant is not the usual polynomial discriminant; rather, it is straightforwardly verified that $\mathrm{disc}(f) = -3^{-3}\mathrm{Disc}(f)$, where $\mathrm{Disc}(f)$ denotes the usual discriminant.

Definition 3.2. A triple (I, δ, s) consisting of a fractional ideal I of S, element $\delta \in K^*$, and element $s \in F^*$ is said to be *valid* if $I^3 \subset \delta S$, the norm of I relative to S is the principal ideal $[S : I] = sD$ in F, and $N(\delta) = s^3$. We define an equivalence relation on valid triples by setting (I, δ, s) and (I', δ', s') to be equal if there exists a κ in K^* such that $I' = \kappa I$, $\delta' = \kappa^3\delta$, and $s' = N(\kappa)s$.

On one side of our bijection are equivalence classes of balanced triples; on the other are $\mathrm{SL}_2(D)$-orbits of $V(D)_{4k}$, i.e., $\mathrm{SL}_2(D)$-orbits of triply symmetric binary cubic forms with coefficients in D and discriminant $4k$. Given a valid triple (I, δ, s), we can construct an element of $V(D)_{4k}$ as follows. Because $S \simeq D[\sqrt{k}]$, we may express S as a D-module as $D + D\sqrt{k}$. Now, we leverage the fact that $[S : I] = sD$ is assumed to be principal to elicit a D-module basis of I. This implies that S and I are isomorphic as D-modules; thus, we may write $I = D\alpha + D\beta$ for some $\alpha, \beta \in I$.[3] The validity of (I, δ, s) implies that $I^3 \subset \delta S$, so we have

$$\begin{aligned} \alpha^3 &= \delta(c_0 + a_0\sqrt{k}) \\ \alpha^2\beta &= \delta(c_1 + a_1\sqrt{k}) \\ \alpha\beta^2 &= \delta(c_2 + a_2\sqrt{k}) \\ \beta^3 &= \delta(c_3 + a_3\sqrt{k}), \end{aligned} \tag{2}$$

where the a_i's and c_i's are all in D. To the valid ideal (I, δ, s) we associate the form

$$f(x, y) = a_0x^3 + 3a_1x^2y + 3a_2xy^2 + a_3y^3.$$

Note that changing (α, β) to some other D-module basis of I would simply transform f by the corresponding element of $\mathrm{SL}_2(D)$. Equivariantly, if π denotes the natural map $S \to S/D \simeq D\sqrt{k}$, we may regard f as

$$f(x, y) = \pi\left(\frac{(\alpha x + \beta y)^3}{\delta}\right),$$

which factors through the map $I \otimes I \otimes \delta^{-1}I \to S/D$ given by $u \otimes v \otimes \delta^{-1}w \mapsto \pi(uvw/\delta)$.

If $I = S$, $\alpha = 1$, and $\beta = \sqrt{k}$, then (2) implies that the corresponding form is $f_S(x, y) = 3x^2y + ky^3$, which has discriminant $4k$. For a general valid triple (I, δ, s), we know that there is an element $g \in \mathrm{GL}_2(F)$ relating the basis $(1, \sqrt{k})$ (of K as an F-vector space) to the basis (α, β). The form corresponding to this general (I, δ, s) is given by regarding f_S as a map from $S \to S/D$, which factors through $S \otimes S \otimes S \to S/D$, and applying g the first two factors of S in $S \otimes S \otimes S$ and applying $\delta^{-1}g$ to the last factor. The discriminant of the resulting form is simply $\det(g)^6 N(\delta)^{-2}\mathrm{disc}(f_S)$. Note that the ideal in D generated by $\det(g)$ is simply $[S : I]$, whose cube is $N(\delta)$ by hypothesis. It follows that the discriminant of the form corresponding to (I, δ, s)

[3] The curious reader might wonder whether a similar result holds for fractional ideals in an arbitrary quadratic extensions of a Dedekind domain. Indeed, Bhargava parametrizes triples (S, I, δ) consisting of a quadratic ring extension S of $\mathbb{Z}$, a fractional ideal I of S, and an element δ of $S \otimes \mathbb{Q}$ such that $I^3 \subset \delta S$ and $N(\delta)S$. Notably, his result permits *all* quadratic extensions of $\mathbb{Z}$, not only those of the form $\mathbb{Z}[\sqrt{k}]$ for $k \in \mathbb{Z}$. While it is beyond the scope of this paper, such a generalization does exist and is one of the results of my senior thesis. The methods used to construct the bijection in this paper fail because the class group interferes: a general quadratic extension S of D is not always isomorphic (as a D-module) to $D \oplus D$; rather, a theorem of Steinitz tells us that S can be written as $D \oplus \mathfrak{a}$ for any ideal $\mathfrak{a}$ representing an invariant of S called its *Steinitz class.* The D-module isomorphism class of S is entirely determined by its rank and this Steinitz ideal class of $\mathrm{Cl}(R)$. In this scenario, more work must be done to construct the parametrization.

is equal to $\mathrm{disc}(f_S) = 4k$. Hence, any form corresponding to a valid triple has discriminant $4k$, allowing us to state the following theorem:

Theorem 3.3 ([2] Theorem 18)**.** *For each $k \in F^*$, the construction outlined in the above gives a canonical bijection*

$$\Phi_k : \{\textit{equivalence classes } (I, \delta, s) \textit{ of valid ideals of } S]\} \longrightarrow \mathrm{SL}_2(D)\backslash V(D)_{4k}.$$

Under this bijection, the stabilizer in $\mathrm{SL}_2(D)$ of f in $V(D)_{4k}$ is isomorphic to $S(I)^*[3]_{N=1}$, where $S(I)$ is the ring of S-module endomorphisms of I.

Proof. That Φ_k is a bijection follows from emulating the proof of Theorem 13 in [1]; we omit this for the sake of brevity. To see that the stabilizer of $f \in V(D)_{4k}$ in $\mathrm{SL}_2(D)$ is isomorphic to $S(I)^*[3]_{N=1}$, suppose f corresponds to (I, δ, s), and regard f as the map

$$I \otimes I \otimes \delta^{-1} I \to S/D.$$

The elements of $K^*_{N=1}$ preserving this map are precisely $S(I)^*[3]_{N=1}$. □

We will be particularly interested in Theorem 3.3 when $D = F$. Of particular importance in this scenario is the *restriction of scalars* group scheme $\mathrm{Res}_F^K(\mu_3)$. For an F-algebra A, we have that $\mathrm{Res}_F^K(\mu_3)(A) = \mu_3(A \otimes_F K)$. From this description, we see that the F-points of $\mathrm{Res}_F^K(\mu_3)$ are simply $\mu_3(K)$, that is, the cube roots of unity in K. The K-points of $\mathrm{Res}_F^K(\mu_3)$ are given by $\mu_3(K \otimes_F K)$, which is isomorphic to $\mu_3(K) \otimes \mu_3(K)$ via the map $a \otimes b \mapsto (ab, a\sigma(b))$, where $\sigma : K \to K$ is the involution taking $\sqrt{k} \mapsto -\sqrt{k}$. Because $F \subset K$, there is an inclusion $\mathrm{Res}_F^K(\mu_3)(F) \hookrightarrow \mathrm{Res}_F^K(\mu_3)(K)$ given by $a \mapsto (a, \sigma(a))$. The norm map $N : K \to F$ gives us a norm map $N : \mathrm{Res}_F^K(\mu_3) \to \mu_3$. On the F-points, the norm map is simply the element norm $\mu_3(K) \to \mu_3(F)$; on the K-points, the norm map takes $(a, b) \mapsto ab$. Note that the restriction of the norm map to the copy of $\mathrm{Res}_F^K(\mu_3)(F)$ contained in $\mathrm{Res}_F^K(\mu_3)(K)$ agrees with the norm map on $\mu(K)$. For a subgroup G of $\mathrm{Res}_F^K(\mu_3)$, let $G_{N=1}$ denote the subgroup of elements of G with norm 1. These descriptions of $\mathrm{Res}_F^K(\mu_3)(K)$ are immaterial at the moment, but they will be necessary when elliptic curves return in the next section.

Corollary 3.4. *For $k \in F^*$, the bijection Φ_k yields a natural bijection between the set of $\mathrm{SL}_2(F)$-orbits on $V(F)_{4k}$ and the group $(K^*/K^{*3})_{N=1}$. Under this bijection, the stabilizer of f in $V(F)_{4k}$ is isomorphic to $\mathrm{Res}_F^K(\mu_3)(F)_{N=1}$.*

Proof. Apply Theorem 3.3 to $D = F$. Then $S = K = F(\sqrt{k})$, which is either a field or isomorphic to $F \times F$. In either case, there is only one fractional ideal of S (the ring S itself), so valid triples are of the form (S, δ, s) with $\delta \in K^*$ and $N(\delta) = s^3$. Up to equivalence, we see that a valid triple corresponds to an element δ of K^*/K^{*3} with $N(\delta) \in F^{*3}$. The statement involving the stabilizer follows immediately. □

We can make the correspondence of Corollary 3.4 quite explicit. Let $\delta \in (K^*/K^{*3})_{N=1}$. Writing $\delta = a + b\sqrt{k}$ so that $a^2 + kb^2 = 1$, (2) in conjunction with Corollary 3.4 implies that the form corresponding to δ is $f(x, y) = -bx^3 + 3ax^2y - 3bkxy^2 + kay^3$.

4. Cubic Forms and Cohomology

Having completed our lengthy excursion investigating the deep connection between valid triples of S and binary cubic forms, we now return to the setting of elliptic curves with j-invariant 0. Let F be a field. Recall the elliptic curves E_k and E_{-27k} from Section 2, which are dual to each other via the 3-isogenies $\phi : E_k \longrightarrow E_{-27k}$ and $\hat{\phi} : E_{-27k} \longrightarrow E_k$. Of crucial importance are the pair of quadratic étale algebras over F,

$$K = F[z]/(z^2 - k) \simeq F(\sqrt{k}) \quad \text{and} \quad \hat{K} = F[z]/(z^2 + 27k) \simeq F(\sqrt{-3k}),$$

over which $E_k(K)$ and $E_{-27k}(K)$ contain $E_k[\phi]$ and $E_{-27k}[\hat{\phi}]$, respectively.

With this in mind, it should not be surprising that the G_F-cohomology of E_{-27k} and the arithmetic of K are intimately connected, and this connection can be made precise as follows. Because E_k and E_{-27k} are dual to each other, there is nondegenerate pairing

$$\langle\, , \rangle : E_{-27k}[\hat{\phi}] \otimes E_k[\phi] \longrightarrow \mu_3$$

defined in exactly the same manner as the Weil pairing on integral torsion points. We define a map of groups $\iota : E_{-27k}[\hat{\phi}] \longrightarrow \mathrm{Res}_F^K(\mu_3)(K)$ given by

$$P \longmapsto (\langle P, T \rangle, \langle P, -T \rangle)$$

(recall from Section 2 that $\pm T = (0, \pm\sqrt{k})$ are the nontrivial ϕ-torsion points of E_k).

Theorem 4.1 ([2] Proposition 24)**.** *The map $\iota : E_{-27k}[\hat{\phi}] \longrightarrow \mathrm{Res}_F^K(\mu_3)(K)$ is an injective group homomorphism and image equal to $\ker(\mathrm{Res}_F^K(\mu_3)(K) \longrightarrow \mu_3(K))$. In other words,*

$$E_{-27k}[\hat{\phi}] \simeq \ker(\mathrm{Res}_F^K(\mu_3)(K) \longrightarrow \mu_3(K)).$$

This induces an isomorphism

$$H^1(G_F, E_{-27k}[\hat{\phi}]) \simeq (K^*/K^{*3})_{N=1},$$

where $(K^*/K^{*3})_{N=1}$ denotes the kernel of the norm map $K^*/K^{*3} \longrightarrow F^*/F^{*3}$.

Proof. Verifying that ι is a group homomorphism is a straightforward manipulation. Its injectivity follows from the nondegeneracy of the Weil pairing, and we see that its image is contained in $\ker(\mathrm{Res}_F^K(\mu_3)(K) \longrightarrow \mu_3(K))$ because $\langle P, T \rangle \langle P, -T \rangle = \langle P, O \rangle = 1$. That $E_{-27k}[\hat{\phi}] \simeq \ker(\mathrm{Res}_F^K(\mu_3)(K) \longrightarrow \mu_3(K))$ follows immediately by counting.

For the second statement, consider the short exact sequence of groups

$$0 \longrightarrow E_{-27k}[\hat{\phi}] \longrightarrow \mathrm{Res}_F^K(\mu_3)(F) \longrightarrow \mu_3(F) \longrightarrow 0,$$

where the third arrow is the norm map. Applying Galois cohomology with G_F gives us a long exact sequence

$$\begin{tikzcd}
0 \arrow[r] & \{O\} \arrow[r] & \mu_3(K) \arrow[r, "N"] & \mu_3(F) \arrow[dlll, "\partial"'] \\
H^1(G_F, E_{-27k}[\hat{\phi}]) \arrow[r] & H^1(G_F, \mu_3(K)) \arrow[r, "N"] & H^1(G_F, \mu_3(F)) \arrow[r] & \cdots
\end{tikzcd}$$

By the Galois cohomology of the Kummer sequence, we know that $H^1(G_F, \mu_3(K)) \simeq K^*/K^{*3}$ and $H^1(G_F, \mu_3(F)) \simeq F^*/F^{*3}$, and the map between them is the usual element norm. The exactness of the above implies that $H^1(G_F, E_{-27k}[\hat{\phi}])$ is isomorphic to the kernel of the norm map $K^*/K^{*3} \longrightarrow F^*/F^{*3}$. □

Remark 4.2. In this section, we just as well could have chosen ϕ and $\hat{K}$ and proven analogous results for $E_k[\phi]$ and $\mathrm{Res}_F^{\hat{K}}(\mu_3)$. Simply make the change of variable $k \longmapsto -27k$ to generate the statements of these results.

Applying Galois cohomology to the sequence $0 \longrightarrow E_{-27k}[\hat{\phi}] \longrightarrow E_{-27k} \longrightarrow E_k \longrightarrow 0$ induces a map

$$\partial : E_k(F) \longrightarrow H^1(G_F, E_{-27k}[\hat{\phi}]) \simeq (K^*/K^{*3})_{N=1}.$$

Explicitly, if $(x, y) \in E_k(F) \setminus E_k[\phi](F)$, then $\partial(x, y) = y + \sqrt{k}$. We also have $\partial(\pm T) = \pm 1/(2\tau)$. A more detailed account is given in Section 14 of [6]. Combining Corollary 3.4 and Theorem 4.1, implies Theorem 3.1, which we restate for convenience.

Theorem 4.3. *There is a canonical bijection between $H^1(G_F, E_{-27k}[\hat{\phi}])$ and the $\mathrm{SL}_2(F)$-orbits of $V_{4k}(F)$. The $\mathrm{SL}_2(F)$-stabilizer of any f in $V_{4k}(F)$ is isomorphic to $E_{-27k}[\hat{\phi}](F)$.*

While Theorem 3.1 gives a parametrization of $H^1(G_F, E_{-27k}[\hat{\phi}])$, recall that our ultimate goal is to study $\mathrm{Sel}_{\hat{\phi}}(E_{-27k})$. To that end, let $V(F)^{\mathrm{sol}}$ denote the set of forms in $V(F)$ corresponding to cohomology classes in $H^1(G_F, E_{-27k}[\hat{\phi}])$ in the image of the connecting map $\partial : E_k(F) \longrightarrow H^1(G_F, E_{-27k}[\hat{\phi}])$ for some $k \in F^*$. Call the elements of $V(F)^{\mathrm{sol}}$ *soluble.* Under the bijection of Theorem 3.1, the orbits of $V(F)^{\mathrm{sol}}$ correspond to $E_k(F)/\hat{\phi}(E_{-27k}(F))$:

Corollary 4.4 ([2] Corollary 28)**.** *There is a canonical bijection between the $\mathrm{SL}_2(F)$-orbits on $V(F)^{\mathrm{sol}}_{4k}$ and the elements of $E_k(F)/\hat{\phi}(E_{-27k}(F))$, under which the identity element of the group $E_k(F)/\hat{\phi}(E_{-27k}(F))$ corresponds to the unique $\mathrm{SL}_2(F)$-orbit of reducible binary cubic forms in $V(F)^{\mathrm{sol}}_{4k}$, namely, the orbit of $f(x, y) = kx^3 + 3xy^2$.*

The moniker *soluble* may initially seem out of place, but it is explained by the following theorem, which reveals the fundamental connection between forms $f \in V(F)_{4k}$ in the image of the connecting homomorphism $\partial : E_k(F) \longrightarrow H^1(G_F, E_{-27k}[\hat{\phi}]) \simeq (K^*/K^{*3})_{N=1}$ and F-rational points on a ϕ-coverings of E_k.

Theorem 4.5. *Given a form $f \in V(F)_{4k}$, its corresponding cohomology class in $H^1(G_F, E_{-27k}[\hat{\phi}])$ lies in the image of $\partial : E_k(F) \longrightarrow H^1(G_F, E_{-27k}[\hat{\phi}]) \simeq (K^*/K^{*3})_{N=1}$ if and only if the variety $C_f : z^3 = f(x, y)$ in $\mathbb{P}^2$ has an F-rational point.*

Proof. Given $f \in V(F)_{4k}$, Corollary 3.4 tells us that there is some element $\delta \in (K^*/K^{*3})_{N=1}$ corresponding to the $\mathrm{SL}_2(F)$-orbit of f. If δ is in the image of ∂, then $\delta = \partial(u, v) = v - \sqrt{k}$ for some $(u, v) \in E_k(F)$. The corresponding form is $g(x, y) = x^3 + 3vx^2y + 3bkxy^2 + kvy^3$, so f is $\mathrm{SL}_2(F)$-equivalent to this form. We see that $1^3 = g(1, 0)$, implying that $1^3 = f(x, y)$ has an F-rational point.

Conversely, if $z^3 = f(u, v)$ for $u, v, z \in F$, then $1 = f(u/z, v/z)$, and we may assume $z = 1$. Then by (2)

$$(u + v\sqrt{k})^3 = \delta(g(u, v) + f(u, v)\sqrt{k}) = \delta(g(u, v) + \sqrt{k})$$

for some $\delta \in K^*$ with $N(\delta) = s^3$ for $s \in F$ and binary cubic form g with coefficients in F. Taking the norm of both sides of the above equation, we see that

$$N(u + v\sqrt{k})^3 = s^3(g(u, v)^2 - k).$$

Hence, $(N(u + v\sqrt{k})/s, g(u, v))$ is a point of $E_k(F)$, and $\partial(N(u + v\sqrt{k})/s, g(u, v)) = g(u, v) - \sqrt{k}$ and δ belong to the same class of $(K^*/K^{*3})_{N=1}$ because

$$\delta = \frac{(u + v\sqrt{k})^3}{(g(u, v) + \sqrt{k})} = (g(u, v) - \sqrt{k})\frac{(u + v\sqrt{k})^3}{g(u, v) - k} = (g(u, v) - \sqrt{k})\left(\frac{(u + v\sqrt{k})s}{N(u + v\sqrt{k})}\right)^3.$$

Therefore, $f(x, y)$ corresponds to an element in the image of ∂. □

Recall that $H^1(G_F, E_{-27k}[\hat{\phi}])$ is in bijection with the group of isomorphism classes of $\hat{\phi}$-coverings of E_k, i.e., isomorphism classes of maps of curves $C \longrightarrow E_k$ over F that are twists of $\hat{\phi}$ (that is, the map $C \longrightarrow E_k$ becomes isomorphic to $\hat{\phi}$ over an algebraic closure $\overline{F}$). Given $\delta \in (K^*/K^{*3})_{N=1} \simeq H^1(G_F, E_{-27k}[\hat{\phi}])$, we may construct the corresponding $\hat{\phi}$-covering as follows. Let $s \in F$ be such that $N(\delta) = s^3$; let f be the corresponding binary cubic form given by applying Corollary 3.4. We take C_f to be the vanishing locus of $z^3 = f(x, y)$ in $\mathbb{P}^2$. There is a map $C_f \longrightarrow E_k$ given by $[u : v : z] \mapsto ((u^2 - kv^2)/s, g(u, v))$, where $g(u, v)$ is the binary cubic form from the proof above this paragraph. We see that this is exactly the $\hat{\phi}$-covering map corresponding to δ. Note that it is this correspondence between $\hat{\phi}$-coverings and binary cubic forms that lies at the heart of our parametrization.

Up to this point, while we have constructed a parametrization of $H^1(G_F, E_{-27k}[\hat{\phi}])$, we have not yet laid out a parametrization of $\mathrm{Sel}_{\hat{\phi}}$, which sits as a subgroup of $H^1(G_F, E_{-27k}[\hat{\phi}])$. Using Theorem 4.5, we are now able to elicit the desired correspondence. Let $V(F)^{\text{loc. sol}}$ denote the set of *locally soluble* triply symmetric binary cubic forms with coefficients in F. By this we mean the forms $f \in V(F)$ such that $C_f : z^3 = f(x, y)$ has a nontrivial solution over F_v for every place v of F. The following is an immediate consequence of Theorem 3.1, Corollary 3.4, and Theorem 4.5.

Theorem 4.6. *Given $k \in F^*$, there is a bijection between the $\mathrm{SL}_2(F)$-orbits on $V(F)^{\text{loc. sol}}$ of discriminant $4k$ and the elements of $\mathrm{Sel}_{\hat{\phi}}(E_{-27k})$ corresponding to the isogeny $\hat{\phi}_{-27k} : E_{-27k} \longrightarrow E_k$. The identity element of $\mathrm{Sel}_{\hat{\phi}}(E_{-27k})$ corresponds to the unique $\mathrm{SL}_2(F)$-orbit of reducible binary cubic forms, that is, the orbit of $f(x, y) = kx^3 + 3xy^2$. The $\mathrm{SL}_2(F)$-stabilizer of any $f \in V(F)_{4k}^{\text{loc. sol}}$ is isomorphic to $E_{-27k}[\hat{\phi}](F)$.*

5. Historical Contextualization and Counting Methods

The connections between ϕ-Selmer groups and binary cubic forms are historic and were initially investigated by Selmer himself [17]. This relationship was further fleshed out in the work of Cassels, Satgé, and Liverance [6, 16, 15]. Using Satgé's work, Fouvry deduced the boundedness of the average rank of elliptic curves over $\mathbb{Q}$ with j-invariant 0 using Davenport and Heilbronn's seminal paper in which they compute the average size of the 3-torsion subgroups of class groups of quadratic fields [12, 10]. The Davenport–Heilbronn method involves counting integer-coefficient binary cubic forms in order to generate a count of cubic fields of bounded discriminant; applying class field theory allows for the transformation of the result on cubic fields to one about the 3-torsion in class groups of quadratic fields. Later, Bhargava and Varma used one of Bhargava's higher composition laws (a special case of Theorem 3.3) and a count of triply symmetric integral binary cubic forms to more directly count the 3-torsion ideal classes in quadratic fields—a new and more direct way of accessing the classical results of Davenport and Heilbronn [4]. By analogy, this led Bhargava, Elkies, and Shnidman to suspect that a similar, more direct correspondence between ϕ-Selmer groups and binary cubic forms might be within reach; this hunch was realized as Theorem 3.1.

Theorem 3.1 reduces the computing the average size of $\mathrm{Sel}_\phi(E_k)$ to counting certain orbits of binary cubic forms of bounded discriminant; the solution to this second problem is a straightforward application of methods developed by Davenport [8, 9]. Essentially, one proceeds by constructing a fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z})$ on $V(\mathbb{R})$ using the Iwasawa decomposition of $\mathrm{SL}_2(\mathbb{R})$. Then $V(\mathbb{Z})$ sits as a lattice inside this fundamental domain, and *Davenport's lemma* tells us that the number of lattice points inside this fundamental domain is given by its volume plus a negligible error term [7]. Once this count of $\mathrm{SL}_2(\mathbb{Z})$-orbits on $V(\mathbb{Z})$ has been given, we apply sieve methods developed by Davenport and Bhargava–Shankar to obtain a formula for the count in terms of a product of local densities, which may be computed explicitly using Tamagawa numbers and Tate's algorithm.

6. Implications

Having proved Theorem 4.6, we now list the plethora of beautiful results that it ultimately implies. Describing the techniques used to generate these results in detail is beyond the scope of this paper—the point of this section is to illustrate the types of results Theorem 4.6 unlocks.[4]

First, we take $F = \mathbb{Q}$ and order the E_k's in the natural way (by the absolute value of k).[5] Let

$$r = \frac{103 \cdot 229}{2 \cdot 3^2 \cdot 7^2 \cdot 13} \prod_{p \equiv 5 \bmod 6} \frac{(1-p^{-1})(1+p^{-1}+5p^{-2}/3+p^{-3}+5p^{-4}/3+p^{-5})}{1-p^{-6}}.$$

We may approximate the product in the above by 1.0337, so $r \approx 2.1265$. Then we have:

Theorem 6.1 ([2] Theorem 1)**.** *The average size of the ϕ_k-Selmer group is $1 + r$ if k is negative and $1 + r/3$ if k is positive.*

Using the approximation from above, we see that the average size of the ϕ_k-Selmer group is approximately 3.1265 when k is negative and 1.7088 when $1 + r/3$ is positive. The methods used to deduce Theorem 6.1 give further insight into the 3-Selmer rank of E_k and the average rank of E_k:

Theorem 6.2 ([2] Theorem 10)**.** *For each $m \geq 0$, a positive proportion of elliptic curves $E_k : y^2 = x^3 + k$ with $k \in \mathbb{Z}$ have 3-Selmer rank m.*

Theorem 6.3 ([2] Theorem 5)**.** *The (limit supremum of the) average rank of the elliptic curves over $\mathbb{Q}$ with j-invariant 0 is less than 1.29.*

It follows immediately that a positive proportion of the E_k's must have rank 0 or rank 1. The methods used to deduce Theorem 6.3 are strong enough to give lower bounds on the proportion of curves having rank 0 or 1. It turns out that the majority of the curves over $\mathbb{Q}$ with j-invariant 0 have rank 0 or 1:

Theorem 6.4 ([2] Theorems 6, 7, and 8)**.** *At least 19.9% of elliptic curves E_k with $k \in \mathbb{Z}$ have rank 0. At least 41.1% of elliptic curves E_k with $k \in \mathbb{Z}$ have rank 1. Thus, at least 61% of all such curves have rank 0 or 1.*

Now, recall that Theorem 3.1 and Theorem 4.6 hold for any field whose characteristic is not 2 or 3. The prior theorems are all over $\mathbb{Q}$ and do not realize the full power of the parametrization from Section 4. To that end, we have the following theorems:

Theorem 6.5. *Let F be a number field such that $\mu_3 \not\subset F$. Order the elliptic curves E_k, where $k \in F^*/F^{*6}$, by the height of k. The average rank of the E_k over F is bounded. A positive proportion of curves E_k have 3-Selmer rank 0 over F and thus also Mordell–Weil rank 0. A positive proportion of curves E_k have 3-Selmer rank 1 over F.*

Theorem 6.6. *Let F be a number field such that $\mu_3 \subset F$. Order the elliptic curves E_k, where $k \in F^*/F^{*6}$, by the height of k. Then the average size of the Selmer groups $\mathrm{Sel}_{\phi_k}(E_k)$ and $\mathrm{Sel}_{\hat{\phi}}(E_{-27k})$ over F is 2. The average rank of the curves E_k is at most 1; at least 50% of the E_k have rank 0 over F.*

[4] For a brief description, see Section 5.

[5] Ordering the E_k's by k is simply ordering them by height.

References

[1] Bhargava, M. Higher composition laws. I. A new view on Gauss composition, and quadratic generalizations. *Ann. of Math. (2) 159*, 1 (2004), 217–250.

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

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

[4] Bhargava, M., and Varma, I. The mean number of 3-torsion elements in the class groups and ideal groups of quadratic orders. *Proc. Lond. Math. Soc. (3) 112*, 2 (2016), 235–266.

[5] Brumer, A. The average rank of elliptic curves. I. *Invent. Math. 109*, 3 (1992), 445–472.

[6] Cassels, J. W. S. *Lectures on elliptic curves*, vol. 24 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1991.

[7] Davenport, H. On a principle of Lipschitz. *J. London Math. Soc. 26* (1951), 179–183.

[8] Davenport, H. On the class-number of binary cubic forms. I. *J. London Math. Soc. 26* (1951), 183–192.

[9] Davenport, H. On the class-number of binary cubic forms. II. *J. London Math. Soc. 26* (1951), 192–198.

[10] Davenport, H., and Heilbronn, H. On the density of discriminants of cubic fields. II. *Proc. Roy. Soc. London Ser. A 322*, 1551 (1971), 405–420.

[11] de Jong, A. J. Counting elliptic surfaces over finite fields. vol. 2. 2002, pp. 281–311. Dedicated to Yuri I. Manin on the occasion of his 65th birthday.

[12] Fouvry, E. Sur le comportement en moyenne du rang des courbes $y^2 = x^3 + k$. In *Séminaire de Théorie des Nombres, Paris, 1990–91*, vol. 108 of *Progr. Math.* Birkhäuser Boston, Boston, MA, 1993, pp. 61–84.

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

[14] Katz, N. M., and Sarnak, P. *Random matrices, Frobenius eigenvalues, and monodromy*, vol. 45 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1999.

[15] Liverance, E. Binary cubic forms with many integral points. No. 998. 1997, pp. 93–101. Algebraic number theory and related topics (Japanese) (Kyoto, 1996).

[16] Satgé, P. Groupes de Selmer et corps cubiques. *J. Number Theory 23*, 3 (1986), 294–317.

[17] Selmer, E. S. The diophantine equation $ax^3 + by^3 + cz^3 = 0$. Completion of the tables. *Acta Math. 92* (1954), 191–197.

[18] Silverman, J. H. *The arithmetic of elliptic curves*, second ed., vol. 106 of *Graduate Texts in Mathematics*. Springer, Dordrecht, 2009.