A STATISTICAL ANALYSIS OF 2-SELMER GROUPS FOR ELLIPTIC CURVES WITH TORSION SUBGROUP $Z_2 \times Z_8$

JESSICA FLORES, KIMBERLY JONES, ANNE ROLLICK, AND JAMES WEIGANDT

Abstract. We consider elliptic curves over $\mathbb{Q}$ with torsion subgroup $Z_2 \times Z_8$. These curves are birationally equivalent to $y^2 = (1 - x^2)(1 - k^2 x^2)$ where $k = (a^4 - 6\,a^2\,b^2 + b^4)/(a^2 + b^2)^2$ for some integers a and b. We perform a computational analysis on the 3 148 208 curves corresponding to $|a|$, $|b| \le 5\,000$

The largest rank known in this family is $r = 3$; there are 13 examples in the literature. We exhibit 3 more. In an attempt to find such curves of larger rank, we perform a statistical analysis of the distribution of the ranks of the 2-Selmer groups.

1. Introduction

An elliptic curve is an object with two distinguishing features. First, such a curve has an equation that allows us to determine which of its points have rational coordinates. Second, its points form an abelian group so that one can ask about its algebraic structure. These two features complement one another. That is, given a few rational points one can find more such points using the group structure. Conversely, one can define the group structure by drawing lines through rational points.

If E is an elliptic curve over $\mathbb{Q}$, then the set $E(\mathbb{Q})$ of rational points forms a finitely generated abelian group. This means that there exists a finite group $E(\mathbb{Q})_{\text{tors}}$ and nonnegative integer r such that $E(\mathbb{Q}) \simeq E(\mathbb{Q})_{\text{tors}} \times \mathbb{Z}^r$. This integer r is called the (Mordell-Weil) rank. Every such elliptic curve with $E(\mathbb{Q})_{\text{tors}} \simeq Z_2 \times Z_8$ is birationally equivalent to

$$y^2 = (1 - x^2)(1 - k^2 x^2), \quad \text{where} \quad k = \frac{t^4 - 6t^2 + 1}{(t^2 + 1)^2}$$

for some rational number t. Currently, the highest known rank for this torsion subgroup is $r = 3$, and the known curves with this rank correspond to

$$t = \frac{5}{29}, \ \frac{18}{47}, \ \frac{15}{76}, \ \frac{74}{207}, \ \frac{47}{219}, \ \frac{19}{220}, \ \frac{87}{407}, \ \frac{143}{419}, \ \frac{17}{439}, \ \frac{145}{444}, \ \frac{159}{569}, \ \frac{230}{923}, \ \frac{223}{1012}.$$

We consider in detail the (Mordell-Weil) ranks of such curves. We focus on two questions:

- What are other examples of curves with $E(\mathbb{Q}) \simeq Z_2 \times Z_8 \times \mathbb{Z}^3$?
- What can we say about the distribution of such ranks?

To answer the first question, we use the following algorithm:

#1. Generate a list of candidate curves with $E(\mathbb{Q})_{\text{tors}} \simeq Z_2 \times Z_8$.

2000 *Mathematics Subject Classification.* Primary 14H52; Secondary 14J27.

Key words and phrases. Elliptic curves, Rank, Selmer group.

#2. Compute the ranks of the 2-Selmer groups of these curves. They will act as upper bounds for the ranks of the Mordell-Weil groups.

#3. Compute the ranks of the Mordell-Weil groups of the remaining curves with 2-Selmer rank at least 3.

We generated a list of curves for $t = \frac{a}{b}$ with $|a|, |b| \leq 5\,000$ containing 3 148 208 elliptic curves. Most of our computations were done on the high-powered computing clusters RedHawk at Miami University and Radon at Purdue University. Although we spread out our computations over 7 weeks, our processing time totaled almost 3 CPU years. We found three new examples of curves with Mordell-Weil group $Z_2 \times Z_8 \times \mathbb{Z}^3$:

$$t = \frac{19}{84}: \qquad Y^2 + X\,Y = X^3 - 7997298575851050590578125618 0\,X$$
$$+\ 8622476952474747423704354086825684364576400$$

$$t = \frac{101}{299}: \qquad Y^2 + X\,Y = X^3 - 977867541351362912053252014560183 00\,X$$
$$+\ 619261854407125870378759391947293497741496443889 0000$$

$$t = \frac{86}{333}: \qquad Y^2 + X\,Y = X^3 - 2508783953934745453167591832093118402 50\,X$$
$$+\ 14799795920221674932249605129107556895742994778089035609 32$$

We found four other curves which probably have rank 3 but we were only able to find two independent points of infinite order.

To answer the second question, we focused on an easier question, namely, the distribution of the ranks of the 2-Selmer groups. By considering our 3 148 208 curves with torsion subgroup $Z_2 \times Z_8$, we found the following distribution.

2-Selmer Rank	0	1	2	3	≥ 4
No. of Curves	461 127	1 110 462	1 004 658	450 939	142 001
Percentages	14.65%	35.27%	31.91%	14.32%	4.51%

We performed an analysis of this data. We realized that the average rank of a 2-Selmer group in our family is nearly 1.6, a value very different from the well-known average rank of 0.5 for Mordell-Weil groups. A plot of the distribution leads one to guess that it is Poisson, but we found through study of a generating function that this cannot be the case. We conclude with some surprising results and conjectures about this generating function which should shed more light on the distribution of these 2-Selmer ranks.

The authors would like to thank the Summer Undergraduate Mathematical Sciences Research Institute (SUMSRI) at Miami University for the opportunity to conduct this research. They would also like to thank the National Science Foundation and the National Security Agency for their funding; Residential Computing (ResComp) at Miami University and the Rosen Center for Advanced Computing (RCAC) at Purdue University for use of their machines; Vasant Waikar and Ashley Swandby for helpful conversations; and Edray Goins and Maria Salcedo for their constant support and guidance.

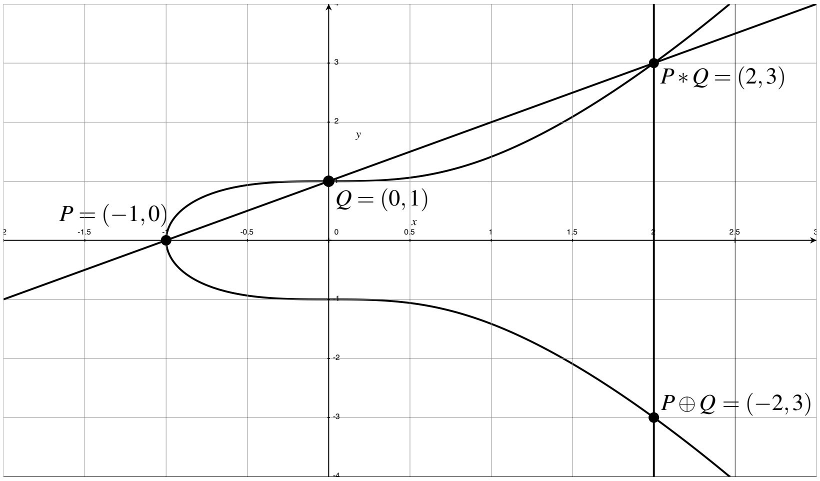


Figure 1. The Group Law on an Elliptic Curve

2. The Mordell-Weil Group

An elliptic curve is a curve that is birationally equivalent to

$$E : Y^2 = X^3 + AX + B,$$

where $4A^3 + 27B^2 \neq 0$. We are interested in the case where A and B are integers. Viewed as a subset of projective space, we consider the set $E(\mathbb{Q})$ of rational points joined with the point at infinity $\mathcal{O}$. Let $P, Q \in E(\mathbb{Q})$ and denote $P * Q$ as the third point of intersection of E and the line through P and Q. Now define $P \oplus Q = (P * Q) * \mathcal{O}$, i.e., the reflection of $P * Q$ about the x-axis. See Figure 1 for an example of this geometry. The set $E(\mathbb{Q})$ under the operation $\oplus$ forms an abelian group with identity element $\mathcal{O}$ and inverses $[-1]P = P * \mathcal{O}$.

In 1901, Henri Poincaré [8] conjectured that the abelian group $E(\mathbb{Q})$ is finitely generated. Louis Mordell proved this in 1922.

Theorem 1 (Mordell, [7])**.** *If E is an elliptic curve over $\mathbb{Q}$, then the abelian group $E(\mathbb{Q})$ is finitely generated. Furthermore, there exists a finite group $E(\mathbb{Q})_{\text{tors}}$ and a nonnegative integer r such that*

$$E(\mathbb{Q}) \simeq E(\mathbb{Q})_{\text{tors}} \times \mathbb{Z}^r.$$

This nonnegative integer r is called the (Mordell-Weil) rank of E. The torsion subgroup $E(\mathbb{Q})_{\text{tors}}$ is the set of points of finite order. Barry Mazur completely classified the structure of this subgroup in 1977.

Theorem 2 (Mazur, [6])**.** *If E is an elliptic curve over $\mathbb{Q}$, then $E(\mathbb{Q})_{\text{tors}}$ is isomorphic to one of the following 15 groups:*

(i) *Z_n, with $1 \leq n \leq 10$ or $n = 12$,*
(ii) *$Z_2 \times Z_{2m}$, with $1 \leq m \leq 4$.*

In this paper, we focus specifically on curves with torsion subgroup $Z_2 \times Z_8$.

In contrast to the torsion subgroup, the rank is not well understood. It is widely believed that for each of the torsion subgroups T given in Mazur's theorem, the

Table 1. Rank Records

Torsion Subgroup T	Lower Bound for $B(T)$	Author(s)
Z_1	28	Elkies (2006)
Z_2	18	Elkies (2006)
Z_3	13	Eroshkin (2007)
Z_4	12	Elkies (2006)
Z_5	6	Dujella - Lecacheux (2001)
Z_6	7	Dujella (2001, 2006)
Z_7	5	Dujella - Kulesz (2001) Elkies (2006)
Z_8	6	Elkies (2006)
Z_9	3	Dujella (2001) MacLeod (2004) Eroshkin (2006)
Z_{10}	4	Dujella (2005) Elkies (2006)
Z_{12}	3	Dujella (2001, 2005, 2006) Rathbun (2003, 2006)
$Z_2 \times Z_2$	14	Elkies (2005)
$Z_2 \times Z_4$	8	Elkies (2005)
$Z_2 \times Z_6$	6	Elkies (2006)
$Z_2 \times Z_8$	3	Connell (2000) Dujella (2000, 2001, 2006) Campbell - Goins (2003) Rathbun (2003, 2006)

ranks of elliptic curves E with $E(\mathbb{Q})_{\text{tors}} \simeq T$ are unbounded. In attempting to verify this belief, one approach is to assume there is an absolute bound $B(T)$, and then search for a curve with torsion subgroup T and rank r greater than this bound. Since March 2001, Andrej Dujella [4] has been tracking the lower bounds of $B(T)$ by recording the highest known rank for each T. Table 1 summarizes the current records. The highest known rank for $T \simeq Z_2 \times Z_8$ is $r = 3$.

We review how one computes the rank of the Mordell-Weil group of an elliptic curve. We give a simplified exposition of that found in [10]. More information can also be found in [5].

Theorem 3 (Complete 2-Descent, [10]). *Let E be an elliptic curve with the following two properties:*

- *E is defined over $\mathbb{Q}$, i.e., we have the equation $Y^2 = X^3 + A\,X + B$ where A and B are integers.*
- *The cubic $X^3 + A\,X + B$ has three rational roots, e_1, e_2, e_3.*

Then we have the following:

(i) *The "connecting homomorphism"*

$$\delta_E : \frac{E(\mathbb{Q})}{2E(\mathbb{Q})} \to \frac{\mathbb{Q}^\times}{(\mathbb{Q}^\times)^2} \times \frac{\mathbb{Q}^\times}{(\mathbb{Q}^\times)^2},$$

defined as that map which sends

$$P = (X, Y) \mapsto \Big(X - e_1 \mod (\mathbb{Q}^\times)^2, \quad X - e_2 \mod (\mathbb{Q}^\times)^2 \Big)$$

is an injective group homomorphism. Its image lies in the finite group

$$G = \left\{ (d_1,\, d_2) \in \frac{\mathbb{Q}^\times}{(\mathbb{Q}^\times)^2} \times \frac{\mathbb{Q}^\times}{(\mathbb{Q}^\times)^2} \;\middle|\; \begin{array}{c} d_i = \pm \ell_1^{a_{i1}} \cdots \ell_r^{a_{ir}}, \textit{ where} \\ \ell_j \textit{ divides } -16(4\,A^3 + 27\,B^2) \end{array} \right\}.$$

(ii) *For each $d = (d_1,\, d_2)$ in G, consider the "principle homogeneous space"*

$$C_d: \quad d_1\, u^2 - d_2\, v^2 = e_2 - e_1, \quad d_1\, u^2 - d_1\, d_2\, w^2 = e_3 - e_1.$$

Assume $P = (X,\, Y)$ is in $E(\mathbb{Q})$. Then the point

$$(u,\, v,\, w) = \left(\sqrt{\frac{X - e_1}{d_1}},\, \sqrt{\frac{X - e_2}{d_2}},\, \frac{Y}{\sqrt{d_1\,(X - e_1) \cdot d_2\,(X - e_2)}} \right)$$

is in $C_d(\mathbb{Q})$ for $d = \delta_E(P)$. In particular, $C_d(\mathbb{Q}) \neq \emptyset$. Conversely, assume $C_d(\mathbb{Q}) \neq \emptyset$. Then $P = (d_1\, u^2 + e_1,\, d_1\, d_2\, u\, v\, w)$ is in $E(\mathbb{Q})$ for $(u,\, v,\, w) \in C_d(\mathbb{Q})$. In particular, $d = \delta_E(P)$.

The assumptions above imply that the Mordell-Weil group of the elliptic curve is $E(\mathbb{Q}) \simeq Z_2 \times Z_{2m} \times \mathbb{Z}^r$. In general, this is an infinite group, but the quotient

$$\frac{E(\mathbb{Q})}{2E(\mathbb{Q})} \simeq \big\{ d \in G \;\big|\; C_d(\mathbb{Q}) \neq \emptyset \big\} \simeq {Z_2}^{r+2}$$

is a finite group. Every point P in this quotient corresponds to a principle homogeneous space C_d containing at least one rational point. Hence, to compute the rank, we must count the number of curves such that $C_d(\mathbb{Q}) \neq \emptyset$.

This identification using homogeneous spaces implies the following composition of homomorphisms:

$$\begin{tikzcd} \{0\} \arrow[r] & \dfrac{E(\mathbb{Q})}{2\,E(\mathbb{Q})} \arrow[r, "\delta_E"] & \mathrm{Sel}^{(2)}(E/\mathbb{Q}) \arrow[r] & Ш(E/\mathbb{Q})[2] \arrow[r] & \{0\} \end{tikzcd}$$

where we define

$$\mathrm{Sel}^{(2)}(E/\mathbb{Q}) = \left\{ d \in G \;\middle|\; C_d(\mathbb{R}) \neq \emptyset \text{ and } C_d(\mathbb{Q}_p) \neq \emptyset \text{ for all primes } p \right\}$$

as the 2-Selmer group of E, and

$$Ш(E/\mathbb{Q})[2] = \left\{ d \in G \;\middle|\; \begin{array}{c} C_d(\mathbb{R}) \neq \emptyset \text{ and } C_d(\mathbb{Q}_p) \neq \emptyset \text{ for all primes } p \\ \text{but } C_d(\mathbb{Q}) = \emptyset \end{array} \right\}$$

as the 2-torsion subgroup of the Shafarevich-Tate group of E. One may think of the 2-Selmer group as the collection of principle homogeneous spaces which have points locally, and the Shafarevich-Tate group as the collection of principle homogeneous spaces which do not have points globally. In practice, the 2-Selmer group is easy to compute, but the Shafarevich-Tate group is quite mysterious. Nonetheless, these are finite abelian groups with orders

$$\left| \frac{E(\mathbb{Q})}{2E(\mathbb{Q})} \right| = 2^{r+2}, \quad \left| \mathrm{Sel}^{(2)}(E/\mathbb{Q}) \right| = 2^{s+2}, \quad \text{and} \quad \left| Ш(E/\mathbb{Q})[2] \right| = 2^{s-r}.$$

We denote r as the "Mordell-Weil" rank of E and s as the "2-Selmer" rank of E. Note that $r \leq s$; this gives an upper bound for the Mordell-Weil rank. We will use this observation later.

3. Curves with Torsion Subgroups $Z_2 \times Z_4$ and $Z_2 \times Z_8$

The aim of this paper is to examine the ranks of curves with $E(\mathbb{Q})_{\text{tors}} \simeq Z_2 \times Z_8$. To do so we must first classify these curves. The following theorems form a basis for the focus of this project.

Theorem 4 (Goins, [5])**.** *Fix a rational number $k \neq -1, 0, 1$ and consider the curve*

$$E : y^2 = (1 - x^2)(1 - k^2 x^2).$$

(i) *E is an elliptic curve.*
(ii) *Its torsion subgroup is either $Z_2 \times Z_4$ or $Z_2 \times Z_8$.*
(iii) *Any elliptic curve over $\mathbb{Q}$ with torsion subgroup $Z_2 \times Z_4$ or $Z_2 \times Z_8$ is birationally equivalent to E for some rational number k.*

For example, consider the elliptic curve

$$(*) \qquad \begin{aligned} E : Y^2 + X\,Y = X^3 &- 79972985758510505905781256180\,X \\ &+ 862247695247474742370435408682568436457640 0 \end{aligned}$$

with Modell-Weil group $E(\mathbb{Q}) \simeq Z_2 \times Z_8 \times \mathbb{Z}^3$. This curve is birationally equivalent to the quartic curve in the statement of Theorem 4 when $k = \frac{34634161}{55011889}$. This can be seen by using the substitutions

$$\begin{aligned} X &= \frac{37080\,(1669429617\,x - 7827678077)}{x - 1} \qquad \text{and} \\ Y &= \frac{18540}{(x-1)^2} \left(\begin{aligned} - 1669429617\,x^2 + 9497107694\,x \\ + 169388440357970470\,y - 7827678077. \end{aligned} \right) \end{aligned}$$

Theorem 4 establishes a birational equivalence between curves E with $E(\mathbb{Q})_{\text{tors}} \simeq Z_2 \times Z_4$ or $Z_2 \times Z_8$ and rational numbers k. To distinguish between the two cases, further restrictions may be placed upon k. The following theorem utilizes this idea to classify curves with torsion subgroup $Z_2 \times Z_8$.

Theorem 5 (Goins, [5])**.** *Say that E is an elliptic curve over $\mathbb{Q}$ with torsion subgroup $Z_2 \times Z_8$.*

(i) *There exists a rational number t such that E is birationally equivalent to the curve*

$$y^2 = (1 - x^2)(1 - k^2 x^2), \quad \text{where} \quad k = \frac{t^4 - 6t^2 + 1}{(t^2+1)^2}.$$

(ii) *Moreover, upon writing $k = \frac{p}{q}$ for some integers p and q, the quartic curve above is birationally equivalent to the cubic curve*

$$Y^2 = X^3 - 27(p^4 + 14p^2q^2 + q^4)X - 54(p^6 - 33p^4q^2 - 33p^2q^4 + q^6).$$

This result allows us to use Theorem 3 rather explicitly. Indeed, we have a cubic curve with integer coefficients such that the cubic $X^3 + A\,X + B$ has three rational roots

$$e_1 = -3\left(p^2 - 6\,p\,q + q^2\right), \quad e_2 = 6\left(p^2 + q^2\right), \quad \text{and} \quad e_3 = -3\left(p^2 + 6\,p\,q + q^2\right).$$

Proof. We will show part (ii) since (i) is shown in [5]. Given a point (x, y) on the quartic curve, the substitutions

$$X = \frac{3(5p^2 - q^2)x + 3(5q^2 - p^2)}{x - 1} \quad \text{and} \quad Y = \frac{54q(p^2 - q^2)y}{(x-1)^2}$$

Table 2. Curves with Mordell-Weil group $Z_2 \times Z_8 \times \mathbb{Z}^3$

Author(s)	Year	Parameter t
Connell	2000	$\frac{5}{29}$
Dujella	2000	$\frac{5}{29}$
Dujella	2001	$\frac{18}{47}$
Campbell, Goins	2003	$\frac{15}{76}$
Rathbun	2003	$\frac{47}{219}$
Campbell, Goins, Watkins	2005	$\frac{19}{220}$
Dujella	2006	$\frac{87}{407}$
Dujella	2006	$\frac{143}{419}$
Dujella	2006	$\frac{145}{444}$
Rathbun	2006	$\frac{74}{207}$
Rathbun	2006	$\frac{17}{439}$
Rathbun	2006	$\frac{159}{569}$
Dujella, Rathbun	2006	$\frac{230}{923}$
Dujella, Rathbun	2006	$\frac{223}{1012}$

yield a point (X, Y) on the cubic curve. Conversely, given a point (X, Y) on the cubic curve, the subsitutions

$$x = \frac{X - 3(5q^2 - p^2)}{X - 3(5p^2 - q^2)} \quad \text{and} \quad y = \frac{6(p^2 - q^2)Y}{q(X - 3(5p^2 - q^2))^2}$$

yield a point (x, y) on the quartic curve. □

We consider the family of elliptic curves $E : y^2 = (1 - x^2)(1 - k^2x^2)$, where $k = \frac{t^4 - 6t^2 + 1}{(t^2+1)^2}$. Dujella's website [4] lists thirteen curves with Mordell-Weil group $E(\mathbb{Q}) \simeq Z_2 \times Z_8 \times \mathbb{Z}^3$. Using the substitutions in the proof of the theorem above, we may list corresponding values of t; they may be found in Table 2.

In the exposition that follows, we explain how we found three previously unknown elliptic curves with the same Mordell-Weil group. They correspond the the following values of t:

$$\frac{19}{84}, \quad \frac{101}{299}, \quad \text{and} \quad \frac{86}{333}.$$

Specifically, the curve of rank 3 in equation $(*)$ is birationally equivalent to the quartic curve in the statement of Theorem 5 when $t = \frac{19}{84}$.

4. Searching for Curves with Large Rank

To search for an elliptic curve defined over $\mathbb{Q}$ with torsion subgroup $Z_2 \times Z_8$ and rank $r \geq 3$, we use the following algorithm:

#1. Generate a list of candidate curves with $E(\mathbb{Q})_{\text{tors}} \simeq Z_2 \times Z_8$.

#2. Compute the ranks of the 2-Selmer groups of these curves. They will act as upper bounds for the ranks of the Mordell-Weil groups.

#3. Compute the ranks of the Mordell-Weil groups of the remaining curves with 2-Selmer rank at least 3.

Here we discuss these steps in further detail and describe our results.

Step 1: Generate a List of Candidate Elliptic Curves. According to Theorem 5, every elliptic curve E defined over $\mathbb{Q}$ with torsion subgroup $Z_2 \times Z_8$ is birationally equivalent to the quartic $y^2 = (1-x^2)(1-k^2x^2)$, where $k = \frac{t^4-6t^2+1}{(t^2+1)^2}$ for some rational number $t = \frac{a}{b}$. We generate a list of candidate curves based on the integer values a and b. By considering symmetries of the expression $k = \frac{t^4-6t^2+1}{(t^2+1)^2}$, we can eliminate redundant curves.

Proposition 6. *Fix a rational number $k \neq -1, 0, 1$ and consider the curve*

$$E : y^2 = (1-x^2)(1-k^2x^2).$$

(i) *E is birationally equivalent to*

$$E' : Y^2 = (1-X^2)\left(1 - \frac{1}{k^2}X^2\right).$$

In particular, we may assume that $0 < k < 1$.

(ii) *Suppose that $k = \frac{t^4-6t^2+1}{(t^2+1)^2}$ for some $t = \frac{a}{b}$. We can choose integers a and b such that $0 < (1+\sqrt{2})a < b$.*

Proof. First we show (i). Given a point (x, y) on E, the point $(X, Y) = (kx, y)$ is on E'. Conversely, given a point (X, Y) on E', the point $(x, y) = (\frac{1}{k}X, Y)$ is on E. Therefore, E and E' are birationally equivalent. It is clear that we may choose $k > 0$. If $k < 1$ we are done. If $k > 1$ then $\frac{1}{k} < 1$, and because E is equivalent to E' we are done.

Next we show (ii). Write $k = \frac{a^4-6a^2b^2+b^4}{(a^2+b^2)^2}$. It is clear that we may assume $0 \leq a \leq b$. If $a = 0$ then $k = 1$. If $a = b$ then $k = -1$. Therefore, we may assume $0 < a < b$. The mapping ϕ: $(a, b) \mapsto (b-a, b+a)$ sends k to $-k$; ϕ is an involution since ϕ^2 is the identity. Denote $c = b-a$ and $d = b+a$. By assumption, $0 < a < b$, hence $0 < c < d$. Assume for the moment that $(1+\sqrt{2})c > d$. This implies

$$\frac{b-a}{b+a} = \frac{c}{d} > \frac{1}{1+\sqrt{2}} \quad \Longrightarrow \quad (b-a)(1+\sqrt{2}) > b+a.$$

After a bit of algebra we find

$$b\sqrt{2} > a(2+\sqrt{2}) \quad \Longrightarrow \quad \frac{a}{b} < \frac{1}{1+\sqrt{2}}.$$

Hence $0 < (1+\sqrt{2})a < b$, so we are done. On the other hand, if $(1+\sqrt{2})c < d$ then we choose c and d instead of a and b. □

A pseudocode for our algorithm is as follows.

```
#1. INPUT: Bound $N$
#2. For integers $a$ and $b$ satisfying $0 < (1+\sqrt{2})a < b \leq N$
      a. Define $t = \frac{a}{b}$ and $k = \frac{p}{q}$, as well as the following integers
           $p = a^4 - 6a^2b^2 + b^4$      $A = -27(p^4 + 14p^2q^2 + q^4)$
           $q = (a^2+b^2)^2$              $B = -54(p^6 - 33p^4q^2 - 33p^2q^4 + q^6)$
      b. Record the elliptic curve $Y^2 = X^3 + AX + B$ to a list.
```

Table 3. Lists of Candidate Curves with $E(\mathbb{Q})_{\text{tors}} \simeq Z_2 \times Z_8$

Bound N	No. of Candidate Curves	Processing Time (seconds)
1 000	126 003	23
2 000	503 923	55
3 000	1 133 364	131
4 000	2 014 563	246
5 000	3 148 208	612

Table 4. 2-Selmer Ranks of Curves with $E(\mathbb{Q})_{\text{tors}} \simeq Z_2 \times Z_8$

Bound N	Processing Time (hh:mm:ss)	No. of Curves with $s < 3$	No. of Curves with $s \geq 3$
1 000	384:53:19	105 160	20 843
2 000	1797:02:13	415 776	88 147
3 000	4108:00:01	931 083	202 281
4 000	8334:36:17	1 650 867	363 696
5 000	11421:51:11	2 576 247	571 961

#3. OUTPUT: List of elliptic curves

We implemented this using `Maple` on desktop PCs. Information regarding various bounds and runtimes can be found in Table 3. For example, a bound of $N = 5\,000$ generated 3 148 208 curves in about ten minutes. We recorded the elliptic curves $Y^2 = X^3 + AX + B$ in the form `[0,0,0,A,B]` for processing in Step 2.

Step 2: Determine the 2-Selmer Ranks. Computing the ranks r of the Mordell-Weil groups $E(\mathbb{Q})$ can be difficult – especially if the Shafarevich-Tate group $\text{Ш}(E/\mathbb{Q})$ is nontrivial. Since we have some 3 million candidate curves to consider, we eliminate many curves by computing the ranks s of their 2-Selmer groups $\text{Sel}^{(2)}(E/\mathbb{Q})$. The idea is to use the latter as upper bounds for the former.

We use John Cremona's `mwrank` [3] to do this. The option `-s` in `mwrank` computes the 2-Selmer ranks s relatively quickly. Miami University has a 128-node high-powered computing cluster known as RedHawk. We divided our lists of candidate curves into separate files, running `mwrank` on these nodes simultaneously. Runtimes for the various lists of candidate curves as listed in Table 3 can be found in Table 4. For example, given a bound of $N = 5\,000$ it took approximately 16 CPU months (cumulative) to compute the 2-Selmer ranks of some 3 million curves.

Step 3: Compute the Mordell-Weil Ranks. Our final step is to compute the Mordell-Weil ranks of the curves from Step 2. This should take the longest because `mwrank` runs slowly on elliptic curves with large coefficients. Recall that the 2-Selmer rank s is an upper bound for the Mordell-Weil rank r. We found it useful to break our lists up according to both the bounds N and their 2-Selmer ranks s. The number of curves in these files can be found in Table 5; the relative percentages can be found in Table 6. Notice that for $N \leq 5\,000$ we found $s \leq 7$.

Due to time constraints, we focused on those curves with bound $N = 1\,000$ and 2-Selmer rank $s = 3$. We know that at least 12 curves in this list have Mordell-Weil

TABLE 5. Distribution of Curves with $\mathrm{Sel}^{(2)}(E/\mathbb{Q}) \simeq {Z_2}^{s+2}$

Bound N	1 000	2 000	3 000	4 000	5 000
$s = 0$	19 309	75 384	167 581	296 135	461 127
$s = 1$	45 807	179 361	401 351	711 392	1 110 462
$s = 2$	40 044	161 031	362 152	643 340	1 004 658
$s = 3$	16 933	70 481	160 695	287 682	450 939
$s = 4$	3 550	15 845	36 956	67 289	106 791
$s = 5$	338	1 707	4 370	8 208	13 371
$s = 6$	22	112	256	509	839
$s = 7$	0	2	4	8	21
$s \geq 8$	0	0	0	0	0
Total	126 003	503 923	1 133 364	2 014 563	3 148 208

TABLE 6. Relative Distribution of Curves with $\mathrm{Sel}^{(2)}(E/\mathbb{Q}) \simeq {Z_2}^{s+2}$

Bound N	1 000	2 000	3 000	4 000	5 000
$s = 0$	0.153242	0.149594	0.147862	0.146997	0.146473
$s = 1$	0.363539	0.355929	0.354124	0.353125	0.352728
$s = 2$	0.317484	0.319555	0.319537	0.319345	0.319121
$s = 3$	0.134386	0.139865	0.141786	0.142801	0.143237
$s = 4$	0.028174	0.031443	0.032607	0.033401	0.033921
$s = 5$	0.002682	0.003387	0.003856	0.004074	0.004247
$s = 6$	0.000175	0.000222	0.000226	0.000253	0.000267
$s = 7$	0.000000	0.000004	0.000004	0.000004	0.000007
$s \geq 8$	0.000000	0.000000	0.000000	0.000000	0.000000

rank $r = 3$; see Table 2. We distributed these 16 933 candidate curves into 100 files for processing at RedHawk. We used the following options in `mwrank`:

- `-q` for quiet output
- `-v 0` for minimum verbosity
- `-b 15` to perform the best possible search for rational points on the principle homogeneous spaces C_d
- `-p 250` to increase the precision since the coefficients of the elliptic curves are rather large
- `-l` to list the generators for $E(\mathbb{Q})/E(\mathbb{Q})_{\mathrm{tors}}$
- `-S` to ensure that we do indeed have the generators

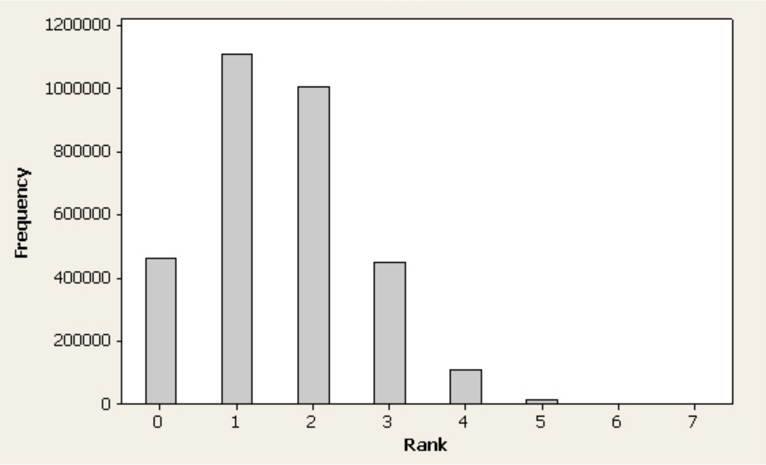
As of the time of publication, RedHawk processed only 11 368 of the 16 933 candidate curves after approximately 2 CPU years (cumulative). We found three new elliptic curves E with Mordell-Weil group $E(\mathbb{Q}) \simeq Z_2 \times Z_8 \times \mathbb{Z}^3$, i.e., curves with Mordell-Weil rank $r = 3$. They correspond to

$$\frac{19}{84}, \quad \frac{101}{299}, \quad \text{and} \quad \frac{86}{333}.$$

Table 7. Mordell-Weil Ranks of $Y^2 + X\,Y = X^3 + A\,X + B$

Parameter t	Rank r	Coefficients A and B
$\frac{19}{84}$	3	$-79972985758510505905781256180$ $862247695247474742370435408682568436457640 0$
$\frac{101}{299}$	3	$-97786754135136291205325201456018300$ $6192618544071258703787593919472934977414964438890000$
$\frac{86}{333}$	3	$-2508783953934745453167591832093118402 50$ $147997959202221674932249605129107556895742994778089035 60932$
$\frac{12}{65}$	≥ 2	$-1569352275530075968688655080$ $23891634009988359905502318672762947150400$
$\frac{21}{92}$	≥ 2	$-3394388652909278601670265324 60$ $7533575475299173812469601053163425555493467 2$
$\frac{9}{296}$	≥ 2	$-71813598680248384341084284771096244120$ $2342384302041141813702521859646228641128533374139589904 00$
$\frac{65}{337}$	≥ 2	$-162119749537661228753090410523372412 0$ $79268859879186507385709442635874834355179816874619040 0$

Figure 2. Histogram of Ranks of 2-Selmer Groups for $N = 5\,000$



We found four other curves with Mordell-Weil rank $2 \leq r \leq 3$ but unfortunately `mwrank` could not determine the exact value. They correspond to

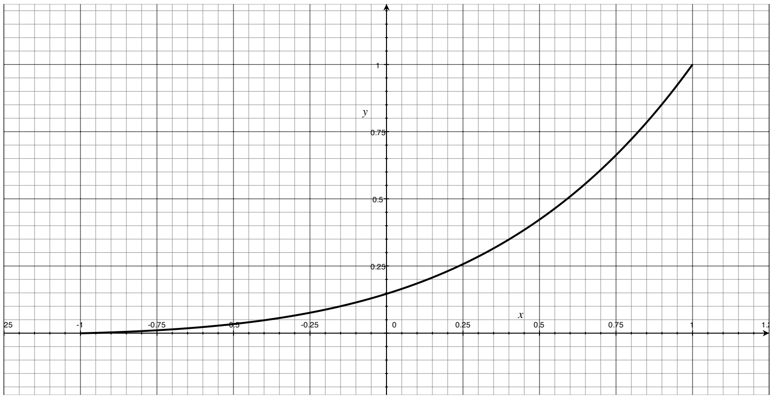
$$\frac{12}{65}, \quad \frac{21}{92}, \quad \frac{9}{296}, \quad \text{and} \quad \frac{65}{337}.$$

Equations for the corresponding curves can be found in Table 7. We did not find any curves with Mordell-Weil rank $r > 3$.

5. Distribution of 2-Selmer Ranks

We noticed a curious pattern by looking at the 2-Selmer ranks. Consider Figure 2. Upon looking at the rows of Table 6, we observe that the relative distributions appear to tend to a limit. We make some formal definitions.

FIGURE 3. Graph of $f_{\mathrm{sel}}(z)$



Denote $\mathcal{F}$ as the family of all elliptic curves E defined over $\mathbb{Q}$ with torsion subgroup $Z_2 \times Z_8$, i.e., curves which are birationally equivalent to

$$y^2 = (1-x^2)(1-k^2x^2), \quad \text{where} \quad k = \frac{a^4 - 6\,a^2\,b^2 + b^4}{(a^2+b^2)^2}$$

for some integers a and b. For each nonnegative integer N, denote $\mathcal{F}_N$ as those $E \in \mathcal{F}$ with $|a|,\ |b| \leq N$. For each nonnegative integer r, denote $\mathcal{F}_{\mathrm{mw}}(r)$ as those $E \in \mathcal{F}$ with Mordell-Weil rank r, i.e.,

$$\frac{E(\mathbb{Q})}{2E(\mathbb{Q})} \simeq {Z_2}^{r+2}.$$

Similarly, for each nonnegative integer s, denote $\mathcal{F}_{\mathrm{sel}}(s)$ as those $E \in \mathcal{F}$ with 2-Selmer rank s, i.e.,

$$\mathrm{Sel}^{(2)}(E/\mathbb{Q}) \simeq {Z_2}^{s+2}.$$

Conjecture 7. *Let* $*$ *be either* mw *or* sel.

(i) *The function*

$$f_*(z) = \lim_{N\to\infty} \sum_{n=0}^{\infty} \frac{\#(\mathcal{F}_N \cap \mathcal{F}_*(n))}{\#\mathcal{F}_N}\ z^n$$

is uniformly convergent in the unit disk $|z| < 1$.

(ii) $f_*(1) = 1$ *and* $f_*(-1) = 0$.

This conjecture allows us to think of the coefficients of these generating functions as probabilities. Our computations of the 2-Selmer ranks suggest that one of the functions looks like

$$\begin{aligned} f_{\mathrm{sel}}(z) = {} & 0.146473 + 0.352728\,z + 0.319121\,z^2 + 0.143237\,z^3 \\ & + 0.033921\,z^4 + 0.004247\,z^5 + 0.000267\,z^6 + 0.000007\,z^7 + \cdots . \end{aligned}$$

A graph of this function can be found in Figure 3. Unfortunately, we do not have enough data to compute $f_{\mathrm{mw}}(z)$.

Proposition 8. *Assume Conjecture 7.*

(i) *The data in Table 5 does not follow a Poisson distribution. The coefficients of $f_{sel}(z)$ do not follow a Poisson distribution.*

(ii) *Consider the family $\mathcal{F}$ of elliptic curves E defined over $\mathbb{Q}$ with torsion subgroup $E(\mathbb{Q}_{tors}) \simeq Z_2 \times Z_8$. The probability of a curve in this family having even 2-Selmer rank is the same as that of a curve in this family having odd 2-Selmer rank.*

We recall the definition of a Poisson distribution. For each nonnegative integer n, let $O(n)$ denote the number of occurrences of n (or perhaps the nth coefficient of a certain generating function). We say that these occurrences follow a Poisson distribution if there exists λ (i.e., the mean) such that

$$\frac{O(n)}{O(0)+O(1)+\cdots} = \frac{\lambda^n}{n!}\, e^{-\lambda}$$

for each nonnegative integer n.

Proof. First we prove (i). Table 5 lists the observed frequencies $O(s)$ for 2-Selmer ranks satisfying $0 \le s \le m(N)$, where $N \le 5\,000$ and $m(N)$ is the largest rank for which there exists a curve in this bound. We perform a "chi-square goodness of fit test" to see how well these observed frequencies match the expected frequencies

$$E(s) = \left[\sum_{s=0}^{m(N)} O(s)\right] \frac{\lambda^s}{s!}\, e^{-\lambda} \quad \text{where we estimate} \quad \lambda = \frac{\sum_{s=0}^{m(N)} s \cdot O(s)}{\sum_{s=0}^{m(N)} O(s)}.$$

Consider the "chi-square value" of our data:

$$\chi^2 = \sum_{s=0}^{m(N)} \frac{[O(s)-E(s)]^2}{E(s)}.$$

For each nonnegative α (the "acceptable error") and nonnegative df (the "degrees of freedom"), denote $\chi^2_{\alpha,\mathrm{df}}$ as that real number such that

$$\alpha = \frac{1}{\Gamma\left(\mathrm{df}/2\right)} \int_{\chi^2_{\alpha,\,\mathrm{df}}}^{\infty} \left(\frac{x}{2}\right)^{\mathrm{df}/2} e^{-x/2}\, \frac{dx}{x}$$

in terms of the Gamma function $\Gamma(n)$. For example, if $\alpha = 5\%$ and $\mathrm{df} = 6$, then $\chi^2_{\alpha,\mathrm{df}} = 12.592$. Our data has $\mathrm{df} = m(N) - 1$ degrees of freedom because there are $m(N)+1$ possible ranks as s ranges between 0 and $m(N)$, but we are constrained by the total number of curves – namely $\sum_{s=0}^{m(N)} O(s)$ – and our estimation of λ as the average rank observed from our data. We accept the hypothesis that the observed frequencies, $O(s)$, follow a Poisson distribution only if $\chi^2 \le \chi^2_{\alpha,\,\mathrm{df}}$ when $\alpha = 5\%$. Table 8 lists the relevant data, thereby showing that the distribution is not Poisson.

Now assume that the coefficients of $f_{\mathrm{sel}}(z)$ do indeed follow a Poisson distribution. As $f_{\mathrm{sel}}(z)$ is uniformly convergent for $|z| < 1$, we may interchange the limit with the summation. That is, for each nonnegative integer s,

$$\lim_{N\to\infty} \frac{\#(\mathcal{F}_N \cap \mathcal{F}_{\mathrm{sel}}(s))}{\#\mathcal{F}_N} = \frac{\lambda^s}{s!}\, e^{-\lambda}$$

Table 8. Chi-Square Distribution for 2-Selmer Ranks

Bound N	$m(N)$	Observed Average λ	χ^2	$\chi^2_{\alpha,\,\mathrm{df}}$
1 000	6	1.529456	7 700.072	11.070
2 000	7	1.558704	29 761.771	12.592
3 000	7	1.569643	65 653.675	12.592
4 000	7	1.575738	115 008.433	12.592
5 000	7	1.579246	177 788.496	12.592

for some λ. This implies

$$f_{\mathrm{sel}}(z) = \sum_{s=0}^{\infty} \left(\frac{\lambda^s}{s!} \, e^{-\lambda} \right) z^s = \left(\sum_{s=0}^{\infty} \frac{\lambda^s}{s!} \, z^s \right) e^{-\lambda} = e^{\lambda(z-1)}.$$

The exponential is a nonzero function, which contradicts $f_{\mathrm{sel}}(-1) = 0$.

Now we prove (ii). It suffices to show that the probability of $E \in \mathcal{F}$ having even 2-Selmer rank s is $\frac{1}{2}$. Consider the following expression:

$$\frac{f_{\mathrm{sel}}(z) + f_{\mathrm{sel}}(-z)}{2} = \sum_{s \text{ even}} \left(\lim_{N \to \infty} \frac{\#(\mathcal{F}_N \cap \mathcal{F}_{\mathrm{sel}}(s))}{\#\mathcal{F}_N} \right) z^s.$$

As $z \to 1$, we have the desired probability

$$\sum_{s \text{ even}} \left(\lim_{N \to \infty} \frac{\#(\mathcal{F}_N \cap \mathcal{F}_{\mathrm{sel}}(s))}{\#\mathcal{F}_N} \right) = \frac{f_{\mathrm{sel}}(1) + f_{\mathrm{sel}}(-1)}{2} = \frac{1}{2}.$$

□

It appears that the average value of the 2-Selmer rank for $E \in \mathcal{F}$ is greater than 1.5. Indeed, we have the following identity which gives this average value:

$$\lim_{N \to \infty} \frac{\sum_{s=0}^{m(N)} s \cdot O(s)}{\sum_{s=0}^{m(N)} O(s)} = \lim_{N \to \infty} \sum_{s=0}^{\infty} s \cdot \frac{\#(\mathcal{F}_N \cap \mathcal{F}_{\mathrm{sel}}(s))}{\#\mathcal{F}_N} = f'_{\mathrm{sel}}(1).$$

We approximate this to have the value 1.579246, as in Table 8. In contrast, it is widely believed that the average value of the Mordell-Weil rank is 0.5.

References

[1] Terris D. Brooks, Elizabeth A. Fowler, Katherine C. Hastings, Danielle L. Hiance, and Matthew A. Zimmerman. Elliptic Curves with Torsion Subgroup $Z_2 \times Z_8$: Does a Rank 4 Curve Exist? *SUMSRI Journal*, 2006.

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

[3] John Cremona. `mwrank` and related programs for elliptic curves over $\mathbb{Q}$. `http://www.maths.nott.ac.uk/personal/jec/mwrank/`, 2006.

[4] Andrej Dujella. High rank elliptic curves with prescribed torsion. `http://www.math.hr/~duje/tors/tors.html`, 2007.

[5] Edray Herber Goins. SUMSRI Number Theory Research Seminar Lecture Notes. In preparation, 2007.

[6] B. Mazur. Modular curves and the Eisenstein ideal. *Inst. Hautes Études Sci. Publ. Math.*, (47):33–186 (1978), 1977.

[7] L. J. Mordell. *Diophantine equations*. Academic Press, London, 1969.

[8] Henri Poincaré. Sur les propriétés arithmétiques des courbes algébriques. *Journal de mathématiques pures et appliquées*, 7(5):161–233, 1901.

[9] Karl Rubin and Alice Silverberg. Ranks of elliptic curves. *Bull. Amer. Math. Soc. (N.S.)*, 39(4):455–474 (electronic), 2002.

[10] Joseph H. Silverman. *The arithmetic of elliptic curves*. Springer-Verlag, New York-Berlin, 1986.

[11] John Stillwell. Elliptic curves. *Amer. Math. Monthly*, 102(9):831–837, 1995.

University of Puerto Rico, Humacao, PR 00791
E-mail address: jeflro@mate.uprh.edu

Savannah State University, Savannah, GA 31404
E-mail address: jonesk3@savstate.edu

John Carroll University, University Heights, OH 44118
E-mail address: arollick08@jcu.edu

Purdue University, West Lafayette, IN 47906
E-mail address: jweigand@math.purdue.edu