MATHEMATICS OF COMPUTATION
Volume 78, Number 268, October 2009, Pages 2397–2425
S 0025-5718(09)02253-4
Article electronically published on June 8, 2009

COMPUTATIONAL VERIFICATION OF THE BIRCH AND SWINNERTON-DYER CONJECTURE FOR INDIVIDUAL ELLIPTIC CURVES

GRIGOR GRIGOROV, ANDREI JORZA, STEFAN PATRIKIS, WILLIAM A. STEIN, AND CORINA TARNIŢĂ

Abstract. We describe theorems and computational methods for verifying the Birch and Swinnerton-Dyer conjectural formula for specific elliptic curves over $\mathbb{Q}$ of analytic ranks 0 and 1. We apply our techniques to show that if E is a non-CM elliptic curve over $\mathbb{Q}$ of conductor ≤ 1000 and rank 0 or 1, then the Birch and Swinnerton-Dyer conjectural formula for the leading coefficient of the L-series is true for E, up to odd primes that divide either Tamagawa numbers of E or the degree of some rational cyclic isogeny with domain E. Since the rank part of the Birch and Swinnerton-Dyer conjecture is a theorem for curves of analytic rank 0 or 1, this completely verifies the full conjecture for these curves up to the primes excluded above.

1. Introduction

Let E be an elliptic curve over $\mathbb{Q}$. The L-function $L(E, s)$ of E is a holomorphic function on $\mathbb{C}$ that encodes deep arithmetic information about E. This paper is about a conjecture of Birch and Swinnerton-Dyer that describes a deep connection between the behavior of $L(E, s)$ at $s = 1$ and the arithmetic of E.

We use theorems and computation to attack the following conjecture for many specific elliptic curves of conductor ≤ 1000 (see Section 1.1 below for more about the notation used in the conjecture):

Conjecture 1.1 (Birch and Swinnerton-Dyer)**.** *Let E be an elliptic curve defined over $\mathbb{Q}$. Then*

1. *The order of vanishing $\mathrm{ord}_{s=1} L(E, s)$ equals the rank r of E.*
2. *The group $\text{Ш}(E)$ is finite, and*

$$\frac{L^{(r)}(E,1)}{r!} = \frac{\Omega_E \cdot \mathrm{Reg}_E \cdot \prod_p c_p \cdot \#\text{Ш}(E)}{(\#E(\mathbb{Q})_{\mathrm{tor}})^2},$$

where $\text{Ш}(E)$ is the Shafarevich-Tate group, Ω_E is the real period, the c_p are the Tamagawa numbers at the primes p of bad reduction, and $E(\mathbb{Q})_{\mathrm{tor}}$ is the torsion subgroup of $E(\mathbb{Q})$.

Received by the editor June 30, 2005 and, in revised form, October 30, 2008.
2000 *Mathematics Subject Classification.* Primary 11Y99.
This material is based upon work supported by the National Science Foundation under Grant No. 0400386.

For more about Conjecture 1.1, see [Lan91, Wil00] and the papers they reference. Henceforth we call it the BSD conjecture, and call (1) the first part of the BSD conjecture and (2) the second part of the BSD conjectural formula.

Definition 1.2 (Analytic Ш)**.** If E has rank r, let

$$\#\text{Ш}(E)_{\mathrm{an}} = \frac{L^{(r)}(E,1)\cdot(\#E(\mathbb{Q})_{\mathrm{tor}})^2}{r!\cdot\Omega_E\cdot\mathrm{Reg}_E\cdot\prod_p c_p}$$

denote the order of $\text{Ш}(E)$ predicted by Conjecture 1.1. We call this the *analytic order* of $\text{Ш}(E)$.

Conjecture 1.3 (BSD(E,p))**.** *Let (E,p) denote a pair consisting of an elliptic curve E over $\mathbb{Q}$ and a prime p. Assume that E satisfies the first part of the BSD conjecture and moreover that $\#\text{Ш}(E)_{\mathrm{an}}$ is a rational number (see Remark* 1.7 *below). We call the assertion that $\text{Ш}(E)[p^\infty]$ is finite and that*

$$\mathrm{ord}_p(\#\text{Ш}(E)[p^\infty]) = \mathrm{ord}_p(\#\text{Ш}(E)_{\mathrm{an}})$$

the BSD conjecture at *p, and denote it by* BSD(E,p)*.*

We emphasize that whenever we write BSD(E,p), we are assuming that the first part of the BSD conjecture is true for E.

Cassels proved that the truth of the BSD conjecture is invariant under isogeny.

Theorem 1.4 (Cassels)**.** *If E and F are $\mathbb{Q}$-isogenous and p is a prime, then* BSD(E,p) *is true if and only if* BSD(F,p) *is true.*

Proof. See [Cas65, Mil86, Jor05]. □

One way to give evidence for the conjecture is to compute $\#\text{Ш}(E)_{\mathrm{an}}$ and note that it is a perfect square, in accord with the following theorem:

Theorem 1.5 (Cassels)**.** *If E is an elliptic curve over $\mathbb{Q}$ and p is a prime such that $\text{Ш}(E)[p^\infty]$ is finite, then $\#\text{Ш}(E)[p^\infty]$ is a perfect square.*

Proof. See [Cas62, PS99]. □

We use the notation of [Crea] to refer to specific elliptic curves over $\mathbb{Q}$. We call the first elliptic curve in each isogeny class *optimal* (see Definition 1.12 below for more about why).

Conjecture 1.6 (Birch and Swinnerton-Dyer ≤ 1000)**.** *For* all *optimal curves of conductor* ≤ 1000 *and rank* ≤ 1*, we have* $\#\text{Ш}(E) = 1$*, except for the following four rank* 0 *elliptic curves, where* $\text{Ш}(E)$ *has the indicated order:*

Curve	571a	681b	960d	960n
$\#\text{Ш}(E)$	4	9	4	4

Remark 1.7. A subtle point is that there is currently no elliptic curve of rank 2 (or larger) for which it is known that $\#\text{Ш}(E)_{\mathrm{an}}$ is a *rational number.* To prove rationality of $\#\text{Ш}(E)_{\mathrm{an}}$ in any example would require proving a deep relationship between the real numbers $L^{(2)}(E,1)$, Ω_E and Reg_E, perhaps analogous to the Gross-Zagier theorem in the case of analytic rank 1. That said, Cremona has computed $\#\text{Ш}(E)_{\mathrm{an}}$ to several digits precision for the curves of rank ≥ 2 with conductor ≤ 1000, and in each case $\#\text{Ш}(E)_{\mathrm{an}}$ is $1.0000\ldots$.

Theorem 1.8. *Conjecture* 1.1 *is true for all elliptic curves of conductor* ≤ 1000 *and rank* ≤ 1 *if and only if Conjecture* 1.6 *is true.*

Proof. By Theorem 1.4 it suffices to consider only the optimal curves, and the four listed in Conjecture 1.6 are the only ones with nontrivial $\#Ш(E)_{\mathrm{an}}$. This assertion about $\#Ш(E)_{\mathrm{an}}$ is verified by computation, as explained below.

In the book [Cre97], Cremona used modular symbols to compute $\#Ш(E)_{\mathrm{an}}$ as an exact integer for every curve of rank 0 and conductor ≤ 1000. For the curves of rank 1, Cremona computed a numerical approximation to $\#Ш(E)_{\mathrm{an}}$ to at least 10 digits of precision. Using the Gross-Zagier formula (see (3.2)) we obtain an explicit formula for $\#Ш(E)_{\mathrm{an}}$ that shows that $\#Ш(E)_{\mathrm{an}}$ is a rational number with bounded denominator. We explicitly computed such a bound for all curves of rank 1 and conductor ≤ 1000 and it was at most 5248800, which is much smaller than 10^{10}. Thus $\#Ш(E)_{\mathrm{an}} = 1$ for each elliptic curve of rank 1 and conductor ≤ 1000. □

In view of Theorem 1.8, the main goal of this paper is to obtain results in support of Conjecture 1.6. The results of Section 4.2 below together imply the theorem we claimed in the abstract:

Theorem 1.9. *Suppose that E is a non-CM elliptic curve of rank* ≤ 1, *conductor* ≤ 1000 *and that p is a prime. If p is odd, assume further that the mod p representation $\overline{\rho}_{E,p}$ is irreducible and p does not divide any Tamagawa number of E. Then* $\mathrm{BSD}(E,p)$ *is true.*

Proof. Combine Theorem 3.31, Theorem 3.25, and Theorem 4.4. □

For example, if E is the elliptic curve 37a, then according to [Cre97], all $\overline{\rho}_{E,p}$ are irreducible and the Tamagawa numbers of E are 1. Thus Theorem 1.9 asserts that the full BSD conjecture for E is true.

There are 18 optimal curves of conductor ≤ 1000 of rank 2 (and none of rank > 2). For these E of rank 2, nobody has proved that $Ш(E)$ is finite in even a single case. We exclude CM elliptic curves from most of our computations. The methods for dealing with both parts of the BSD conjecture for CM elliptic curves are different than for general curves; methods for the second part will be the subject of another paper. The same is true for $\mathrm{BSD}(E,p)$ when $\overline{\rho}_{E,p}$ is reducible, where we attack this problem theoretically in a forthcoming paper by Stein and Wuthrich, and later apply the theory in a further paper.

1.1. **Notation and background.**

If G is an abelian group, let G_{tor} denote the torsion subgroup and $G_{/\,\mathrm{tor}}$ denote the quotient G/G_{tor}. For an integer m, let $G[m]$ be the kernel of multiplication by m on G. For a commutative ring R, we let R^* denote the group of units in R.

1.1.1. *Galois cohomology of elliptic curves.* For a number field K, let $G_K = \mathrm{Gal}(\overline{\mathbb{Q}}/K)$. Let E be an elliptic curve defined over a number field K, and consider the first Galois cohomology group $\mathrm{H}^1(K,E) = \mathrm{H}^1(G_K, E(\overline{K}))$, and the local Galois cohomology groups $\mathrm{H}^1(K_v,E) = \mathrm{H}^1(\mathrm{Gal}(\overline{K}_v/K_v), E(\overline{K}_v))$, for each place v of K.

Definition 1.10 (Shafarevich-Tate group)**.** The *Shafarevich-Tate group*

$$Ш(E/K) = \mathrm{Ker}\Big(\mathrm{H}^1(K,E) \to \bigoplus_v \mathrm{H}^1(K_v,E)\Big)$$

of E measures the failure of global cohomology classes to be determined by their localizations at all places.

If E is an elliptic curve over a field F and if the field F is clear from the context, we write $Ш(E) = Ш(E/F)$. For example, if E is an elliptic curve over $\mathbb{Q}$, then $Ш(E)$ means $Ш(E/\mathbb{Q})$.

Definition 1.11 (Selmer group)**.** For each positive integer m, the *m-Selmer group* is

$$\mathrm{Sel}^{(m)}(E/K) = \mathrm{Ker}\Big(\mathrm{H}^1(K, E[m]) \to \bigoplus_v \mathrm{H}^1(K_v, E)\Big).$$

The Selmer group relates the Mordell-Weil and Shafarevich-Tate groups of E via the exact sequence

$$0 \to E(K)/mE(K) \to \mathrm{Sel}^{(m)}(E/K) \to Ш(E/K)[m] \to 0,$$

where $Ш(E/K)[m]$ denotes the m-torsion subgroup of $Ш(E/K)$. Note that every element of $Ш(E/K)$ has finite order since every element of $\mathrm{H}^1(K, E)$ has finite order.

1.1.2. *Elliptic curves over* $\mathbb{Q}$. See [Sil92, pp. 360–361] for the definition of $L(E, s)$ and [Wil95, BCDT01] for why $L(E, s)$ is entire.

Let E be an elliptic curve over $\mathbb{Q}$. As mentioned above, we use the notation of [Crea] to refer to elliptic curves over $\mathbb{Q}$. Thus, e.g., 37b3 refers to the third elliptic curve in the second isogeny class of elliptic curves of conductor 37, i.e., the curve $y^2 + y = x^3 + x^2 - 3x + 1$. The ordering of isogeny classes and curves in isogeny classes is as specified in [Crea]. If the last number is omitted, it is assumed to be 1, so 37b refers to the first curve in the second isogeny class of curves of conductor 37.

Let Reg_E be the absolute value of the determinant of the canonical height pairing on $E(\mathbb{Q})_{/\,\mathrm{tor}}$. Let $c_p = [E(\mathbb{Q}_p) : E_0(\mathbb{Q}_p)]$ be the Tamagawa number of E at p, where $E_0(\mathbb{Q}_p)$ is the subgroup of points that reduce to a nonsingular point modulo p. Let $\Omega_E = \int_{E(\mathbb{R})} |\omega|$, where $\omega = dx/(2y + a_1x + a_3)$ is the invariant differential attached to a minimal Weierstrass model

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

for E.

For any prime p, let $\overline{\rho}_{E,p} : G_{\mathbb{Q}} \to \mathrm{Aut}(E[p])$ denote the mod p representation and $\rho_{E,p} : G_{\mathbb{Q}} \to \mathrm{Aut}(T_pE)$ the representation on the p-adic Tate module T_pE of E.

It follows from [BCDT01] that every elliptic curve E over $\mathbb{Q}$ is modular, i.e., is a quotient of the modular curve $X_0(N)$, where N is the conductor of E.

Definition 1.12 (Optimal)**.** An elliptic curve E over $\mathbb{Q}$ is *optimal* if for every elliptic curve F and surjective morphism $X_0(N) \to F \to E$, we have $E \cong F$.

Optimal curves are also called "strong Weil curves" in the literature, and they are always the first curve listed in each isogeny class in the Cremona tables [Crea].

We say that E is a *complex multiplication* (CM) curve if $\mathrm{End}(E/\overline{\mathbb{Q}}) \neq \mathbb{Z}$.

2. Elliptic curve algorithms

2.1. **Images of Galois representations.** Let E be an elliptic curve over $\mathbb{Q}$. Many theorems that provide explicit bounds on $\#Ш(E)[p^\infty]$ have as a hypothesis that $\overline{\rho}_{E,p}$ or $\rho_{E,p}$ be either surjective or irreducible. In this section we explain how to prove that $\overline{\rho}_{E,p}$ or $\rho_{E,p}$ is surjective or irreducible in particular cases.

2.1.1. *Irreducibility.* Regarding irreducibility, note that $\overline{\rho}_{E,p}$ is irreducible if and only if there is no isogeny $E \to F$ over $\mathbb{Q}$ of degree p. The degrees of all such isogenies for curves of conductor ≤ 1000 are recorded in [Cre97], which were computed using Cremona's program `allisog`. This program uses results of Mazur [Maz78] along with computations involving modular curves of genus 0.

2.1.2. *Surjectivity.* We discuss surjectivity of the p-adic representation $\rho_{E,p}$ and the mod p representation $\overline{\rho}_{E,p}$ in the rest of this section.

Theorem 2.1 (Mazur)**.** *If E is semistable and $p \geq 11$, then $\overline{\rho}_{E,p}$ is surjective.*

Proof. See [Maz78, Thm. 4]. □

Example 2.2. Mazur's theorem implies that the representations $\overline{\rho}_{E,p}$ attached to the semistable elliptic curve $E = X_0(11)$ are surjective for $p \geq 11$. In fact $\overline{\rho}_{E,5}$ is reducible and all other $\overline{\rho}_{E,p}$ for $p \neq 5$ are surjective.

Theorem 2.3 (Cojocaru, Kani, and Serre)**.** *If E is a non-CM elliptic curve of conductor N, and*

$$p \geq 1 + \frac{4\sqrt{6}}{3} \cdot N \cdot \prod_{\textit{prime } \ell \mid N} \left(1 + \frac{1}{\ell}\right)^{1/2},$$

then $\overline{\rho}_{E,p}$ is surjective.

Proof. See Theorem 2 of [CK], whose proof relies on the results of [Ser72]. □

Example 2.4. When $N = 11$, the bound of Theorem 2.3 is ~ 38.52. When $N = 997$, the bound is ~ 3258.8. For $N = 40000$, the bound is ~ 143109.35.

Proposition 2.5. *Let E be a non-CM elliptic curve over $\mathbb{Q}$ of conductor N and let $p \geq 5$ be a prime. For each prime $\ell \nmid p \cdot N$ with $a_\ell \not\equiv 0 \pmod{p}$, let*

$$s(\ell) = \left(\frac{a_\ell^2 - 4\ell}{p}\right) \in \{0, -1, +1\},$$

where the symbol $\left(\frac{\cdot}{\cdot}\right)$ is the Legendre symbol. If -1 and $+1$ both occur as values of $s(\ell)$, then $\overline{\rho}_{E,p}$ is surjective. If $\mathrm{Im}(\overline{\rho}_{E,p})$ is contained in a Borel subgroup (i.e., reducible), then $s(\ell) \in \{0, 1\}$ for all ℓ, and if $\mathrm{Im}(\overline{\rho}_{E,p})$ is a nonsplit torus, then $s(\ell) \in \{0, -1\}$ for all ℓ.

Proof. This proposition follows from the proof of Proposition 19 [Ser72, §2.8], where we use the quadratic formula to convert the condition that certain polynomials modulo p be reducible or irreducible into the above quadratic residue symbol condition. The only thing one needs to add is that case iii) does not concern us because the image of the Galois representation cannot be one of the exceptional groups (this is mentioned in the introduction [Ser72, p. 261] and at the end of [Ser72, §4.2.b)]). □

For computational applications we apply Proposition 2.5 as follows. We choose a bound B and compute values $s(\ell)$; if both -1 and $+1$ occur as values of $s(\ell)$, we stop computing $s(\ell)$ and conclude that $\overline{\rho}_{E,p}$ is surjective. If for all $\ell \leq B$ we find that $s(\ell) \in \{0, 1\}$, we suspect that $\mathrm{Im}(\overline{\rho}_{E,p})$ is Borel, and attempt to show this, which is a finite computation (see Section 2.1.1). If for all $\ell \leq B$, we have $s(\ell) \in \{0, -1\}$, we suspect that $\mathrm{Im}(\overline{\rho}_{E,p})$ is contained in a nonsplit torus, and we try to show this by computing and analyzing the p-division polynomial of E. If this approach is inconclusive, we can always increase B and eventually the process terminates. We often apply some theorem under the hypothesis that $\overline{\rho}_{E,p}$ is surjective, which is something that in practice we verify for a particular p using Proposition 2.5.

Example 2.4 suggests that the bound of Theorem 2.3 is probably far larger than necessary. Nonetheless, it is small enough that in a reasonable amount of time we can determine whether $\overline{\rho}_{E,p}$ is surjective, using the above process, for all p up to the bound.

For $p \leq 3$ we can also determine surjectivity of the mod p representations by directly using the p-division polynomial of E.

Proposition 2.6. *Let $p = 2, 3$. Then $\overline{\rho}_{E,p}$ is surjective if and only if the p-division polynomial (of degree $p+1$) has Galois group the full symmetric group S_{p+1}.*

Proof. Fix a choice $y^2 = x^3 + ax + b$ of a (short) Weierstrass equation for E. We take division polynomials relative to this fixed choice of Weierstrass equation.

First suppose $p = 2$. Since $\# \mathrm{GL}_2(\mathbb{F}_2) = 6$, the representation $\overline{\rho}_{E,2}$ is surjective onto $\mathrm{GL}_2(\mathbb{F}_2)$ if and only if $\mathbb{Q}(E[2])$ has degree 6. By definition, the 2-division polynomial $f(x)$ is $4(x^3 + ax + b)$. Then $y^2 = f(x)/4$ is the Weierstrass equation that we fixed for E, so the y-coordinates of the 2-torsion points with respect to this equation are all 0. Thus $\mathbb{Q}(E[2])$ is the splitting field of the 2-division polynomial $f(x)$, which proves the proposition in the case $p = 2$.

Next suppose that $p = 3$. Then $f(x) = 3x^4 + 6ax^2 + 12bx - a^2$ is the 3-division polynomial of E. First assume that the Galois group of f is S_4, so S_4 is a quotient of the image of $\overline{\rho}_{E,3}$. Since S_4 has order 24 and $\mathrm{GL}_2(\mathbb{F}_3)$ has order 48, either the image of $\overline{\rho}_{E,3}$ is $\mathrm{GL}_2(\mathbb{F}_3)$ and we are done or it is isomorphic to S_4. But the image cannot be isomorphic to S_4, since the image would be normal (since it has index 2), and S_4 is not the unique normal subgroup of $\mathrm{GL}_2(\mathbb{F}_3)$ of index 2 (that normal subgroup is $\mathrm{SL}_2(\mathbb{F}_3)$, which is not isomorphic to S_4, since S_4 has a normal subgroup of index 2 and $\mathrm{SL}_2(\mathbb{F}_3)$ does not). This proves that $\overline{\rho}_{E,3}$ is surjective.

Next we assume that $\overline{\rho}_{E,3}$ is surjective, and we show that the 3-division polynomial f has Galois group S_4. That $\overline{\rho}_{E,3}$ is surjective implies that $\mathbb{Q}(E[3])$ has Galois group $\mathrm{GL}_2(\mathbb{F}_3)$ over $\mathbb{Q}$. Let $\mathbb{Q}(E[3])^+$ be the subfield fixed by the element -1 of $\mathrm{GL}_2(\mathbb{F}_3)$. Then $\mathbb{Q}(E[3])^+$ has Galois group $\mathrm{PGL}_2(\mathbb{F}_3)$, which is identified with S_4 by its action on the four 3-element subgroups of $E[3]$. Each such subgroup is in turn determined by the x-coordinate shared by its two nonzero points. So since $\overline{\rho}_{E,3}$ is surjective, any permutation of those x-coordinates is realized by some element of $\mathrm{Gal}(\mathbb{Q}(E[3])^+/\mathbb{Q})$. Thus the Galois group of the division polynomials (whose roots are those x-coordinates) maps surjectively to S_4, which means it equals S_4. □

Theorem 2.7 (Serre)**.** *If $p \geq 5$ is a prime of good reduction, then $\rho_{E,p}$ is surjective if and only if $\overline{\rho}_{E,p}$ is surjective.*

Proof. This follows from [Ser72, Thm. 4′, p. 300]. □

Remark 2.8. This result does not extend to $p = 3$ (see [Ser98, Ex. 3, p. IV-28]). In fact, there are infinitely many elliptic curves with $\overline{\rho}_{E,p}$ surjective, but $\rho_{E,p}$ not surjective (see forthcoming work of Noam Elkies).

2.2. **Special values of L-functions.** Let E be an elliptic curve over $\mathbb{Q}$ of conductor N, and let $f = \sum a_n q^n$ be the corresponding cusp form.

The following lemma will be useful in determining how many terms of the L-series of E are needed to compute the L-series to a given precision. (We could give a strong bound, but for our application this will be enough, and it is the simplest to apply in practice.)

Lemma 2.9. *For any positive integer n, we have $|a_n| \leq n$.*

Proof. For p prime we know that $a_p = \alpha + \beta$, where α and β are the roots of $x^2 - a_p x + p = 0$. Note that $|\alpha| = |\beta| = \sqrt{p}$.

Since a_n is multiplicative, it is enough to show $|a_n| \leq n$ for prime powers p^r. Let $r > 1$. If p divides N, the conductor of E, then $a_{p^r} = a_p a_{p^{r-1}}$ and by induction it follows that $|a_{p^r}| \leq p^{r/2} \leq p^r$. If $p \nmid N$, then $a_{p^r} = a_p a_{p^{r-1}} - p a_{p^{r-2}}$, and by induction

$$a_{p^r} = \alpha^r + \alpha^{r-1}\beta + \cdots + \alpha\beta^{r-1} + \beta^r.$$

Then

$$|a_{p^r}| \leq (r+1)p^{r/2}.$$

If $r+1 \leq p^{r/2}$, then the conclusion follows. This fails only for $p = 2$ and $r = 1, 2, 3, 4$ or $p = 3$ and $r = 1$. But if $r = 1$, then $|a_{p^r}| \leq p^{1/2} \leq p$, so we may assume $p = 2$ and $r = 2, 3, 4$. Therefore we finish the proof by observing that

$$\begin{aligned} |a_{2^2}| &= |a_2^2| \leq 2 \leq 2^2, \\ |a_{2^3}| &= |a_2^3 - 2a_2| \leq 2^{5/2} \leq 2^3, \\ |a_{2^4}| &= |a_2^4 - 4a_2^2| \leq 12 < 2^4. \end{aligned}$$ □

Suppose E has even analytic rank. By [Cre97, §2.13] or [Coh93, Prop. 7.5.8], we have

$$L(E,1) = 2\sum_{n=1}^{\infty} \frac{a_n}{n} e^{-2\pi n/\sqrt{N}}. \tag{2.1}$$

Using the bound $|a_n| \leq n$ of Lemma 2.9, we see that if we truncate the series (2.1) at the $(k-1)$th term, the error is at most

$$\varepsilon = 2\sum_{n=k}^{\infty} e^{-2\pi n/\sqrt{N}} = \frac{2e^{-2\pi k/\sqrt{N}}}{1 - e^{-2\pi/\sqrt{N}}},$$

and the quantity on the right can easily be evaluated.

Next suppose E has odd analytic rank. In [Cre97, §2.13] or [Coh93, Prop. 7.5.9] we find that

$$L'(E,1) = 2\sum_{n=1}^{\infty} \frac{a_n}{n} G_1(2\pi n/\sqrt{N}).$$

We have

$$G_1(x) = \int_1^{\infty} e^{-xy}\frac{dy}{y} = \int_x^{\infty} e^{-y}\frac{dy}{y} \leq e^{-x},$$

and we obtain the same error bound as for $L(E,1)$. (In fact, $G_1(x) \leq e^{-x}/x$ but we will not need this stronger bound.)

2.3. **Mordell-Weil groups.** If E is an elliptic curve over $\mathbb{Q}$ of analytic rank ≤ 1, there are algorithms to compute $E(\mathbb{Q})$ that are guaranteed to succeed. This is because $\#Ш(E)$ is finite, by [Kol91]. Independent implementations of these algorithms are available as part of `mwrank` [Creb] and Magma [BCP97]. We did most of our computations of $E(\mathbb{Q})$ using Sage [Sage] via `mwrank`, but used Magma in a few cases, since it is currently the only software in existence that implements 3-descents and 4-descents (thanks to Michael Stoll, Tom Womack, Mark Watkins, Geoff Bailey and others).

2.4. **Other algorithms.** We use many other elliptic curve algorithms, for example, for computing root numbers and the coefficients a_n of the modular form associated to E. For the most part, we used the PARI (see [ABC]) C-library via Sage (see [Sage]). For descriptions of these general elliptic curve algorithms, see [Coh93, Cre97].

3. The Kolyvagin bound

In this section we describe a bound due to Kolyvagin on $\#Ш(E) = \#Ш(E/\mathbb{Q})$, and we compute it for many specific elliptic curves over $\mathbb{Q}$. In fact, each bound in this section is stated as a bound on $\#Ш(E/K)$, where K is a quadratic imaginary field; this is not a problem, because the natural map $Ш(E/\mathbb{Q}) \to Ш(E/K)$ has kernel of order a power of 2, so the bound is also a bound on the odd part of $\#Ш(E/\mathbb{Q})$.

Proposition 3.1. *If p is an odd prime, then* $\mathrm{ord}_p(\#Ш(E/\mathbb{Q})) \leq \mathrm{ord}_p(\#Ш(E/K))$.

Proof. We see using the inf-res sequence of Galois cohomology that the inclusion $\mathrm{H}^1(\mathbb{Q}, E) \to \mathrm{H}^1(K, E)$ has kernel $\mathrm{H}^1(K/\mathbb{Q}, E)$. Since $K/\mathbb{Q}$ is a quadratic extension, the group $\mathrm{H}^1(K/\mathbb{Q}, E)$ is a 2-torsion group. The kernel G of the res map $Ш(E/\mathbb{Q}) \to Ш(E/K)$ is a subgroup of $\mathrm{H}^1(K/\mathbb{Q}, E)$, so G is also a 2-torsion group. □

Let E be an elliptic curve over $\mathbb{Q}$ of conductor N. For any quadratic imaginary field $K = \mathbb{Q}(\sqrt{-D})$, let E^D denote the twist of E by D. If E is defined by $y^2 = x^3 + ax + b$, then E^D is defined by $y^2 = x^3 + D^2ax + D^3b$, and

$$L(E/K, s) = L(E, s) \cdot L(E^D, s).$$

Definition 3.2 (Heegner hypothesis)**.** We say that K satisfies the *Heegner hypothesis* for E if $K \neq \mathbb{Q}(\sqrt{-1}), \mathbb{Q}(\sqrt{-3})$, the discriminant D of K is coprime to N, and every prime factor of N splits as a product of two distinct primes in the ring of integers of K.

Remark 3.3. Slight variants of many of the results below can be proved for $K = \mathbb{Q}(\sqrt{-1}), \mathbb{Q}(\sqrt{-3})$, but we exclude these two cases for simplicity.

If K satisfies the Heegner hypothesis for E, then there is a Heegner point $y_K \in E(K)$, which is the sum of images of certain complex multiplication (CM) points on $X_0(N)$ (see [GZ86, §I.3]). Properties of this point affect the arithmetic of E over K.

3.1. **Bounds on $\#Ш(E/K)$.** Let E be a non-CM elliptic curve over $\mathbb{Q}$. Suppose that K is an imaginary quadratic extension of $\mathbb{Q}$ that satisfies the Heegner hypothesis for E, and that the Heegner point y_K has infinite order. Under these hypotheses, Kolyvagin proved the following theorem (see [McC91] for a nice account of a generalization of this theorem):

Theorem 3.4 (Kolyvagin). *Suppose that p is an odd prime such that*

$$\overline{\rho}_{E,p} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}(E[p])$$

is surjective. Then

$$\mathrm{ord}_p(\#Ш(E/K)) \leq 2 \cdot \mathrm{ord}_p([E(K) : \mathbb{Z}y_K]). \tag{3.1}$$

Cha [Cha03, Cha05] extended Kolyvagin's method to provide better bounds on $Ш(E/K)$ in some cases. Let K be a number field, let D_K be the discriminant of K, and let N be the conductor of E, and again assume that the Heegner point y_K has infinite order.

Theorem 3.5 (Cha). *If $p \nmid D_K$, $p^2 \nmid N$, and $\overline{\rho}_{E,p}$ is irreducible, then*

$$\mathrm{ord}_p(\#Ш(E/K)) \leq 2 \cdot \mathrm{ord}_p([E(K) : \mathbb{Z}y_K]).$$

Remark 3.6. As we will see in the proof of Theorem 4.3 below, there is exactly one curve that satisfies the hypotheses of that theorem, but for which we cannot use Theorem 3.4 to prove $\mathrm{BSD}(E, 5)$ since for that curve $\overline{\rho}_{E,5}$ is not surjective. Fortunately, we can use Cha's Theorem 3.5 in this case.

Cha's assumption on the reduction of E at p and the assumption that $p \nmid D_K$ is problematic when there is a prime $p \geq 5$ of additive reduction or when one uses only one K. This situation does occur in several cases, which motivated us to prove the following theorem:

Theorem 3.7. *Suppose E is a non-CM elliptic curve over $\mathbb{Q}$. Suppose K is a quadratic imaginary field that satisfies the Heegner hypothesis, that the Heegner point y_K has infinite order, and suppose p is an odd prime such that $p \nmid [E(K)_{/tor} : \mathbb{Z}y_K]$, that $p \nmid \#E'(K)_{\mathrm{tor}}$ for all curves E' that are $\mathbb{Q}$-isogenous to E, and that $\mathrm{disc}(K)$ is divisible by some odd prime other than p. Then*

$$p \nmid \#Ш(E).$$

Since the proof of Theorem 3.7 is long, we defer it until Section 5.

Remark 3.8. If in Theorem 3.7, $\overline{\rho}_{E,p}$ is irreducible, then $p \nmid \#E'(K)_{\mathrm{tor}}$ for all E' isogenous to E. This is because the isogeny $E \to E'$ has degree coprime to p, so $E[p] \cong E'[p]$. Also, since $E[p]$ is irreducible, if $E'(K)$ were to contain a p-torsion point, it would have to contain all of them, a contradiction since $\boldsymbol{\mu}_p \not\subset K$ (recall that we exclude $\mathbb{Q}(\sqrt{-3})$ as a possibility for K).

Theorem 3.9 (Bump-Friedberg-Hoffstein, Murty-Murty, Waldspurger). *There are infinitely many quadratic imaginary extensions $K/\mathbb{Q}$ such that K satisfies the Heegner hypothesis and $\mathrm{ord}_{s=1} L(E/K) = 1$.*

Proof. If $\mathrm{ord}_{s=1}L(E, s) = 0$, then the papers [MM91] and [BFH90] both imply the existence of infinitely many K such that y_K has infinite order. If $\mathrm{ord}_{s=1}L(E, s) = 1$, then a result of Waldspurger ([Wal85]) applies, as does [BFH90]. □

Theorem 3.9 is not used in our computations, but ensures that our procedure for bounding $\#Ш(E)$, when E has analytic rank ≤ 1, is an algorithm; i.e., it always terminates with a nontrivial upper bound.

3.2. **The Gross-Zagier formula.** We use the Gross-Zagier formula to compute the index $[E(K) : \mathbb{Z}y_K]$ without explicitly computing y_K.

The modularity theorem of [BCDT01] asserts that there exists a surjective morphism $\pi : X_0(N) \to E$. Choose π to have minimal degree among all such morphisms. Let $\pi^*(\omega)$ be the pullback of a minimal invariant differential ω on E. Then $\pi^*(\omega) = \alpha \cdot f$, for some constant α and some normalized cusp form f. By [Edi91, Prop. 2], we know that $\alpha \in \mathbb{Z}$.

Definition 3.10 (Manin constant)**.** The *Manin constant* of E is $c = |\alpha|$.

Manin conjectured in [Man72, §5] that $c = 1$ for the optimal curve in the $\mathbb{Q}$-isogeny class of E. Cremona has shown that $c = 1$ for every curve over $\mathbb{Q}$ of conductor up to 60000 (see Section 3.5.1 below).

Theorem 3.11 (Gross-Zagier, Zhang)**.** *If K satisfies the Heegner hypothesis for E, then the Néron-Tate canonical height over K of y_K is*

$$h(y_K) = \frac{\sqrt{D}}{c^2 \cdot \int_{E(\mathbb{C})} \omega \wedge \overline{i\omega}} \cdot L'(E/K, 1). \tag{3.2}$$

Proof. Gross and Zagier proved the above formula in [GZ86] under the hypothesis that D is odd (see the bottom of p. 227 of [GZ86] for the restriction that D be odd). For the general assertion, see [Zha04, Thm. 6.1]. □

3.3. **Remarks on the index.** Suppose that E is an elliptic curve over $\mathbb{Q}$ of conductor N and that E has analytic rank 1 over a quadratic imaginary field K that satisfies the Heegner hypothesis. In [GZ86, Conj. 2.2, p. 311], Gross and Zagier rephrase the analogue of Conjecture 1.1 for E over K using the Gross-Zagier formula as follows:

Conjecture 3.12 (Birch and Swinnerton-Dyer)**.** *Suppose K is a quadratic imaginary field (not $\mathbb{Q}(i)$ or $\mathbb{Q}(\sqrt{-3})$) that satisfies the Heegner hypothesis, and that E has analytic rank 1 over K. Then*

$$\#Ш(E/K) = \left(\frac{[E(K) : \mathbb{Z}y_K]}{c \cdot \prod_{p|N} c_p} \right)^2 . \tag{3.3}$$

Here the c_p are the Tamagawa numbers of E over $\mathbb{Q}$, c is the Manin constant of E, and $\mathbb{Z}y_K$ is the cyclic group generated by y_K.

Remark 3.13. A serious issue is that Conjecture 3.12 implies that the index $I_K = [E(K) : \mathbb{Z}y_K]$ will be divisible by the Tamagawa numbers c_p. One sees using Tate curves that these Tamagawa numbers can be very large; hence the bound in Theorems 3.4, 3.5 and 3.7 can be very weak. In many cases when E has analytic rank 0, we could instead apply Theorem 4.1 below, but when E has analytic rank 1 a different approach is required such as computation of p-adic regulators and use of results of K. Kato, B. Perrin-Riou, P. Schneider and others toward p-adic analogues of the BSD conjecture. This is the subject of a forthcoming paper by W. Stein and C. Wuthrich. In general, it is natural to ask for a refinement of Kolyvagin's bound

(3.1) that takes into account the Tamagawa numbers. The forthcoming Berkeley Ph.D. thesis of Dimitar Jetchev makes a substantial improvement in this direction.

Remark 3.14. Conjecture 3.12 has interesting implications in certain special cases. For example, if E is the curve 91b, then $c_7 = c_{13} = 1$. Also $c = 1$, as Cremona has verified, and $\#E(\mathbb{Q})_{\mathrm{tor}} = 3$. Thus for any K, we have $3 \mid [E(K) : \mathbb{Z}y_K]$. If y_K has infinite order, then Conjecture 3.12 implies that $3^2 \mid \#Ш(E/K)$. For $K = \mathbb{Q}(\sqrt{-103})$, the point y_K is torsion, and in this case $E(K)$ has rank 3 and (conjecturally) $Ш(E/K)[3] = 0$. See Remark 3.27 for another example along these lines.

3.4. **Mordell-Weil groups and quadratic imaginary fields.** Let E be an elliptic curve over $\mathbb{Q}$ and $K = \mathbb{Q}(\sqrt{D})$ a quadratic imaginary field such that $E(K)$ has rank 1. In this section we explain how to understand $E(K)$ in terms of $E(\mathbb{Q})$ and $E^D(\mathbb{Q})$.

The following proposition is well known, but we were unable to find a reference.

Proposition 3.15. *Let $R = \mathbb{Z}[1/2]$ and $K = \mathbb{Q}(\sqrt{D})$. For any square-free integer $D \neq 1$, we have*

$$E(K) \otimes R = (E(\mathbb{Q}) \otimes R) \oplus (E^D(\mathbb{Q}) \otimes R).$$

Proof. Let τ be the complex conjugation automorphism on $E(K) \otimes R$. The characteristic polynomial of τ is $x^2 - 1$, which is square-free, so $E(K) \otimes R$ is a direct sum of its $+1$ and -1 eigenspaces for τ. The natural map $E(\mathbb{Q}) \hookrightarrow E(K)$ identifies $E(\mathbb{Q}) \otimes R$ with the $+1$ eigenspace for τ since $E(K)^{G_{\mathbb{Q}}} = E(\mathbb{Q})$; likewise, $E^D(\mathbb{Q}) \hookrightarrow E(K)$ identifies $E^D(\mathbb{Q}) \otimes R$ with the -1 eigenspace for τ. □

The following slightly more refined proposition will be important for certain explicit Heegner point computations (directly after Equation (3.4)).

Proposition 3.16. *Suppose $E(K)$ has rank 1. Then the image of either $E(\mathbb{Q})_{/\,\mathrm{tor}}$ or $E^D(\mathbb{Q})_{/\,\mathrm{tor}}$ has index at most 2 in $E(K)_{/\,\mathrm{tor}}$.*

Proof. Since $E(K)$ has rank 1, Proposition 3.15 implies that exactly one of $E(\mathbb{Q})$ and $E^D(\mathbb{Q})$ has rank 1 and the other has rank 0. We may assume that $E(\mathbb{Q})$ has rank 1 (otherwise, swap E and E^D). Let i be the natural inclusion $E(\mathbb{Q}) \hookrightarrow E(K)$, and let τ denote the automorphism of $E(K)$ induced by complex conjugation. Then $P \mapsto (1+\tau)P$ induces a map $E(K) \to E(K)^+ = E(\mathbb{Q})$ that, upon taking quotients by torsion, induces a map $\psi : E(K)_{/\,\mathrm{tor}} \to E(\mathbb{Q})_{/\,\mathrm{tor}}$. Let P_1 be a generator for $E(\mathbb{Q})_{/\,\mathrm{tor}}$ and P_2 a generator for $E(K)_{/\,\mathrm{tor}}$, and write $i(P_1) = nP_2$, for some nonzero integer n. Then

$$[2]P_1 = \psi(i(P_1)) = \psi(nP_2) = [n]\psi(P_2) = [nm]P_1 \pmod{E(\mathbb{Q})_{\mathrm{tor}}},$$

for some nonzero integer m. Thus $2 = nm$, so $n \leq 2$. □

The root number $\varepsilon_E = \pm 1$ of E is the sign of the functional equation of $L(E, s)$. If $\varepsilon_E = +1$, then the analytic rank $\mathrm{ord}_{s=1} L(E, s)$ is even, and if $\varepsilon_E = -1$, then it is odd.

Proposition 3.17. *Let E be an elliptic curve, let $D = D_K$ be a discriminant that satisfies the Heegner hypothesis such that $\mathrm{ord}_{s=1} L(E/K, s) = 1$, and let $R = \mathbb{Z}[1/2]$. Then*

(1) *If $\varepsilon_E = +1$, then a generator of $E(K) \otimes R$ is the image of a generator of $E^D(\mathbb{Q}) \otimes R$ and $L'(E/K,1) = L(E,1) \cdot L'(E^D,1)$.*
(2) *If $\varepsilon_E = -1$, then a generator of $E(K) \otimes R$ is the image of a generator of $E(\mathbb{Q}) \otimes R$ and $L'(E/K,1) = L'(E,1) \cdot L(E^D,1)$.*

Proof. Since D satisfies the Heegner hypothesis, by computing the residue symbol $\left(\frac{N}{D}\right)$ and taking into account how the sign of the functional equation changes under twist, we see that

$$\mathrm{ord}_{s=1} L(E,s) \not\equiv \mathrm{ord}_{s=1} L(E^{(D)},s) \pmod 2.$$

The factorization

$$L(E/K,s) = L(E/\mathbb{Q},s) \cdot L(E^D/\mathbb{Q},s)$$

then implies the formulas for $L'(E/K,1)$.

Since K satisfies the Heegner hypothesis and $\mathrm{ord}_{s=1} L(E/K,s) = 1$, work of Kolyvagin and Gross-Zagier (see [Kol91, Kol88, GZ86]) implies that $E(K)$ has rank 1. This implies the statement about generators. □

We will use the above proposition to relate computation of $E(K) \otimes R$ to computation of Mordell-Weil groups of elliptic curves defined over $\mathbb{Q}$.

3.5. **Computing the index of the Heegner point.** We assume throughout this section that E is an elliptic curve over $\mathbb{Q}$ and that K is a quadratic imaginary field that satisfies the Heegner hypothesis for E such that the corresponding Heegner point y_K has infinite order.

A key input to the theo rems of Section 3.1 is computation of the index $[E(K) : \mathbb{Z}y_K]$. We have

$$[E(K)_{/\mathrm{tor}} : \mathbb{Z}y_K]^2 = h(y_K)/h(z), \tag{3.4}$$

where z is a generator of $E(K)_{/\,\mathrm{tor}}$.

In the Gross-Zagier formula we have $h = h_K$, the Néron-Tate canonical height on $E(K) = E^D(K)$ relative to K. Let $h_{\mathbb{Q}}$ denote the height on $E(\mathbb{Q})$ or $E^D(\mathbb{Q})$. Note that if $P \in E(\mathbb{Q})$ or $E^D(\mathbb{Q})$, then

$$h_{\mathbb{Q}}(P) = \frac{1}{[K:\mathbb{Q}]} \cdot h_K(P) = \frac{h_K(P)}{2}. \tag{3.5}$$

Using Proposition 3.16 and the algorithms for computing Mordell-Weil groups (see Section 2.3), we can compute z or $2z$ (where z is a generator of $E(K)/_{\mathrm{tor}}$), so we can compute $h(z)$. In practice, even for curves of conductor up to 1000, it can take a huge amount of time to compute z. This section is about practical methods to either compute the index or at least bound it.

It is not difficult to compute $h(y_K)$, without computing y_K itself, using the Gross-Zagier formula (Section 3.2). We compute $L'(E/K,1)$ by computing L-functions of elliptic curves defined over $\mathbb{Q}$ as explained in Proposition 3.17. It remains to compute

$$\alpha = \frac{\sqrt{|D|}}{c^2 \int_{E(\mathbb{C})} \omega \wedge \overline{i\omega}}, \tag{3.6}$$

where c is the Manin constant of E.

3.5.1. *The Manin constant.* Manin conjectured that the Manin constant c for any optimal elliptic curve factor E of $X_0(N)$ is 1, and there are bounds on the possibilities for c (see, e.g., [Edi91, ARS05]). There is an algorithm to verify in any particular case that $c = 1$, as explained in the proof of the following proposition in Cremona's appendix to [ARS05].

Proposition 3.18 (Cremona)**.** *If E is an optimal elliptic curve of conductor at most* 60000*, then the Manin constant of E is* 1*.*

3.5.2. *The integral.* We have the following well-known lemma regarding the integral in (3.6), for which we were unable to find a suitable reference:

Lemma 3.19. *We have $\int_{E(\mathbb{C})} \omega \wedge i\overline{\omega} = 2 \cdot \mathrm{Vol}(\mathbb{C}/\Lambda)$, where the volume $\mathrm{Vol}(\mathbb{C}/\Lambda)$ is the absolute value of the determinant of a matrix formed from a basis for the lattice in $\mathbb{C}$ obtained by integrating the Néron differential ω_E against all homology classes in $\mathrm{H}_1(E, \mathbb{Z})$.*

Proof. Fix the Weierstrass equation $y^2 = 4x^3 + g_4x + g_6$ for E, so $x = \wp(z)$ and $y = \wp'(z)$. First note that

$$\omega = \frac{dx}{y} = \frac{d\wp(z)}{\wp'(z)} = \frac{\wp'(z)dz}{\wp'(z)} = dz.$$

Thus

$$\begin{aligned}\int_{E(\mathbb{C})} \omega \wedge i\overline{\omega} &= i \int_{\mathbb{C}/\Lambda} dz \wedge \overline{dz} \\ &= i \int_{\mathbb{C}/\Lambda} (dx + idy) \wedge (dx - idy) \\ &= i(-2i) \int_{\mathbb{C}/\Lambda} dx \wedge dy = 2 \cdot \mathrm{Vol}(\mathbb{C}/\Lambda). \qquad \square\end{aligned}$$

Note that $\mathrm{Vol}(\mathbb{C}/\Lambda)$ can be computed to high precision using the Gauss arithmetic-geometric mean, as described in [Cre97, §3.7].

3.5.3. *Mordell-Weil groups and heights.* For the curves on which we run our computation, we use [Creb] (via [Sage]), which computes a basis for $E^D(\mathbb{Q})$.

Cremona describes the computation of heights of points on curves defined over $\mathbb{Q}$ in detail in [Cre97, §3.4]. There is an explicit bound on the error in the height computation, which shrinks exponentially in terms of the precision of approximating series, and can be made arbitrarily small. For the L-function computations, see Section 2.2.

3.5.4. *Indices of Heegner points on rank* 1 *curves.* Suppose E is an elliptic curve over $\mathbb{Q}$ of analytic rank 1, and suppose we wish to compute indexes

$$i_K = [E(K)_{/\,\mathrm{tor}} : \mathbb{Z}y_K]$$

for various K. Assume that $E(\mathbb{Q})$ is known, so we can compute $h(z)$ to high precision, where z generates $E(\mathbb{Q})_{/\mathrm{tor}}$. Then computing the indexes i_K is relatively easy. For each K, compute $h(y_K)$ as described above using the Gross-Zagier formula, so

$$h(y_K) = \alpha \cdot L'(E, 1) \cdot L(E^D, 1).$$

Then if z also generates $E(K)_{/\,\mathrm{tor}}$ we have

$$i_K = \sqrt{\frac{h_K(y_K)}{h_K(z)}} = \sqrt{\frac{h_K(y_K)}{2h_{\mathbb{Q}}(z)}}. \tag{3.7}$$

The other possibility is that z generates a subgroup of $E(K)_{/\,\mathrm{tor}}$ of index 2, in which case there is a w that generates $E(K)_{/tor}$ such that $2w = z$, so $h(z) = 4h(w)$; hence

$$i_K = \sqrt{\frac{h_K(y_K)}{h_K(w)}} = \sqrt{\frac{h_K(y_K)}{\frac{1}{4}h_K(z)}} = 2 \cdot \sqrt{\frac{h_K(y_K)}{2h_{\mathbb{Q}}(z)}}. \tag{3.8}$$

We emphasize that computation of the Heegner point itself is not necessary. For the results of this index computation for E of conductor ≤ 1000, see Section 3.6.2.

Algorithm 3.20. Given an elliptic curve E of analytic rank 1 and a Heegner quadratic imaginary field K of discriminant D, this algorithm computes the odd part of the index i_K of the Heegner point in $E(K)/_{\mathrm{tor}}$.

(1) *Since E has analytic rank 1, the BSD rank conjecture is known for E, so we can compute $E(K)$. We can hence compute the regulator* Reg_E *of E, correct to precision at least* 10^{-10} *(i.e., we find a very small interval that contains* Reg_E*).*

(2) *Compute $L'(E,1)$ to some bounded precision ε, using* $2\sqrt{N}+10$ *terms. The bound ε is determined as explained in Section* 2.2.

(3) *Compute $L(E^D,1)$ to some bounded precision ε' using* $2\sqrt{N}+10$ *terms.*

(4) *Compute $\alpha = \sqrt{|D|}/(2\,\mathrm{Vol}(\mathbb{C}/\Lambda))$ to precision at least* 10^{-10} *using PARI.*

(5) *Using interval arithmetic (in* [Sage]*) and the bound above we compute an interval in which the real number*

$$\alpha \cdot L'(E,1) \cdot L'(E^D,1)/(\mathrm{Reg}_E/2)$$

must lie. We thus obtain a (small) interval that contains i_K^2 or $i_K^2/4$ using (3.7)–(3.8)*. Multiply by* 4 *if there is a unique integer divided by* 4 *in this interval.*

(6) *If there is a unique integer in the resulting interval, then by Theorem* 3.11 *this integer must be the square index $[E(K) : \mathbb{Z}y_K]^2$; that we find a square provides a good double check on our calculation. If there are at least two integers in this interval, we increase the precision of the computation of α,* Reg_E*, $L'(E,1)$, and $L(E,1)$ and repeat the above steps.*

Example 3.21. Let E be the elliptic curve 540b given by the Weierstrass equation $y^2 = x^3 + 3x + 1$, which has rank 1 and conductor $540 = 2^2 \cdot 3^3 \cdot 5$. The first K that satisfies the Heegner hypothesis is $\mathbb{Q}(\sqrt{-71})$. The group $E(\mathbb{Q})$ is generated by $z = (0,1)$, and we have $h_{\mathbb{Q}}(z) \sim 0.656622630$. We have $|D| = 71$, $c = 1$, and $\mathrm{Vol}(\mathbb{C}/\Lambda) \sim 3.832955$, so

$$\alpha \sim \frac{\sqrt{71}}{2 \cdot 3.832955} \sim 1.09917.$$

Also, $L'(E,1) \sim 1.9340458$ and $L(E^D,1) = 5.559761726$; hence

$$h(y_K) \sim 1.09917 \cdot 1.9340458 \cdot 5.559761726 \sim 11.819.$$

Thus

$$i_K = \sqrt{\frac{11.819}{2 \cdot 0.656622630}} \sim \sqrt{8.99999} \sim 3.$$

Note that here we have computed the *integer* i_K approximately to several decimal places of precision, so are justified in rounding to 3.

3.5.5. *Indices of Heegner points on rank* 0 *curves.* Assume that the analytic rank of E is 0. In practice, computing the indexes of Heegner points in this case is substantially more difficult than in the rank 1 case. For a Heegner quadratic imaginary field $K = \mathbb{Q}(\sqrt{D})$, we have for z a generator of $E(K)_{/\,\mathrm{tor}}$ that

$$i_K^2 = [E(K)_{/\,\mathrm{tor}} : \mathbb{Z}y_K]^2 = \frac{h(y_K)}{h(z)} = \alpha \cdot \frac{L(E,1) \cdot L'(E^D,1)}{h(z)}.$$

Thus one method for finding i_K is to find a generator $z' \in E^D(\mathbb{Q})$ exactly using descent algorithms, which will terminate since we know that $Ш(E^D)$ is finite, by Kolyvagin's theorem. However, since E^D has potentially large conductor and rank 1, in practice the Mordell-Weil group will sometimes be generated by a point of large height, hence be extremely time consuming to find. One can use 2-descent, 3-descent, 4-descent, and Heegner points methods (i.e., explicitly compute the coordinates of the Heegner point as decimals and try to recognize them using continued fractions). In some cases these methods produce in a reasonable amount of time an element of $E^D(\mathbb{Q})$ of infinite order, and one can then saturate the point using [Creb] to find a generator $z \in E^D(\mathbb{Q})$. However, we will explain a trick below to get information about the index without actually computing it.

Example 3.22. Let E be the curve 11a, with Weierstrass equation $y^2 + y = x^3 - x^2 - 10x - 20$. The first field that satisfies the Heegner hypothesis is $K = \mathbb{Q}(\sqrt{-7})$. The conductor of the quadratic twist $F = E^{-7}$ is 539, and we find a generator $z \in F(\mathbb{Q})$ for the Mordell-Weil group of this twist. This point has height $h_{\mathbb{Q}}(z) \sim 0.1111361471$. We have $|D| = 7$ and $\mathrm{Vol}(\mathbb{C}/\Lambda) \sim 1.8515436234$, so

$$\alpha \sim \frac{\sqrt{7}}{2 \cdot 1.8515436234} \sim 0.71447177.$$

Also, $L(E,1) \sim 0.25384186$ and $L'(E^D,1) \sim 1.225566874$, so the height over K of the Heegner point is thus

$$h(y_K) \sim 0.71447177 \cdot 0.25384186 \cdot 1.225566874 \sim 0.2222722925.$$

Thus by (3.5),

$$i_K^2 = \frac{h(z)}{h(y_K)} = \frac{2h_{\mathbb{Q}}(z)}{h(y_K)} \sim 1.$$

As mentioned above, there is a trick to bounding the index i_K without computing *any* elements of $E(K)$. This is useful when the algorithms mentioned above for computing a generator of $E^D(\mathbb{Q})$ produce no information in a reasonable amount of time. First compute the height $h(y_K)$ using the Gross-Zagier formula. Next compute the Cremona-Prickett-Siksek [CPS06] bound B for E^D, which is a number such that if $P \in E^D(\mathbb{Q})$, then the naive logarithmic height of P differs from the canonical height of P by at most B. Using standard sieving methods implemented in [Creb], we compute all points on E of naive logarithmic height up to some number h_0. If we find any point of infinite order, we saturate, and hence compute $E^D(\mathbb{Q})$, then use the above methods. If we find no point of infinite order, we conclude that there is no point in $E^D(\mathbb{Q})$ of canonical height $\leq h_0 - B$. If $h_0 - B > 0$, we obtain an upper bound on i_K as follows. If z is a generator for $E^D(\mathbb{Q})$, then $h_{\mathbb{Q}}(z) > h_0 - B$, so using (3.5) we have

$$h_{\mathbb{Q}}(z) = \frac{1}{2} \cdot h_K(z) = \frac{h(y_K)}{2 \cdot i_K^2} > h_0 - B.$$

Solving for i_K gives

$$i_K < \sqrt{\frac{h(y_K)}{2(h_0 - B)}}, \tag{3.9}$$

so to bound i_K we consider many K (e.g., the first 30 ordered by the absolute value of the discriminant), and for each compute the quantity on the right side of (3.9) for a fixed choice of h_0. We then use a K that minimizes this quantity.

Remark 3.23. Another approach to finding some Heegner point is to search for small points on $E(K)$ over various fields K, until finding a K that satisfies the Heegner hypothesis and is such that $E(K)$ has rank 1. For example, if E is given by $y^2 = x^3 + ax + b$, and x_0 is a small integer, write $y_0^2 \cdot D = x_0^3 + ax_0 + b$, where y_0 and D are integers, and D is square-free. Then (x_0, y_0) is a point on the quadratic twist of E by D. We did not use this approach, since it was not necessary in order to prove Theorem 1.9. It would be needed to continue these computations to a much larger conductor.

Example 3.24. Let E be the elliptic curve 546e. Then $K = \mathbb{Q}(\sqrt{-311})$ satisfies the Heegner hypothesis, since the prime divisors of $546 = 2\cdot 3\cdot 7\cdot 13$ split completely in K. We compute the height of the Heegner point y_K. Let F be the quadratic twist of E by -311. We have

$$\alpha \sim \frac{\sqrt{311}}{2 \cdot 0.034096494268966216 8001} \sim 258.60711587.$$

Thus

$$\begin{aligned} h(y_K) &\sim \alpha \cdot L(E,1) \cdot L'(F,1) \\ &\sim 258.60711587 \cdot 2.2783578 \cdot 12.41550 \sim 7315.20688, \end{aligned}$$

where in each case we compute the L-series using enough terms to obtain a value correct to $\pm 10^{-5}$. Thus 7320 is a conservative upper bound on $h(y_K)$. The Cremona-Prickett-Siksek bound for F is $B = 13.0825747$. We search for points on F of naive logarithmic height ≤ 18, and we find no points. Thus (3.9) implies that

$$i_K < \sqrt{7320/(2 \cdot (18 - 13.0825747))} \sim 27.28171 < 28.$$

It follows that if $p \mid i_K$, then $p \leq 23$. Searching up to height 21 would (presumably) allow us to remove 23, but this might take much longer.

For the results of our computations for all E of conductor $\leq$ 1000, see Section 3.6.3.

3.6. **Results of computations.**

3.6.1. *Two descent.* In this section, we explain how descent computations imply that $\mathrm{BSD}(E, 2)$ is true for curves of conductor $N \leq 1000$.

Theorem 3.25. *If E is an elliptic curve with $N \leq 1000$, then $\mathrm{BSD}(E, 2)$ is true.*

Proof. According to Theorem 1.4, it suffices to prove the theorem for the set S of optimal elliptic curves with $N \leq 1000$. By doing an explicit 2-descent, Cremona computed $\mathrm{Sel}^{(2)}(E/\mathbb{Q})$ for every curve $E \in S$, as explained in [Cre97]. His calculations show that $Ш(E)[2]$ has order the predicted order of $Ш(E)[2^\infty]$ for all $E \in S$. Using Magma's `FourDescent` command, we compute $\mathrm{Sel}^{(4)}(E/\mathbb{Q})$ in the three cases in which $Ш(E)[2] \neq 0$, and we find that $Ш(E)[4] = Ш(E)[2]$. By Theorem 1.8, it follows that $\mathrm{BSD}(E, 2)$ is true for all $E \in S$. □

3.6.2. *Curves of rank* 1. First we consider curves of rank 1. Recall from Conjecture 1.6 that we expect Ш to be trivial for all optimal rank 1 curves of conductor at most 1000.

Proposition 3.26. *Suppose (E, p) is a pair with E an optimal elliptic curve of conductor up to* 1000 *of rank* 1*. Let I be the greatest common divisor of $[E(K)_{/\,\mathrm{tor}} : \mathbb{Z}y_K]$ for the first four quadratic imaginary fields $K = \mathbb{Q}(\sqrt{D})$ (ordered by absolute value of the discriminant) that satisfy the Heegner hypothesis. If $p \mid I$, then*

$$p \mid 2 \cdot \#E(\mathbb{Q})_{\mathrm{tor}} \cdot \prod_{q \mid N} c_{E,q},$$

except if (E, p) is $(540b, 3)$ or $(756b, 3)$.

Proof. For each rank 1 curve E of conductor up to 1000 we apply Algorithm 3.20 with the first four Heegner discriminants $D = D_0, D_1, D_2, D_3$ (smallest in absolute value) to compute the index i_K and observe that the conclusion of the proposition holds. □

Remark 3.27. For the curves 540b and 756b there is no 3-torsion, but there is a rational 3-isogeny. In each case we verified in addition that 3 divides the GCD of the indexes $[E(K)_{/\mathrm{tor}} : \mathbb{Z}y_K]$ for at least the first 16 fields K (ordered by the absolute value of the discriminant) that satisfy the Heegner hypothesis. Thus as in Remark 3.14, Conjecture 3.12 asserts that $9 \mid \#Ш(E/K)$ for the first sixteen K. This illustrates that not only Tamagawa numbers but also isogenies can make it impossible to apply Kolyvagin's theorem to give a tight upper bound on $\#Ш(E/\mathbb{Q})$, even if Kolyvagin's theorem did not require that $\overline{\rho}_{E,p}$ is surjective.

Proposition 3.28. *Suppose E is a non-CM optimal curve of conductor ≤ 1000 and that p is an odd prime such that $\overline{\rho}_{E,p}$ is irreducible but not surjective. If E has rank* 0*, then (E, p) is one of the following:* (245b,3), (338d,3), (352e,3), (608b,5), (675d,5), (675f,5), (704h,3), (722d,3), (726f,3), (800e,5), (800f,5), (864d,3), (864f,3), (864g,3), (864i,3)*. If E has rank* 1*, then (E, p) is one of the following:* (245a,3), (338e,3), (352f,3), (608e,5), (675b,5), (675i,5), (704l,3), (722b,3), (726a,3), (800b,5), (800i,5), (864a,3), (864b,3), (864j,3), (864l,3)*. There are no curves of rank ≥ 2 with the above property.*

Proof. Using Proposition 2.5 we make a list of pairs (E, p) such that $\overline{\rho}_{E,p}$ might not be surjective, and such that if (E, p) is not in this list, then $\overline{\rho}_{E,p}$ is surjective. Then using the program `allisog`, we compute for each curve E, a list of all degrees of isogenies emanating from E, and remove those pairs (E, p) for which p divides the degree of one of those isogenies. The curves listed above are the ones that remain. □

Remark 3.29. In Proposition 3.28, the nonsurjective irreducible (E, p) come in pairs, one of rank 0 and one of rank 1 having the same conductor. Each pair of curves is related by a quadratic twist. This pattern is common, but does not always occur. For example, (1184f,3) and (1184h,3) are both of rank 0 and have nonsurjective irreducible representations, and no curve of conductor 1184 and rank 1 has this property. Note that $1184 = 2^5 \cdot 37$ and 1184f and 1184h are quadratic twists of each other by -1.

Remark 3.30. Proposition 3.28 suggests that it is rare for $\overline{\rho}_{E,p}$ to be nonsurjective yet irreducible. When this does occur, frequently $p^2 \mid N$, though not always. Continuing the computation to conductor 10000 we find that $p^2 \mid N$ about half the time in which $\overline{\rho}_{E,p}$ is nonsurjective yet irreducible. This gives a sense of the extent to which Theorem 3.5 improves upon Theorem 3.4.

Theorem 3.31. *Suppose (E, p) is a pair consisting of a rank 1 non-CM elliptic curve E of conductor ≤ 1000 and a prime p such that $\overline{\rho}_{E,p}$ is irreducible and p does not divide any Tamagawa number of E. Then* $\mathrm{BSD}(E, p)$ *is true.*

Proof. By Theorem 3.25 we may assume that p is odd. The pairs that do not satisfy the Heegner point divisibility hypothesis in Proposition 3.26 are those in $S = \{(540b, 3), (756b, 3)\}$. However, both of these curves admit a rational 3-isogeny, so are excluded by the hypothesis of Theorem 3.31.

Let

$$\begin{aligned} T = \{&(245a, 3), (338e, 3), (352f, 3), (608e, 5), (675b, 5), (675i, 5),\\ &(704l, 3), (722b, 3), (726a, 3), (800b, 5), (800i, 5), (864a, 3),\\ &(864b, 3), (864j, 3), (864l, 3)\}. \end{aligned}$$

Then Proposition 3.28, Theorem 1.8, and Theorem 3.4 together imply $\mathrm{BSD}(E, p)$ for all pairs as in the hypothesis of Theorem 3.31, except the pairs in T. Note that for each $(E, p) \in T$, we have $p^2 \mid N$, so Theorem 3.5 does not apply either. We eliminate the pairs $(245a, 3), (338e, 3), (352f, 3), (608e, 5), (704l, 3), (864j, 3), (864l, 3)$ from consideration because in each case $p \mid \prod c_\ell$.

For each $(E, p) \in T$ (except those eliminated above) a computation shows that the representation $\overline{\rho}_{E,p}$ is irreducible and E does not have CM. Note that for the pairs $\{(245a, 3), (338e, 3), (352f, 3), (608e, 5), (704l, 3), (864j, 3), (864l, 3)\}$ we have $p \mid [E(K) : \mathbb{Z}y_K]$ for the first six Heegner K, but that is not a problem since we eliminated these pairs from consideration. For the remaining 8 pairs, in each case we find a K that satisfies the hypotheses of Theorem 3.7, so $p \nmid \#Ш(E/K)$. See Table 1.

By Proposition 3.1, since p is odd, this implies that $\mathrm{BSD}(E, p)$ is true. □

Table 1

E	p	first K satisfying Theorem 3.7 for (E,p)
675b	5	$\mathbb{Q}(\sqrt{-11})$
675i	5	$\mathbb{Q}(\sqrt{-11})$
722b	3	$\mathbb{Q}(\sqrt{-15})$
726a	3	$\mathbb{Q}(\sqrt{-95})$
800b	5	$\mathbb{Q}(\sqrt{-39})$
800i	5	$\mathbb{Q}(\sqrt{-31})$
864a	3	$\mathbb{Q}(\sqrt{-23})$
864b	3	$\mathbb{Q}(\sqrt{-23})$

3.6.3. *Curves of rank* 0.

Proposition 3.32. *Suppose* (E,p) *is a pair with* E *an optimal elliptic curve of conductor* ≤ 1000 *of rank* 0. *Let* I *be the greatest common divisor of* $[E(K)_{/\,\mathrm{tor}} : \mathbb{Z}y_K]$ *as* K *varies over quadratic imaginary fields that satisfy the Heegner hypothesis. If* $p \mid I$ *and* $\overline{\rho}_{E,p}$ *is irreducible, then*

$$p \mid 2 \cdot \#E(\mathbb{Q})_{\mathrm{tor}} \cdot \prod_{q \mid N} c_{E,q},$$

except possibly for the curves in Table 2.

Table 2

E	$p \mid I$?	D used
258e	3	-983
378g	3	-47
594f	3	-359
600g	3	-71
612d	3	-359
626b	3	-39
658a	3	-31
676e	5	-23
681b	3	-8
735b	3	-479
738b	3	-23
742f	3, 5	-199
777b	3	-215
780b	3,7	-191
819d	3,5	-404
850i	3	-151
858d	5, 7	-95
858k	7	-1031
900a	3	-71
906e	$p \leq 19$	-23
924a	5	-1679
978c	3	-431
980i	3	-671

In Table 2*, the first column gives an elliptic curve, the second column gives the primes* p *(with* $\overline{\rho}_{E,p}$ *irreducible) that might divide the* GCD *of indexes, and the third column gives the discriminant used to make this deduction.*

Proof. We use the methods described in Section 3.5.5, including computing with intervals as in Algorithm 3.20. In many cases we combine the explicit computation of a Heegner point for one prime with the bounding technique explained in Section 3.5.5, or only compute information using the bound.

For the curve 910e, we used four-descent via Magma to compute the point $(3257919871/16641, 133897822473008/2146689)$ on the -159 twist E^D, found using

[Creb] that it generates $E^D(\mathbb{Q})$, and obtained an index that is a power of 2 and 3. Since 3 divides a Tamagawa number, we do not include 910e in our table. Likewise, for 930f and $D = -119$, we used Magma's four-descent commands to find a point of height ~ 85.3, and deduced that the only odd prime that divides the index is 11; since 11 is a Tamagawa number, we do not include 930f. Similar remarks apply for 966j with $D = -143$. We were unable to use 4-descent to find a generator for a twist of 906e1. (Fortunately, $906 = 2 \cdot 3 \cdot 151$, so Theorem 4.3 implies $\mathrm{BSD}(E, p)$ except for $p = 2, 3, 151$, and for our purposes we will only need that 151 does not divide the Heegner point index.) □

Remark 3.33. We may be able to further reduce the number of entries in Table 2 in Proposition 3.32 by using Magma's four-descent command. However, we will not need this for our ultimate application to the BSD conjecture (Theorem 4.4).

Theorem 3.34. *Suppose (E, p) is a pair with E a rank 0 non-CM curve of conductor ≤ 1000 and p a prime such that $\overline{\rho}_{E,p}$ is irreducible and p does not divide any Tamagawa number of E. Then $\mathrm{BSD}(E, p)$ is true except possibly if (E, p) appears in Table 2 in the statement of Proposition 3.32, i.e., if E appears in column 1 and p appears in the column directly to the right of E.*

Proof. The argument is similar to the proof of Theorem 3.31. By Theorem 3.25 we may assume that p is odd. Let S be the set of pairs (E, p) in the table in Proposition 3.32 Let

$$\begin{aligned} T = \{&(245b, 3), (338d, 3), (352e, 3), (608b, 5), (675d, 5), (675f, 5), \\ &(704h, 3), (722d, 3), (726f, 3), (800e, 5), (800f, 5), (864d, 3), \\ &(864f, 3), (864g, 3), (864i, 3)\}. \end{aligned}$$

Then Proposition 3.28, Theorem 1.8, and Theorem 3.4 together imply $\mathrm{BSD}(E, p)$ for all pairs as in the hypothesis of Theorem 3.34, except the pairs in $S \cup T$, since the representation $\overline{\rho}_{E,p}$ is surjective and we have verified that $p \nmid [E(K) : \mathbb{Z}y_K]$ for some K. We eliminate the pairs $(722d, 3)$ and $(726f, 3)$ from consideration because in each case $p \mid \prod c_\ell$.

For each $(E, p) \in T$ the representation $\overline{\rho}_{E,p}$ is irreducible and E does not have CM. Next, for each pair $(E, p) \in T$ except for $(722d, 3)$ and $(726f, 3)$, which we already eliminated, we find a K that satisfies all the hypotheses of Theorem 3.7, so for each pair (E, p), we have that $p \nmid \#\text{Ш}(E)$; hence $\mathrm{BSD}(E, p)$ is true. □

3.6.4. *Three-descent.* We sharpen Theorem 3.34 using Stoll's 3-descent package (see [Sto05]).

Proposition 3.35. *We have $3 \nmid \#\text{Ш}(E)$ for each of the curves listed in Table 2 in Proposition 3.32 with 3 in the second column and $\overline{\rho}_{E,3}$ irreducible, except for 681b where $\#\text{Ш}(E)[3^\infty] = 9$.*

Proof. We use Stoll's package [Sto05] to compute each of the Selmer groups

$$\mathrm{Sel}^{(3)}(E) \cong \text{Ш}(E)[3]$$

and obtain the claimed dimensions. For E the curve 681b, that $\#\text{Ш}(E)[3^\infty] = 9$ follows by applying Theorem 3.4 with $K = \mathbb{Q}(\sqrt{-8})$, and noting that $\overline{\rho}_{E,3}$ is surjective and the index of the Heegner point y_K in $E(K)_{/\mathrm{tor}}$ is exactly divisible by 3. □

Remark 3.36.

(1) In the case of 681b, one can alternatively use [CM00] and [AS05, App.] instead of an explicit 3-descent to see that $9 \mid \#\text{Ш}(E)$.
(2) When computing class groups in Stoll's package one must take care to not assume any unproven conjectures that speed up class group computations (by modifying the call to `ClassGroup` in `3descent.m`).

4. The Kato bound

The following theorem bounds $\text{Ш}(E)$ from above when $L(E,1) \neq 0$.

Theorem 4.1 (Kato). *Let E be an optimal elliptic curve over $\mathbb{Q}$ of conductor N, and let p be a prime such that $p \nmid 6N$ and $\rho_{E,p}$ is surjective. If $L(E,1) \neq 0$, then $\text{Ш}(E)$ is finite and*

$$\operatorname{ord}_p(\#\text{Ш}(E)) \leq \operatorname{ord}_p\left(\frac{L(E,1)}{\Omega_E}\right).$$

This theorem follows from the existence of an "optimal" Kato Euler system (see [Kat04], [Rub98, Cor. 8.9] and [MR04]). The precise statement in Theorem 4.1 follows from the discussion in Sections 8.1 and 8.3 of [SW08]. See also [Gri05] for further discussion and recent results on lower bounds on $\#\text{Ш}(E)$ that make use of optimal Kato Euler systems. In addition, [SW08] gives a bound that also applies in the case when $\overline{\rho}_{E,p}$ is reducible.

4.1. **Computations.** When $L(E,1) \neq 0$ the group $\text{Ш}(E)$ is finite. Therefore, $\operatorname{ord}_p(\#\text{Ш}(E))$ is even. Thus if $\operatorname{ord}_p\left(\frac{L(E,1)}{\Omega_E}\right)$ is odd, we conclude that

$$\operatorname{ord}_p(\#\text{Ш}(E)) \leq \operatorname{ord}_p\left(\frac{L(E,1)}{\Omega_E}\right) - 1.$$

Lemma 4.2. *There are no pairs (E,p) that satisfy the conditions of Theorem 4.1 with $N \leq 1000$, such that*

$$\operatorname{ord}_p(\#\text{Ш}(E)_{\text{an}}) < \operatorname{ord}_p\left(\frac{L(E,1)}{\Omega_E}\right) - 1.$$

Proof. First we use the table `allbsd` from [Crea] to make a table of ratios $L(E,1)/\Omega_E$ for all curves of conductor ≤ 1000. For each of these with $L(E,1) \neq 0$, we factor the numerator of the rational number $L(E,1)/\Omega_E$. We then observe that the displayed inequality in the statement of the proposition does not occur. □

Theorem 4.3. *Suppose (E,p) is a pair such that $N \leq 1000$, $p \nmid 3N$, E is a non-CM elliptic curve of rank 0, and $\overline{\rho}_{E,p}$ is irreducible. Then $\mathrm{BSD}(E,p)$ is true.*

Proof. The statement for $p = 2$ follows from Theorem 3.25.

Let S be the set of pairs (E,p) as in the statement of Theorem 4.3 for which E is optimal and $p > 2$. By Theorem 1.8 it suffices to prove that $p \nmid \#\text{Ш}(E)$ for all $(E,p) \in S$. Using Proposition 2.5 with $A = 1000$, we compute for each rank 0 non-CM elliptic curve of conductor $N \leq 1000$, all primes $p \nmid 6N$ such that $\rho_{E,p}$ might not be surjective. This occurs for 53 pairs (E,p), with the E's all distinct. For these 53 pairs (E,p), we find that the representation $\overline{\rho}_{E,p}$ is reducible (since there is an explicit p-isogeny listed in [Cre97]), except for the pair $(608b, 5)$, for which $\overline{\rho}_{E,5}$ is irreducible.

Thus Theorem 4.1 implies that for each pair $(E, p) \in S$, except $(608b, 5)$, we have the bound

$$\mathrm{ord}_p(\#\text{Ш}(E)) \leq \mathrm{ord}_p(L(E,1)/\Omega_E).$$

By Theorem 1.5, $\mathrm{ord}_p(\#\text{Ш}(E))$ is even, so $\text{Ш}(E)[p^\infty]$ is trivial whenever

$$\mathrm{ord}_p(L(E,1)/\Omega_E) \leq 1.$$

By Theorem 1.8, $\mathrm{ord}_p(\#\text{Ш}(E)_{\mathrm{an}}) = 0$ for all $p \geq 5$. Thus by Lemma 4.2, there are no pairs $(E, p) \in S$ with $\mathrm{ord}_p(L(E,1)/\Omega_E) > 1$ (since otherwise some $\mathrm{ord}_p(\#\text{Ш}(E)_{\mathrm{an}})$ would be nontrivial).

Finally, let E be the curve 608b and $p = 5$. Since E admits no 5-isogeny (see [Cre97]), $\overline{\rho}_{E,5}$ is irreducible. Also, $5^2 \nmid 608$, so for any Heegner K of discriminant coprime to 5 we can apply Theorem 3.5. Taking $K = \mathbb{Q}(\sqrt{-79})$, we find that the odd part of $[E(K) : \mathbb{Z}y_K]$ is 1, so $5 \nmid \#\text{Ш}(E/K)$. It follows that $5 \nmid \#\text{Ш}(E)$, so $\mathrm{BSD}(E, 5)$ is true, according to Theorem 1.8. This completes the proof. □

4.2. **Combining Kato and Kolyvagin.** In this section we bound $\text{Ш}(E)$ for rank 0 curves by combining the Kato and Kolyvagin approaches.

Theorem 4.4. *Suppose E is a non-CM elliptic curve of rank 0 with conductor $N \leq 1000$, that $\overline{\rho}_{E,p}$ is irreducible, and that p does not divide any Tamagawa number of E. Then $\mathrm{BSD}(E,p)$ is true.*

Proof. Let (E, p) be as in the hypotheses to Theorem 4.4. By Theorem 4.3, $\mathrm{BSD}(E, p)$ is true, except possibly if $p \mid 3N$. Theorem 3.34 implies $\mathrm{BSD}(E, p)$, except if (E, p) appear in Table 2 of Proposition 3.32. Inspecting the table, we see that whenever a prime $p \geq 5$ is in the second column, then p does not divide the conductor N of E. This proves $\mathrm{BSD}(E, p)$ for $p \geq 5$.

Let E be the curve 681b. Then $\mathrm{BSD}(E, 3)$ asserts that $\#\text{Ш}(E)[3^\infty] = 9$, which follows from Proposition 3.35.

Finally Proposition 3.35 implies $\mathrm{BSD}(E, 3)$ for the remaining curves, which proves the theorem. □

5. Proof of Theorem 3.7

In this section we prove Theorem 3.7. Assume that E and K are as in the statement of the theorem, and assume that $\mathrm{ord}_{s=1} L(E/K, 1) = 1$. Then the Heegner point y_K has infinite order. Kolyvagin ([Kol90]) shows that in this case the rank of $E(K)$ is 1 and $\text{Ш}(E/K)$ is finite.

5.1. **Gross's account.** Gross's account of Kolyvagin's work in [Gro91] contains a proof of the following theorem:

Theorem 5.1. *Suppose that E is an elliptic curve over $\mathbb{Q}$ that does not have complex multiplication, that K is a quadratic imaginary field that satisfies the Heegner hypothesis such that y_K has infinite order, and that p is an odd prime such that $\overline{\rho}_{E,p} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}(E[p])$ is surjective and $p \nmid [E(K)_{/tor} : \mathbb{Z}y_K]$. Then*

$$p \nmid \#\text{Ш}(E/K).$$

Our Theorem 3.7 provides a better bound in that it relaxes the surjectivity hypothesis on $\overline{\rho}_{E,p}$. Gross uses surjectivity of $\overline{\rho}_{E,p}$ as a hypothesis only to prove the following two propositions. We will prove analogous propositions below, but under weaker hypotheses, which yields our claimed improvement to [Gro91].

Proposition 5.2 (Gross). *Assume that $\overline{\rho}_{E,p}$ is surjective. For any integer n, let K_n be the ring class field of K of conductor n. The restriction map*

$$\mathrm{Res} : \mathrm{H}^1(K, E[p]) \to \mathrm{H}^1(K_n, E[p])^{\mathrm{Gal}(K_n/K)} \tag{5.1}$$

is an isomorphism.

Proof. This proposition is implicit in [Gro91, p. 241]. By [Gro91, Lemma 4.3], the fact that $\overline{\rho}_{E,p}$ is surjective implies that $E(K_n)[p] = 0$. That $E(K_n)[p] = 0$ implies that $\mathrm{H}^0(K_n/K, E[p](K_n)) = \mathrm{H}^2(K_n/K, E[p](K_n)) = 0$, so the inflation-restriction-transgression sequence then implies that Res is an isomorphism. □

Let $L = K(E[p])$ and consider the pairing

$$\mathrm{H}^1(K, E[p]) \otimes \mathrm{Gal}(L/K) \to E[p] \tag{5.2}$$

given by

$$(c, \sigma) = \mathrm{res}_L(c)(\sigma).$$

Gross also uses surjectivity of $\overline{\rho}_{E,p}$ when proving that the pairing (5.2) is nondegenerate, as follows. Setting $L = K(E[p])$, we have that

$$\mathrm{H}^1(L/K, E(L)[p]) \to \mathrm{H}^1(K, E[p]) \to \mathrm{H}^1(L, E[p])^{\mathrm{Gal}(L/K)} \to \mathrm{H}^2(L/K, E(L)[p]).$$

To see that the pairing is nondegenerate, it suffices to know that the groups $\mathrm{H}^i(L/K, E[p])$ vanish for $i = 1, 2$. This is because we have

$$\mathrm{H}^1(L, E[p])^{\mathrm{Gal}(L/K)} = \mathrm{Hom}(G_L, E[p])^{\mathrm{Gal}(L/K)}$$

since $K(E[p]) \subset L$ and the pairing is $(c, \sigma) = \mathrm{res}_L(c)(\sigma)$. Thus nondegeneracy of the pairing then follows from Proposition 5.3 below. The hypotheses are satisfied since $\gcd(D, Np) = 1$, so the fields K and $\mathbb{Q}(E[p])$ are linearly disjoint; hence

$$\mathrm{Gal}(K(E[p])/K) \cong \mathrm{Gal}(\mathbb{Q}(E[p])/\mathbb{Q}) \cong \mathrm{GL}_2(\mathbb{F}_p).$$

Proposition 5.3 (Gross). *Let E be an elliptic curve over any number field K and let p be an odd prime. Assume that the mod p Galois representation $\overline{\rho}_{E,p}$: $\mathrm{Gal}(\overline{\mathbb{Q}}/K) \to \mathrm{GL}_2(\mathbb{F}_p)$ is surjective. Then $\mathrm{H}^i(K(E[p])/K, E[p]) = 0$ for all $i \geq 1$.*

Proof. This proposition is implicit in [Gro91, pp. 249–250] and is proved in [Ste02, p. 146], but for the convenience of the reader we give a proof here. Set $L = K(E[p])$. If $Z \subset G$ is the subgroup corresponding to the scalars in $\mathrm{GL}_2(\mathbb{F}_p)$, then the Hochschild-Serre spectral sequence implies that

$$\mathrm{H}^i(G/Z, \mathrm{H}^j(Z, E(L)[p])) \Longrightarrow \mathrm{H}^{i+j}(L/K, E(L)[p]).$$

Since $\#Z = p-1$, and $E(L)[p]$ is a p-group, and p is odd, we have $\mathrm{H}^j(Z, E(L)[p]) = 0$ for all $j \geq 1$. Also, since p is odd, and since nonidentity scalars have no nonzero fixed points, $\mathrm{H}^0(Z, E(L)[p]) = 0$. Thus for all i, j we have

$$\mathrm{H}^i(G/Z, \mathrm{H}^j(Z, E(L)[p])) = 0,$$

which implies that the groups $\mathrm{H}^{i+j}(L/K, E(L)[p])$ are all 0. □

Thus our goal is to prove analogues of Propositions 5.2–5.3 under hypotheses that are more amenable to computational verification.

5.2. **Analogue of Proposition 5.2.** In this section we verify that

$$\mathrm{H}^i(K_n/K, E(K_n)[p]) = 0$$

for all $i \geq 0$ under a simple condition on p-torsion over K. As in the proof of Proposition 5.2, the inflation-restriction-transgression sequence then implies that the res map (5.1) is an isomorphism.

Proposition 5.4. *Let E be an elliptic curve over $\mathbb{Q}$ and K be a quadratic imaginary extension of $\mathbb{Q}$. Assume that p is a prime with $p \nmid \#E(K)_{\mathrm{tor}}$ and if $p = 3$ assume that $K \neq \mathbb{Q}(\zeta_3)$. Then for every finite abelian extension L of K we have*

$$\mathrm{H}^i(L/K, E(L)[p]) = 0 \qquad \textit{for all } i \geq 1.$$

Proof. Write the abelian group $\mathrm{Gal}(L/K)$ as a direct sum $P \oplus P'$, where P is its Sylow p-subgroup, so $p \nmid \#P'$. First we show that the subgroup of $E(L)[p]$ invariant under P' is trivial. Let $G = \mathrm{Gal}(L/K)/H$, where H is the subgroup of $\mathrm{Gal}(L/K)$ that acts trivially on $E(L)[p]$. Thus $G \subset \mathrm{Aut}(E(L)[p])$.

Case 1. If $p \nmid \#G$, then $P \subseteq H$, so P' surjects onto G. There is no nonzero element of $E(L)[p]$ invariant under $\mathrm{Gal}(L/K)$ by our assumption that $p \nmid \#E(K)$, so the same holds for P'.

Case 2. If $p \mid \#G$, we cannot have $E(L)[p] = \mathbb{F}_p$, since $\mathbb{F}_p$ has automorphism group isomorphic to $\mathbb{F}_p^*$, of order $p-1$, but $G \subset \mathrm{Aut}(E(L)[p])$ and $\#G > p-1$. Thus, $E(L)[p]$ is the full p-torsion subgroup of E, and we identify G with a subgroup of $\mathrm{GL}_2(\mathbb{F}_p)$ acting on $E(L)[p] = \mathbb{F}_p^2$.

We can choose a basis of $\mathbb{F}_p^2$ so that G contains the subgroup generated by $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. Since G is abelian, it must be contained in the normalizer of this subgroup, so $G \subseteq \{\left(\begin{smallmatrix} a & b \\ 0 & a \end{smallmatrix}\right) : a \in \mathbb{F}_p^*, b \in \mathbb{F}_p\}$. We claim that G contains an element with $a \neq 1$. Since $E[p] = E(L)[p]$, the representation $\mathrm{Gal}(\overline{\mathbb{Q}}/K) \to \mathrm{Aut}(E[p])$ factors through $\mathrm{Gal}(L/K)$. The determinant of $\overline{\rho}_{E,p} : G_{\mathbb{Q}} \to \mathrm{Aut}(E[p])$ is surjective onto $\mathbb{F}_p^*$, and $[K : \mathbb{Q}] = 2$, so the character $\mathrm{Gal}(\overline{K}/K) \to \mathbb{F}_p^*$ has image of index at most 2 in F_p^*. That is, it contains at least $(p-1)/2$ elements, the squares in $\mathbb{F}_p^*$. Thus, for $p > 3$, the group G contains an element with nontrivial determinant having the form $\left(\begin{smallmatrix} a & b \\ 0 & a \end{smallmatrix}\right)$ with $a \neq 1$. Now, $\left(\begin{smallmatrix} a & b \\ 0 & a \end{smallmatrix}\right)^p = \left(\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right)$ since $a, b \in \mathbb{F}_p$, so $\mathrm{Gal}(L/K)$ contains an element that acts as a nontrivial scalar. Since the group of scalars in $\mathrm{GL}_2(\mathbb{F}_p)$ has $p-1$ elements, this nontrivial scalar must be in P', so $E(L)[p]^{P'} = 0$.

We have shown in each case that $E(L)[p]^{P'} = 0$. Because $p \nmid \#P'$, we have $\mathrm{H}^i(P', E(L)[p]) = 0$ for all $i \geq 1$, so for each $i \geq 1$ there is an exact inflation-restriction sequence

$$0 \to \mathrm{H}^i(P, E(L)[p]^{P'}) \to \mathrm{H}^i(L/K, E(L)[p]) \to \mathrm{H}^i(P', E(L)[p]).$$

The first group vanishes since $E(L)[p]^{P'} = 0$, and the third group vanishes as mentioned above. We conclude that $\mathrm{H}^i(L/K, E(L)[p]) = 0$, as claimed.

Finally we deal with the case $p = 3$. The only situation in the above argument where $p = 3$ is relevant is in Case 2, when $3 \mid \#G$. Our hypothesis that $K \neq \mathbb{Q}(\zeta_3)$ implies that $\det(\rho_{E,3}) : \mathrm{Gal}(\overline{K}/K) \to \mathbb{F}_3^*$ is surjective, since the fixed field of the kernel of the mod 3 cyclotomic character is $\mathbb{Q}(\zeta_3)$. If we are in Case 2, then the image of $\mathrm{Gal}(\overline{K}/K)$ in $\mathrm{GL}_2(\mathbb{F}_3)$ is contained in $\{\left(\begin{smallmatrix} a & b \\ 0 & a \end{smallmatrix}\right) : a \in \mathbb{F}_p^*, b \in \mathbb{F}_p\}$. Since no such upper triangular matrix has determinant 2, this contradicts the surjectivity of $\det(\rho_{E,3})$. Thus our hypothesis that $K \neq \mathbb{Q}(\zeta_3)$ implies that Case 2 does not occur. □

Corollary 5.5. *Let E be an elliptic curve with $p \nmid \#E(K)_{\mathrm{tor}}$, where $p > 3$ or, if $p = 3$, $K \neq \mathbb{Q}(\zeta_3)$. Let K_n be the ring class field of conductor n of K. Then $\mathrm{H}^i(K_n/K, E(K_n)[p]) = 0$ for all $i \geq 1$.*

5.3. **Analogue of Proposition 5.3.**

Lemma 5.6. *Let p be an odd prime. The determinant of $\overline{\rho}_{E,p}$ is the cyclotomic character; hence $\det(\overline{\rho}_{E,p})$ is surjective.*

Proof. For the convenience of the reader, we give a proof of this very standard fact here. The Weil pairing induces an isomorphism of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-modules

$$E[p] \wedge E[p] \cong \mu_p.$$

Let $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Fix a basis $\{e_1, e_2\}$ of $E[p]$, with respect to which $\overline{\rho}_{E,p}(\sigma)$ has the form $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Then

$$\sigma(e_1 \wedge e_2) = (ae_1 + ce_2) \wedge (be_1 + de_2) = \det(\rho_p(\sigma)) \cdot e_1 \wedge e_2.$$

It follows that composition with the determinant gives the cyclotomic character (i.e., the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on μ_p), which is surjective since no nontrivial pth roots of unity lie in $\mathbb{Q}$. □

By hypothesis, the discriminant $\mathrm{disc}(K)$ is divisible by a prime other than p and is coprime to N. If K were contained in $\mathbb{Q}(E[p])$, all primes that ramify in K would ramify in $\mathbb{Q}(E[p])$, so they would divide Np, which contradicts our hypothesis. Thus the quadratic field K is linearly disjoint from $\mathbb{Q}(E[p])$, so the restriction map

$$\mathrm{Gal}(K(E[p])/K) \to \mathrm{Gal}(\mathbb{Q}(E[p])/\mathbb{Q})$$

is an isomorphism. The action of $\mathrm{Gal}(K(E[p])/K)$ on the module $E[p]$ is through $\mathrm{Gal}(\mathbb{Q}(E[p])/\mathbb{Q})$, so for our application it will suffice to show that for $i > 0$,

$$\mathrm{H}^i(\mathbb{Q}(E[p])/\mathbb{Q}, E[p]) = 0.$$

Let $G \subseteq \mathrm{Gal}(\mathbb{Q}(E[p])/\mathbb{Q})$ be the image of $\overline{\rho}_{E,p}$. If $p \nmid \#G$, then for $i > 0$ we have $\mathrm{H}^i(G, E[p]) = 0$ since $E[p]$ is a p-group. Therefore we may assume that $p \mid \#G$. By [Ser72, Prop. 15], the image G either contains $\mathrm{SL}_2(\mathbb{F}_p)$ or is contained in a Borel subgroup of $\mathrm{GL}_2(\mathbb{F}_p)$. First consider the case when G contains $\mathrm{SL}_2(\mathbb{F}_p)$. By Lemma 5.6, the determinant $\det : G \to \mathbb{F}_p^*$ is surjective, so in fact $G = \mathrm{GL}_2(\mathbb{F}_p)$. In this case, we already know Propositions 5.2–5.3. Thus we turn next to the case when G is contained in a Borel subgroup.

Lemma 5.7. *Assume that G is contained in a Borel subgroup of $\mathrm{GL}_2(\mathbb{F}_p)$. Moreover, assume that there is a basis of $E[p]$ so that G acts as $\left(\begin{smallmatrix} \chi & * \\ 0 & \psi \end{smallmatrix}\right)$ where χ and ψ are nontrivial characters. Then $\mathrm{H}^i(G, E[p]) = 0$ for all $i \geq 0$.*

Proof. Let $W = \left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right)$ be the unique p-Sylow subgroup of $\left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right) \subset \mathrm{GL}_2(\mathbb{F}_p)$. We may assume $W \subset G$, for otherwise G has order prime to p, and the cohomology vanishes.

We begin by explicitly computing $\mathrm{H}^j(W, E[p])$ using the fact that W is cyclic, generated by $w = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. Recall that for cyclic groups we can compute cohomology using the projective resolution

$$\cdots \to \mathbb{Z}[W] \to \mathbb{Z}[W] \to \mathbb{Z} \to 0,$$

where the boundary maps alternate between multiplication by $w-1$ and $\mathrm{Norm}(w) = \sum_{i=0}^{p-1} w^i$.

Then we see that

$$\mathrm{H}^j(W, E[p]) = \begin{cases} \mathrm{Ker}(1-w)/\,\mathrm{Im}(\mathrm{Norm}(w)) = \langle \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) \rangle\,, & \text{if } j \text{ is even,} \\ \mathrm{Ker}(\mathrm{Norm}(w))/\,\mathrm{Im}(1-w) = \mathbb{F}_p^2/\,\langle \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) \rangle\,, & \text{if } j \text{ is odd.} \end{cases}$$

Since χ and ψ are nontrivial by assumption, the G/W-invariants for both of these groups are trivial. Thus $\mathrm{H}^j(W, E[p])^{G/W} = 0$ for $j \geq 0$. Consider the Hochschild-Serre spectral sequence

$$\mathrm{H}^i(G/W, \mathrm{H}^j(W, E[p])) \Rightarrow \mathrm{H}^{i+j}(G, E[p]).$$

For $i > 0$, since $\#(G/W)$ is prime to p, and $\mathrm{H}^j(W, E[p])$ is a p-group for all j, the group $\mathrm{H}^i(G/W, \mathrm{H}^j(W, E[p]))$ is trivial. But when $i = 0$ we have just computed that $\mathrm{H}^i(G/W, \mathrm{H}^j(W, E[p])) = \mathrm{H}^j(W, E[p])^{G/W} = 0$, so the entire spectral sequence is trivial, and we conclude that $\mathrm{H}^n(G, E[p]) = 0$ for all $n \geq 0$. □

In this section we show how the vanishing of $\mathrm{H}^i(\mathbb{Q}(E[p])/\mathbb{Q}, E[p])$ follows from a statement about torsion and rational isogenies.

Note that E has no $\mathbb{Q}$-rational p-isogeny if and only if $\overline{\rho}_{E,p}$ is irreducible.

Proposition 5.8. *If p is an odd prime and E has no $\mathbb{Q}$-rational p-isogeny, then $\mathrm{H}^i(\mathbb{Q}(E[p])/\mathbb{Q}, E[p]) = 0$ for all $i > 0$.*

Proof. Our hypothesis that E has no $\mathbb{Q}$-rational p-isogeny implies that $\overline{\rho}_{E,p}$ is irreducible. As we already noted, the problem reduces to the case when either G is contained in a Borel subgroup or $G = \mathrm{GL}_2(\mathbb{F}_p)$. The latter case follows from Proposition 5.3. The former case contradicts the hypothesis since the module $E[p]$ is reducible as a module over a Borel subgroup. □

For the above result, we used the irreducibility of the representation to deal with the case when G was contained in a Borel subgroup. The following proposition completes the proof of the general case:

Proposition 5.9. *Suppose p is an odd prime and that $E(\mathbb{Q})[p] = 0$ and for all elliptic curves E' that are p-isogenous to E over $\mathbb{Q}$ we have $E'(\mathbb{Q})[p] = 0$. Then*

$$\mathrm{H}^i(\mathbb{Q}(E[p])/\mathbb{Q}, E[p]) = 0 \qquad \textit{for all } i > 0.$$

Proof. If E admits no p-isogeny, then Proposition 5.8 implies the required vanishing. Thus we may assume that E admits a rational p-isogeny, so $E[p]$ is reducible, and $G = \mathrm{Im}(\overline{\rho}_{E,p})$ is contained in a Borel subgroup. In particular, for some basis of $E[p]$, the image G acts as $\left(\begin{smallmatrix}\chi & *\\0 & \psi\end{smallmatrix}\right)$ for characters χ and ψ. If both χ and ψ are nontrivial, then Lemma 5.7 implies the proposition and we are done. Thus assume that either χ or ψ is trivial.

First suppose that χ is trivial. Then all matrices of the above form fix $\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$. Therefore there is a point of $E[p]$ fixed by the action of G, which contradicts the assumption that $E(\mathbb{Q})[p] = 0$.

Next suppose that ψ is trivial. Matrices of the above form preserve the line generated by $\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$, so this line forms a $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-stable subspace of $E[p]$. In particular, there exists an isogeny over $\mathbb{Q}$ to a curve E' having this line as kernel. The image under this isogeny of the line generated by $\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ is a 1-dimensional subspace of $E'[p]$, and since $\psi = 1$, $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts trivially on this subspace (we have an isomorphism of Galois modules $E/\langle\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)\rangle \cong E'$). Thus, $E'(\mathbb{Q})[p]$ is nontrivial, contradicting our assumption. □

Acknowledgement

We thank Michael Stoll for suggesting this project at an American Institute of Mathematics meeting and for initial feedback and ideas, and Stephen Donnelly and Michael Stoll for much key input to Section 5. We thank John Cremona for many discussions and his immensely useful computer software. We thank Don Zagier for Remark 1.7. Finally we thank Benedict Gross and Noam Elkies for helpful feedback and encouragement throughout the project, and the anonymous referee for a massive number of extremely helpful comments. Elkies also provided part of the proof of Proposition 2.6.

We thank the Harvard College Research Program for funding Pătraşcu's work on this paper, and the Herchel Smith Harvard Undergraduate Research Fellowship for supporting Patrikis. Jorza and Stein were supported by the National Science Foundation under Grant No. 0400386.

References

[ABC] B. Allombert, K. Belabas, H. Cohen, X. Roblot, and I. Zakharevitch, PARI/GP, http://pari.math.u-bordeaux.fr/.

[ARS05] A. Agashe, K. A. Ribet, and W. A. Stein, *The Manin constant, congruence primes, and the modular degree*, Preprint, http://www.williamstein.org/papers/manin-agashe/, With an appendix by J. Cremona (2005).

[AS05] A. Agashe and W. Stein, *Visible evidence for the Birch and Swinnerton-Dyer conjecture for modular abelian varieties of analytic rank zero*, Math. Comp. **74** (2005), no. 249, 455–484 (electronic), With an appendix by J. Cremona and B. Mazur. MR2085902

[BCDT01] C. Breuil, B. Conrad, F. Diamond, and R. Taylor, *On the modularity of elliptic curves over $\mathbb{Q}$: Wild* 3*-adic exercises*, J. Amer. Math. Soc. **14** (2001), no. 4, 843–939 (electronic). MR2002d:11058

[BCP97] W. Bosma, J. Cannon, and C. Playoust, *The Magma algebra system. I. The user language*, J. Symbolic Comput. **24** (1997), no. 3–4, 235–265, Computational algebra and number theory (London, 1993). MR1484478

[BFH90] Daniel Bump, Solomon Friedberg, and Jeffrey Hoffstein, *Nonvanishing theorems for L-functions of modular forms and their derivatives*, Invent. Math. **102** (1990), no. 3, 543–618. MR1074487 (92a:11058)

[Cas62] J. W. S. Cassels, *Arithmetic on curves of genus* 1*. III. The Tate-Šafarevič and Selmer groups*, Proc. London Math. Soc. (3) **12** (1962), 259–296. MR29:1212

[Cas65] J. W. S. Cassels, *Arithmetic on curves of genus* 1*. VIII. On conjectures of Birch and Swinnerton-Dyer*, J. Reine Angew. Math. **217** (1965), 180–199. MR31:3420

[Cha03] Byungchul Cha, *Vanishing of Some Cohomology Groups and Bounds for the Shafarevich-Tate Groups of Elliptic Curves*, Johns-Hopkins Ph.D. Thesis (2003).

[Cha05] ———, *Vanishing of some cohomology groups and bounds for the Shafarevich-Tate groups of elliptic curves*, J. Number Theory **111** (2005), 154–178. MR2124047 (2006g:11130)

[CK] Alina Carmen Cojocaru and Ernst Kani, *On the surjectivity of the Galois representations associated to non-CM elliptic curves*, Canad. Math. Bull. **48** (2005), 16–31. MR2118760 (2005k:11109)

[CM00] J. E. Cremona and B. Mazur, *Visualizing elements in the Shafarevich-Tate group*, Experiment. Math. **9** (2000), no. 1, 13–28. MR1758797

[Coh93] H. Cohen, *A course in computational algebraic number theory*, Springer-Verlag, Berlin, 1993. MR94i:11105

[CPS06] J. E. Cremona, M. Prickett, and Samir Siksek, *Height difference bounds for elliptic curves over number fields*, J. Number Theory **116** (2006), no. 1, 42–68. MR2197860 (2006k:11121)

[Crea] J. E. Cremona, *Elliptic curves of conductor* ≤ 25000, `http://www.maths.nott.ac.uk/personal/jec/ftp/data/`.

[Creb] ———, `mwrank` *(computer software)*, `http://www.maths.nott.ac.uk/personal/jec/mwrank/`

[Cre97] ———, *Algorithms for modular elliptic curves*, second ed., Cambridge University Press, Cambridge, 1997, `http://www.maths.nott.ac.uk/personal/jec/book/`. MR1628193 (99e:11068)

[Edi91] B. Edixhoven, *On the Manin constants of modular elliptic curves*, Arithmetic algebraic geometry (Texel, 1989), Birkhäuser Boston, Boston, MA, 1991, pp. 25–39. MR92a:11066

[Gri05] G. Grigorov, *Kato's Euler System and the Main Conjecture*, Harvard Ph.D. Thesis (2005).

[Gro91] B. H. Gross, *Kolyvagin's work on modular elliptic curves*, L-functions and arithmetic (Durham, 1989), Cambridge Univ. Press, Cambridge, 1991, pp. 235–256. MR1110395 (93c:11039)

[GZ86] B. Gross and D. Zagier, *Heegner points and derivatives of* L*-series*, Invent. Math. **84** (1986), no. 2, 225–320. MR87j:11057

[Jor05] A. Jorza, *The Birch and Swinnerton-Dyer Conjecture for Abelian Varieties over Number Fields*, Harvard University Senior Thesis (2005).

[Kat04] Kazuya Kato, p*-adic Hodge theory and values of zeta functions of modular forms*, Astérisque (2004), no. 295, ix, 117–290, Cohomologies p-adiques et applications arithmétiques. III. MR2104361

[Kol88] V. A. Kolyvagin, *Finiteness of* $E(\mathbf{Q})$ *and* $Ш(E, \mathbf{Q})$ *for a subclass of Weil curves*, Izv. Akad. Nauk SSSR Ser. Mat. **52** (1988), no. 3, 522–540, 670–671. MR89m:11056

[Kol90] ———, *Euler systems*, The Grothendieck Festschrift, Vol. II, Birkhäuser Boston, Boston, MA, 1990, pp. 435–483. MR92g:11109

[Kol91] V. A. Kolyvagin, *On the Mordell-Weil group and the Shafarevich-Tate group of modular elliptic curves*, Proceedings of the International Congress of Mathematicians, Vols. I, II (Kyoto, 1990) (Tokyo), Math. Soc. Japan, 1991, pp. 429–436. MR1159231 (93c:11046)

[Lan91] S. Lang, *Number theory. III.* Diophantine geometry, Springer-Verlag, Berlin, 1991. MR93a:11048

[Man72] J. I. Manin, *Parabolic points and zeta functions of modular curves*, Izv. Akad. Nauk SSSR Ser. Mat. **36** (1972), 19–66. MR47:3396

[Mat03] Kazuo Matsuno, *Finite* Λ*-submodules of Selmer groups of abelian varieties over cyclotomic* $\mathbb{Z}_p$*-extensions*, J. Number Theory **99** (2003), no. 2, 415–443. MR1969183 (2004c:11098)

[Maz78] B. Mazur, *Rational isogenies of prime degree (with an appendix by D. Goldfeld)*, Invent. Math. **44** (1978), no. 2, 129–162. MR482230 (80h:14022)

[McC91] W. G. McCallum, *Kolyvagin's work on Shafarevich-Tate groups*, L-functions and arithmetic (Durham, 1989), Cambridge Univ. Press, Cambridge, 1991, pp. 295–316. MR92m:11062

[Mil86] J. S. Milne, *Arithmetic duality theorems*, Academic Press Inc., Boston, Mass., 1986. MR881804 (88e:14028)

[MM91] M. Ram Murty and V. Kumar Murty, *Mean values of derivatives of modular* L*-series*, Ann. of Math. (2) **133** (1991), no. 3, 447–475. MR1109350 (92e:11050)

[MR04] Barry Mazur and Karl Rubin, *Kolyvagin systems*, Mem. Amer. Math. Soc. **168** (2004), no. 799, viii+96. MR2031496 (2005b:11179)

[PS99] B. Poonen and M. Stoll, *The Cassels-Tate pairing on polarized abelian varieties*, Ann. of Math. (2) **150** (1999), no. 3, 1109–1149. MR2000m:11048

[Rub98] K. Rubin, *Euler systems and modular elliptic curves*, Galois representations in arithmetic algebraic geometry (Durham, 1996), Cambridge Univ. Press, Cambridge, 1998, pp. 351–367. MR2001a:11106

[Ser72] J-P. Serre, *Propriétés galoisiennes des points d'ordre fini des courbes elliptiques*, Invent. Math. **15** (1972), no. 4, 259–331. MR0387283 (52:8126)

[Ser98] ———, *Abelian ℓ-adic representations and elliptic curves*, A K Peters Ltd., Wellesley, MA, 1998, With the collaboration of Willem Kuyk and John Labute, Revised reprint of the 1968 original. MR0263823 (41:8422)

[Sil92] J. H. Silverman, *The arithmetic of elliptic curves*, Springer-Verlag, New York, 1992, Corrected reprint of the 1986 original. MR1329092 (95m:11054)

[Sage] W. A. Stein, *Sage: Open Source Mathematics Software*, `http://www.sagemath.org`.

[Ste02] W. A. Stein, *There are genus one curves over* $\mathbf{Q}$ *of every odd index*, J. Reine Angew. Math. **547** (2002), 139–147. MR1900139 (2003c:11059)

[SW08] W. A. Stein and C. Wuthrich, *Computations About Tate-Shafarevich Groups Using Iwasawa Theory*, in preparation (2008).

[Sto05] M. Stoll, *Explicit* 3*-descent in Magma* `http://www.faculty.iu-bremen.de/stoll/magma/explicit-3descent/`.

[Wal85] J.-L. Waldspurger, *Sur les valeurs de certaines fonctions L automorphes en leur centre de symétrie*, Compositio Math. **54** (1985), no. 2, 173–242. MR783511 (87g:11061b)

[Wil95] A. J. Wiles, *Modular elliptic curves and Fermat's last theorem*, Ann. of Math. (2) **141** (1995), no. 3, 443–551. MR1333035 (96d:11071)

[Wil00] ———, *The Birch and Swinnerton-Dyer Conjecture*, `http://www.claymath.org/prize_problems/birchsd.htm`.

[Zha04] Shou-Wu Zhang, *Gross-Zagier formula for* GL(2)*. II*, Heegner points and Rankin L-series, Math. Sci. Res. Inst. Publ., vol. 49, Cambridge Univ. Press, Cambridge, 2004, pp. 191–214. MR2083213

Department of Mathematics, Harvard University, Cambridge, Massachusetts 02138

Department of Mathematics, Princeton University, Fine Hall, Princeton, New Jersey 08544-1000

Department of Mathematics, Princeton University, Fine Hall, Princeton, New Jersey 08544-1000

Department of Mathematics, University of Washington, Seattle, Box 354350, Seattle, Washington 98195-4350

Department of Mathematics, Harvard University, Cambridge, Massachusetts 02138