ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 25, Number 4, Fall 1995

INFINITE DESCENT ON ELLIPTIC CURVES

SAMIR SIKSEK

Dedicated to my parents

ABSTRACT. We present an algorithm for computing an upper bound for the difference of the logarithmic height and the canonical height on elliptic curves. Moreover, a new method for performing the infinite descent on elliptic curves is given, using ideas from the geometry of numbers. These algorithms are practical and are demonstrated by a few examples.

1. Introduction. Recently there has been much interest in the computation of Mordell-Weil groups of elliptic curves, both for specific families of curves (such as in [**3, 4, 25**]), and in the development of new algorithms for computing the Mordell-Weil group (see, for example, [**11**]). Not only is this an interesting problem in itself, but it is also an essential ingredient for the popular algorithm for calculating the integral points on elliptic curves using elliptic logarithms (see any of [**12, 23, 24, 26**]).

Let E be an elliptic curve defined over a number field K. The computation of the Mordell-Weil group naturally falls into two parts:

(1) The 2-descent. Here, with some luck, a basis for $E(K)/2E(K)$ is computed.

(2) The infinite descent. This is the name given to the process by which, given a basis for $E(K)/mE(K)$ for some $m \geq 2$, we can obtain a basis for $E(K)$.

Over the rationals, the best (unconditional) algorithm known to me for the 2-descent is the one given in [**1**] and in [**9**, pp. 68–76]. This has recently been (re-)implemented by J. Cremona as the program `mwrank`. For most curves of reasonably small discriminant `mwrank` can calculate $E(\mathbf{Q})/2E(\mathbf{Q})$ in a very short time. In contrast to this, the method

Received by the editors on February 18, 1995.
1991 *Mathematics Subject Classification.* Primary 11G05, Secondary 11Y16.
Key words and phrases. Elliptic curves, Diophantine equations, computational number theory, Mordell-Weil group.
The author's research was funded by a studentship from the SERC/EPSRC.

found in the literature for performing the infinite descent usually takes longer and is often impossible to carry out in practice. Below we explain why this is so, and we present a new more efficient algorithm for carrying out the infinite descent. Our algorithm is practical, and its practicality is demonstrated by a few examples.

The standard method for infinite descent (see [**20**, pp. 739–742] or [**9**, pp. 58–61]) normally goes via Zagier's theorem, and explicit bounds on the difference between the logarithmic and canonical heights of points on elliptic curves defined over number fields (such as Silverman's given below).

We return to the generality of E being a curve over a number field K.

Theorem 1.1 (Zagier). *Let $B > 0$ be such that*

$$S = \{P \in E(K) : \hat{h}(P) \leq B\} \tag{1}$$

contains a complete set of coset representatives for $mE(K)$ in $E(K)$. Then the set S generates $E(K)$.

Proof. See [**9**, p. 61] or [**20**, p. 740]. □

Theorem 1.2[1] (Silverman [**20**]). *Let K be a number field, and let E/K be given by the Weierstrass equation*

$$E : Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6 \tag{2}$$

whose coefficients are in the ring of integers of K. Let Δ be the discriminant of the equation (2), and let j be the j-invariant of E. Further, let

$$b_2 = a_1^2 + 4a_2 \quad \textit{and} \quad 2^* = \begin{cases} 2 & \textit{if } b_2 \neq 0 \\ 1 & \textit{if } b_2 = 0. \end{cases}$$

Define "height of E" (really of the Weierstrass equation (2)) by

$$\mu(E) = \frac{1}{12}h(\Delta) + \frac{1}{12}h_\infty(j) + \frac{1}{2}h_\infty\frac{b_2}{12} + \frac{1}{2}\log(2^*),$$

where, for $t \in K$,

$$h_\infty(t) = \frac{1}{[K:\mathbf{Q}]} \sum_{v \in M_K^\infty} n_v \log(\max(1, |t_v|)).$$

Then, for all $P \in E(\overline{K})$,

$$h(P) - \hat{h}(P) \leq \frac{1}{12} h(j) + 2\mu(E) + 1.946.$$

Proof. See [**20**]. □

Having obtained a set of generators for $E(K)/mE(K)$ we can compute all the coset representatives for $E(K)/mE(K)$ and hence their canonical heights. If B is an upper bound for these canonical heights, then, by Zagier's Theorem 1.1, we get an upper bound for the canonical heights of all the points of a set S (defined above) which generates $E(K)$. Combining this with Silverman's result 1.2 we get an upper bound B' for the logarithmic heights of all the points of S. It follows that the set S can be enumerated, provided of course that this upper bound is not too large.

Unhappily, as indicated above, practical experience suggests that the upper bound B' involved in this method is often too large. This can be the case for several reasons:

(1) It is possible that the Silverman estimate on the difference between the logarithmic and canonical height is very large.

(2) It is possible that the canonical heights of the generators of $E(K)/mE(K)$ are large.

(3) It is also possible, even though the generators of $E(K)/mE(K)$ have small canonical heights, that some of the coset representatives (particularly if the rank is large) will have large heights.

We stress that the size of the search region for the point of S increases exponentially with B'. To illustrate, if say $K = \mathbf{Q}$, and if $P = (X, Y) \in S$, then we can write $X = x/z^2$ where x and z are in $\mathbf{Z}$ and satisfy $|x| \leq \exp(B')$ and $|z| \leq \exp(B'/2)$. It follows that the search region here is roughly proportional to $\exp(1.5B')$. For a number field K of

degree n over the rationals, the search region is, very roughly, between $\exp(1.5nB')$ and $\exp(2nB')$ in size. Hence, small savings on B' can translate into big savings in the actual size of the search region.

We will adopt a different approach to the infinite descent:

(1) We will give an algorithm which will allow us, in most cases, to calculate a sharper upper bound for the quantity $h(P) - \hat{h}(P)$.

(2) We will show how a basis of a submodule of the torsion-free part of $E(K)$, having full rank, can be enlarged efficiently to a basis for $E(K)$.

The algorithm for infinite descent we will give uses both of these ingredients and involves searching much smaller regions than the above.

In computing our examples we have found Cremona's programs [2] `mwrank` and `findinf` very useful. `findinf` is a program for searching for points on a given curve up to a given logarithmic height. In the little programming we needed, we use the population package `Pari/GP` (see [**16**]). This has many functions for doing arithmetic on elliptic curves, including elliptic logarithms and canonical height computations.

2. The bound on the difference $h(P) - \hat{h}(P)$.

2.1. Preliminaries. Let E be an elliptic curve given by the Weierstrass equation

$$E : Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6 \tag{3}$$

where $a_1, \ldots, a_6$ are in the ring of integers $\mathcal{O}_K$ of a number field K. In this section we shall give an algorithm for obtaining an upper bound for the quantity $h(P) - \hat{h}(P)$. This is based on the traditional method of estimating the difference $h(2P) - 4h(P)$. Generally speaking, when this has been done in the past, it relied on the use of elimination theory, which leads to poor upper bounds. The method we shall give bypasses elimination theory using explicit calculations over some local completions of K.

Apart from Silverman's Theorem 1.2, there are other results which give bounds on the quantity $h(P) - \hat{h}(P)$, most notably in [**27**] and [**10**]. The reason why we make specific comparisons only with Silverman's theorem is that this is currently the most widely used and quoted in the literature.

As our method is very different from Silverman's method for obtaining his estimate 1.2, we have no easy way of deciding a priori which should give the smaller bound. We can only note that, in practice, we have found that our method gives much smaller bounds most of the time, or exceptionally bounds which are slightly better. For example, a straightforward application of Silverman's Theorem 1.2 for the curve

$$Y^2 + Y = X^3 - 7X + 6$$

gives

$$h(P) - \hat{h}(P) \leq 5.4.$$

In [**2**], Buhler, Gross and Zagier derive that

$$h(P) - \hat{h}(P) \leq 0 \quad \text{for all } P \in E(\mathbf{Q}),$$

and we get this also by applying our Theorem 2.1. Needless to say, here our method gave a much better bound than Silverman's. In contrast to this, for the curve

$$Y^2 = X(X^2 - p^2)$$

where p is prime and greater than 2, Silverman's theorem gives

$$h(P) - \hat{h}(P) \leq \log(p) + 4.505$$

and our Theorem 2.1 gives

$$h(P) - \hat{h}(P) \leq \log(p) + 0.347 \quad \text{for all } P \in E(\mathbf{Q}).$$

Here for small primes p our bound looks much better and for large p it looks roughly the same as Silverman's. However, even here, the extra work we had to do to get our bound was worthwhile, since to search for all rational points on the curve of canonical height less than or equal to B, the size of the search region if we apply our bound is roughly

$$1.682p^{1.5}\exp(1.5B),$$

and if we apply Silverman's bound it is roughly

$$860.488p^{1.5}\exp(1.5B).$$

Accordingly, we believe that the small amount of work that goes into obtaining our bound will usually be amply rewarded by the time saved through searching smaller regions.

We employ some standard notation to do with number fields and elliptic curves. Given a number field K, we let M_K be the set of all valuations on K. We write M_K^0 and M_K^∞ for the sets of non-archimedean and archimedean valuations on K, respectively. For an elliptic curve E given by a Weierstrass equation of the form (3), we define some associated constants (see [**21**, p. 46]):

$$\begin{aligned} b_2 &= a_1^2 + 4a_2, \\ b_4 &= 2a_4 + a_1a_3, \\ b_6 &= a_3^2 + 4a_6, \\ b_8 &= a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2, \\ \Delta &= -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6. \end{aligned} \tag{4}$$

Let

$$\begin{aligned} f(X) &= 4X^3 + b_2X^2 + 2b_4X + b_6 \\ g(X) &= X^4 - b_4X^2 - 2b_6X - b_8. \end{aligned} \tag{5}$$

It will be seen that the polynomials f and g arise in the duplication formula for a point on the curve E, and a little study of these polynomials essentially gives us our required bound for $h(P) - \hat{h}(P)$.

As usual, we denote the residue field of a completion K_υ with respect to a non-archimedean prime υ by k_υ, and we denote the canonical map $K_\upsilon \to k_\upsilon \cup \{\infty\}$ by $x \to \bar{x}$. We let π be a prime element for υ, i.e., $\pi \in K_\upsilon$ such that $\upsilon(\pi) = 1$.

Lemma 2.1. *Suppose that υ is a non-archimedean valuation on K and $P = (x, y) \in E(K_\upsilon)$ is such that its reduction $\overline{P} = (\bar{x}, \bar{y}) \in E(K_\upsilon)$ is nonsingular. Then*

$$\max\{|f(x)|_\upsilon, |g(x)|_\upsilon\} = \max\{1, |x|_\upsilon\}^4.$$

Proof. If $|x|_\upsilon > 1$, then $|f(x)|_\upsilon \leq |x|_\upsilon^3$ and $|g(x)|_\upsilon = |x|_\upsilon^4$ and in this case the conclusion is obvious.

Hence we can suppose that $|x|_v \leq 1$. Now we are required to prove that

$$\max\{|f(x)|_v, |g(x)|_v\} = 1.$$

Hence it is enough to show that when $f(x) \equiv 0 \pmod{\pi}$ and $g(x) \equiv 0 \pmod{\pi}$, $\overline{P}$ is singular on $E(k_v)$.

By a change of variable over K_v which is nonsingular modulo π, we may suppose that $(x, y) = (0, 0)$. Now the condition for $(0, 0)$ to be on the Weierstrass equation is that $a_6 = 0$. Moreover, since $f(0) \equiv g(0) \equiv 0 \pmod{\pi}$ we get that $b_6 \equiv b_8 \equiv 0 \pmod{\pi}$. Hence, from the formulae for b_6 and b_8, we get that $a_3 \equiv a_4 \equiv 0 \pmod{\pi}$. This is a sufficient condition for $(0, 0)$ to be singular on $E(k_v)$. □

Here is some more notation which we will find useful:

$$\begin{aligned} f'(X') &= X'^4 f(1/X') \\ g'(X') &= X'^4 g(1/X'). \end{aligned} \tag{6}$$

Further, for each $v \in M_K$, let

$$\begin{aligned} D_v &= \{X \in K_v : |X|_v \leq 1 \text{ and } f(X) \in K_v^2\} \\ D'_v &= \{X' \in K_v : |X'|_v \leq 1 \text{ and if } X' \neq 0, \text{ then } f(1/X') \in K_v^2\}. \end{aligned}$$

Lemma 2.2. *Define constants d_v, d'_v by*

(1) $d_v = \inf_{X \in D_v} \max\{|f(X)|_v, |g(X)|_v\}$,

(2) $d'_v = \inf_{X' \in D'_v} \max\{|f'(X')|_v, |g'(X')|_v\}$.

Then d_v and d'_v are nonzero.

Proof. We begin by noting that the sets D_v, D'_v, are compact subsets of K_v (with respect to the v-adic topology), and hence the infima d_v, d'_v must be attained. If, say d_v was zero, then there would exist $X_1 \in D_v$ such that $f(X_1) = g(X_1) = 0$. However, from [**22**, p. 347], We have that

$$\text{Resultant}\,(f, g) = \text{Resultant}\,(f', g') = \Delta^2$$

where Δ is the discriminant of the elliptic curve E. Accordingly, as this cannot be zero, $d_v \neq 0$. Similarly, $d'_v \neq 0$. □

If E is minimal at some non-archimedean valuation v, then we define

$$c_v = [E(K_v) : E^0(K_v)],$$

i.e., c_v is the Tamagawa index at v.

Lemma 2.3. *For any valuation v on K, let*

$$\varepsilon_v^{-1} = \inf_{(X,Y)\in E(K_v)} \frac{\max(|f(X)|_v, |g(X)|_v)}{\max(1, |X|_v)^4}. \tag{7}$$

Then

(1) *ε_v exists, i.e., the quantity on the right exists and is nonzero. Moreover, $\varepsilon_v^{-1} = \min(d_v, d'_v)$.*

(2) *$\varepsilon_v \geq 1$.*

(3) *If v is non-archimedean, E is minimal at v, and the local Tamagawa index $c_v = 1$, then $\varepsilon_v = 1$.*

(4) *If v is non-archimedean, then $\varepsilon_v = d_v^{-1}$ where d_v is as defined in Lemma 2.2.*

(5) *If v is non-archimedean, and*

$$\left\lfloor \frac{v(4\Delta)}{2} \right\rfloor = n,$$

then $\varepsilon_v \leq |\pi|_v^{-2n}$ (where $\lfloor\ \rfloor$ denotes the integer part of a number).

Proof. Suppose $(X, Y) \in E(K_v)$. Then, by a standard manipulation of the Weierstrass equation (3), we get

$$(2Y + a_1X + a_3)^2 = f(X). \tag{8}$$

Hence, if $|X|_v \leq 1$, then $X \in D_v$ and

$$\frac{\max(|f(X)|_v, |g(X)|_v)}{\max(1, |X|_v)^4} = \max(|f(X)|_v, |g(X)|_v).$$

If $|X|_v \geq 1$, then $X' = X^{-1} \in D'_v$ and

$$\frac{\max(|f(X)|_v, |g(X)|_v)}{\max(1, |X|_v)^4} = \max(|f'(X')|_v, |g'(X')|_v).$$

Hence, it is clear that the quantity on the right of (7) exists and is equal to $\min(d_v, d'_v)$, and so is nonzero (by Lemma 2.2). This proves the first part of the above.

For the second part we note that we may take $(X, Y) \in E(K_v)$ to be arbitrarily close to 0. Hence, X is unbounded with respect to the metric $|\ |_v$, and so

$$\frac{\max(|f(X)|_v, |g(X)|_v)}{\max(1, |X|_v)^4}$$

is arbitrarily close to 1. It follows that $\varepsilon_v^{-1} \leq 1$, and hence that $\varepsilon_v \geq 1$, as required for part 2.

Part 3 is clear from Lemma 2.1.

For part 4 we note that if v is non-archimedean and $|X|_v > 1$, then by the proof of Lemma 2.1,

$$\frac{\max(|f(X)|_v, |g(X)|_v)}{\max(1, |X|_v)^4} = 1,$$

and if $|X|_v \leq 1$, then

$$\frac{\max(|f(X)|_v, |g(X)|_v)}{\max(1, |X|_v)^4} \leq 1,$$

so by the definition of ε_v we get

$$\varepsilon_v^{-1} \leq \inf_{(X,Y)\in E(K_v), |X|_v \leq 1} \max(|f(X)|_v, |g(X)|_v)$$

which immediately gives part 4.

Let us now prove part 5. Let n be as defined in the lemma. Suppose that

$$\inf_{x \in D_v} \max(|f(X)|_v, |g(X)|_v) \leq |\pi|_v^{2n+1},$$

and it is sufficient to derive a contradiction. If this was the case, then there would exist $(X, Y) \in E(K_v)$, with

$$f(X) \equiv g(X) \equiv 0 \pmod{\pi^{2n+1}}.$$

But from equation (8) we must deduce that $f(X) \equiv 0 \pmod{\pi^{2n+2}}$. We now invoke the following identity:

$$4g(X) = (6X^2 + b_2X + b_4)^2 - (8X + b_2)f(X). \tag{9}$$

This is easily verified. It follows that $(6X^2+b_2X+b_4)^2 \equiv 0 \pmod{\pi^{2n+2}}$. Finally, we use the congruence

(10)
$$[48X^2 + 8b_2X + (-b_2^2 + 32b_4)](6X^2 + b_2X + b_4)^2 \equiv -4\Delta \pmod{f(X)}$$

in $\mathbf{Z}[X, a_1, \ldots, a_6]$. This is straightforward but rather tedious to verify (it is a slightly more general form of the congruence on page 51 [**5**]). We can now conclude that π^{2n+2} divides 4Δ as required. □

Definition 2.1. For a non-archimedean valuation v, we let (as usual) $E^0(K_v)$ be the set of points on $E(K_v)$ with non-singular reduction modulo π. It is useful to define $\mu_v = \mu_v(E)$ as follows:

(1) if v is archimedean, then $\mu_v = 1/3$,

(2) if v is non-archimedean and E is not minimal at v, then $\mu_v = 1/3$,

(3) if v is non-archimedean and E is minimal at v, then

$$\mu_v = \begin{cases} 0 & \text{if } [E(K_v) : E^0(K_v)] = 1 \\ 1/4 & \text{if } E(K_v)/E^0(K_v) \cong \mathbf{Z}/2\mathbf{Z} \text{ or } (\mathbf{Z}/2\mathbf{Z})^2 \\ (1 - 1/4^\alpha)/3 & \text{if } E(K_v)/E^0(K_v) \cong \mathbf{Z}/2^\alpha\mathbf{Z} \text{ where } \alpha \geq 1 \\ 1/3 & \text{if } [E(K_v) : E^0(K_v)] \text{ is not a power of 2.} \end{cases}$$

Here we recall that, for non-archimedean v at which E is minimal, the group $E(K_v)/E^0(K_v)$ is either cyclic or is equal to $(\mathbf{Z}/2\mathbf{Z})^2$ (see, for example, Theorem 7.6.1 on page 183 of [**21**]). Hence, the above definition for v covers all the possible cases.

We are now ready to state our main theorem on the bound $h - \hat{h}$.

Theorem 2.1. *Let M_K be a complete set of inequivalent valuations on K. For each $v \in M_K$, let $n_v = [K_v : \mathbf{Q}_v]$. Define a function*

$$\varepsilon : M_K \times E(K) \to \mathbf{R}_{\geq 1} \tag{11}$$

by

$$\varepsilon(v, P) = \begin{cases} 1, & \text{if } M_K^0, E \text{ is minimal at } v, \text{ and } P \in E^0(K_v) \\ \varepsilon_v & \text{otherwise.} \end{cases} \tag{12}$$

Then for all $P \in E(K)$, we have

$$\begin{aligned} h(P) - \hat{h}(P) &\leq \frac{1}{[K:\mathbf{Q}]}\Bigg(\sum_{v\in M_K} \mu_v n_v \log(\varepsilon(v,P))\Bigg) \\ &\leq \frac{1}{[K:\mathbf{Q}]}\Bigg(\sum_{v\in M_K} \mu_v n_v \log(\varepsilon_v)\Bigg). \end{aligned} \tag{13}$$

We note here that if v is non-archimedean, E is minimal at v, and the Tamagawa index $c_v = 1$, then by the definition for μ_v above, and Lemma 2.3, we have that $\mu_v = \log(\varepsilon(v,P)) = \log(\varepsilon_v) = 0$. Hence, only finitely many terms in the above sums are nonzero.

Proof. We begin by noting that, for all $P \in E(K)$, $v \in M_K$,

$$\max(|f(X)|_v, |g(X)|_v) \geq \varepsilon(v,P)^{-1} \max(1, |X|_v)^4 \tag{14}$$

using the definition of ε_v on page 6, and the definition of $\varepsilon(v,P)$ above, and Lemma 2.1.

Now if $P = (X,Y) \in E(K)$, then by the duplication formula (see [**21**, p. 59]) the x-coordinate of $2P$ is $g(X)/f(X)$. Hence, using the product definition for naive heights and Lemma 2.1 above, we get

$$\begin{aligned} H_K(2P) &= \prod_{v\in M_K} \max\{|f(X)|_v, |g(X)|_v\}^{n_v} \\ &\geq \prod_{v\in M_K} (\varepsilon(v,P)^{-1} \max\{1, |X|_v\}^4)^{n_v} \\ &= \Bigg(\prod_{v\in M_K} \varepsilon(v,P)^{-n_v}\Bigg) H_K(P)^4. \end{aligned} \tag{15}$$

Recall that

$$h(P) = \frac{1}{[K:\mathbf{Q}]} \log(H_K(P))$$

and so

$$h(2P) - 4h(P) \geq \frac{1}{[K:\mathbf{Q}]}\Bigg(\sum_{v\in M_K} n_v \log(\varepsilon(v,P)^{-1})\Bigg).$$

Rearranging, we get

$$h(P) \leq \frac{1}{4}h(2P) + \frac{1}{4[K:\mathbf{Q}]}\Bigg(\sum_{v \in M_K} n_v \log(\varepsilon(v,P))\Bigg).$$

Using

$$\hat{h}(P) = \lim_{n\to\infty} 4^{-n} h(2^n P),$$

we get

$$h(P) \leq \frac{1}{[K:\mathbf{Q}]}\Bigg(\sum_{v \in M_K} n_v \Bigg(\sum_{n=1}^{\infty} \frac{1}{4^n} \log(\varepsilon(v, 2^n P))\Bigg)\Bigg) + \hat{h}(P).$$

However, from the definition of the function ε, we find that

$$\log(\varepsilon(v,2^n P)) = \begin{cases} 0 & v \in M_K^0,\ E \text{ is minimal at } v, \\ & \text{and } 2^n P \in E^0(K_v), \\ \log(\varepsilon_v) & \text{otherwise.} \end{cases}$$

It is now an easy matter to show that, for all $v \in M_K$,

$$\sum_{n=1}^{\infty} \frac{1}{4^n} \log(\varepsilon(v,2^n P)) \leq \mu_v \log(\varepsilon(v,P))$$

where μ_v is as defined above. This completes the proof. □

It is apparent from our theorem above that to get an upper bound on $h - \hat{h}$, all that remains is to calculate the values ε_v at the finitely many valuations for which μ_v is not zero: recall these are the cases when either v is archimedean (i.e., where $K_v = \mathbf{R}$ or $\mathbf{C}$), or where v is non-archimedean but E is not minimal at v, or it is minimal but the Tamagawa index $c_v \neq 1$.

We give separate algorithms for calculating $\varepsilon_v = \min(d_v, d_v')^{-1}$ for three different cases:

i) $K_v = \mathbf{R}$

ii) $K_v = \mathbf{C}$

iii) v is non-archimedean.

2.2 Case (i): $K_v = \mathbf{R}$. Suppose that $K_v = \mathbf{R}$. Note that there exists $\sigma \in \mathrm{Gal}(K/\mathbf{Q})$ such that $K^\sigma \subset \mathbf{R}$ and for all $x \in K$, $|x|_v = |x^\sigma|$ where $|\ |$ is the ordinary absolute value. Hence, by replacing f, g, f', g' by f^σ, g^σ, f'^σ, g'^σ if necessary, we can assume f, g, f' and g' are all real polynomials. Now the problem is reduced to finding

$$d_v = \inf_{X \in D_v} \max\{|f(X)|_v, |g(X)|_v\},$$
$$d'_v = \inf_{X' \in D'_v} \max\{|f'(X')|_v, |g'(X')|_v\},$$

where

$$D_v = \{X \in \mathbf{R} : |X| \leq 1 \text{ and } f(X) \geq 0\}$$

and

$$D'_v = \{X' \in \mathbf{R} : |X'| \leq 1 \text{ and either } X' = 0 \text{ or } f(1/X') \geq 0\}$$

are clearly finite unions of intervals. Finally, we use the following elementary lemma.

Lemma 2.4. *If f and g are continuous real functions and I is an interval, then the infimum of the continuous function* $\max\{|f(X)|, |g(X)|\}$ *over the interval I is attained at one of the following points*

(i) *an endpoint of I,*

(ii) *at one of the roots of f, g, $f+g$, $f-g$ in the interval I,*

(iii) *at a turning point of one of the functions f, g.*

Proof. We simply note that at any point in I not listed in (i) or (ii), the function $\max\{|f(X)|, |g(X)|\}$ is equal to one of $\pm f$, $\pm g$ and its infimum must be a local supremum or infimum of f or g. □

Hence, to calculate d_v, we write D_v as a union of intervals (I) and calculate the infimum of $\max\{|f(X)|, |g(X)|\}$ over each interval separately using the above lemma, and then d_v will be the minimum of these (finitely many) infima. Similarly, we calculate d'_v, and then $\varepsilon_v = \min(d_v, d'_v)^{-1}$.

2.3 Case (i): $K_v = \mathbf{C}$. Suppose that $K_v = \mathbf{C}$. In the same way as the real case, we can, if necessary, replace f, g, f', g' by appropriate

conjugates so that

$$d_v = \inf_{X \in D_v} \max\{|f(X)|_v, |g(X)|_v\},$$
$$d'_v = \inf_{X' \in D'_v} \max\{|f'(X')|_v, |g'(X')|_v\},$$

where $D_v = D'_v = D = \{z \in \mathbf{C} : |z| \leq 1\}$ is the closed unit disk. We make use of the following lemma.

Lemma 2.5. *Let f and g be as above. Then the continuous function $h : \mathbf{C} \to \mathbf{R}_{>0}$ defined by*

$$k(z) = \max\{|f(z)|, |g(z)|\}$$

attains its infimum over D at a point z_0 satisfying either

(1) $|z_0| = 1$ (*i.e., it is on the boundary of D*), or

(2) $|f(z_0)| = |g(z_0)|$.

Proof. For each $\rho \in \mathbf{C}$ there are, counting multiplicities, four solutions to the equation $f(X) = \rho g(X)$. In fact, by Cardano's formulae, there exist four functions $\phi_1, \ldots, \phi_4$: $\mathbf{C} \to \mathbf{C}$ such that $\phi_1(\rho), \ldots, \phi_4(\rho)$ are solutions to $f(X) = \rho g(X)$.

Let

$$S = \{\rho \in \mathbf{C} : |\rho| = 1\}.$$

It follows that each $\phi_i(S)$ is a path in $\mathbf{C}$. We note that for all $z \in \mathbf{C}$, $|f(z)| = |g(z)|$ if and only if there exist $\rho \in s$ such that $f(z) = \rho g(z)$ and hence if and only if $z \in \phi_i(S)$ for some i.

Now the paths $\phi_1(S), \ldots, \phi_4(S)$ divide the unit disk D into finitely many connected regions $U_1, \ldots, U_n$. Consider a region U_j; denote the interior of U_j by $\mathrm{int}\,(U_j)$ and its closure by $\overline{U}_j$. We note that the intersection of $\mathrm{int}\,(U_j)$ and $\phi_j(S)$ is empty for $i = 1, \ldots, 4$. Hence, by the connectedness of U_j, we get that either $|f| > |g|$ or $|g| > |f|$ on all of $\mathrm{int}\,(U_j)$. Suppose, without loss of generality, that $|f| > |g|$ on all of $\mathrm{int}\,(U_j)$. Then $k(z) = |f(z)|$ for all $z \in \overline{U}_j$. It is easy to see that f is never zero on $\overline{U}_j$: if f is zero at some point of $\overline{U}_j$, then g is also zero at that point, contradicting Lemma 2.2. Let $w(z) = 1/f(z)$. Then w is holomorphic on $\mathrm{int}\,(U_j)$ and continuous on

$\overline{U}_j$ and so by the maximum modulus theorem of complex analysis (See [**17**, p. 76]), it attains its maximum modulus over $\overline{U}_j$ on the boundary $\overline{U}_j \backslash \text{int}\,(U_j)$. Hence, $k(z) = |f(z)|$ attains its infimum over $\overline{U}_j$ on the boundary $\overline{U}_j \backslash \text{int}\,(U_j)$. But each of these boundaries is a subset of $S \cup \phi_1(S) \cup \cdots \cup \phi_4(S)$. Since the $\overline{U}_j$ cover D, we get that k attains its infimum over D on $S \cup \phi_1(S) \cup \cdots \cup \phi_4(S)$. This is the statement of the theorem. □

It is plain that the lemma is true for f' and g' instead of f and g. Now it is necessary to estimate $\inf\{|f|, |g|\}$ over the boundary S and over the sections of the paths $\phi_i(S)$ inside the unit disc D. We will use the following naive method. Fix some $n \geq 2$ (this should be roughly one more than the number of significant digits we want to determine d_v to). Let $\theta_j = 10^{-n} j$ for $j = 1, \ldots, 10^n$. For each θ_j we solve (numerically) the equation

$$f(X) = e^{2\pi\theta_j} g(X),$$

and let

$$\kappa_j = \min\{\max(|f(e^{2\pi\theta_j})|, |g(e^{2\pi\theta_j})|)\} \cup \{|f(X)| : X \in D \text{ and } f(X) = e^{2\pi\theta_j} g(X)\}.$$

Finally, we take $d_v = \min(\kappa_j)$. Similarly, we estimate d'_v and take $\varepsilon_v = \min(d_v, d'_v)^{-1}$. Of course, this method is crude, and great improvements must be possible, but we will not do this.

2.4. Case (iii): v is non-archimedean. In this section we want to calculate

$$\varepsilon_v^{-1} = \inf_{(X,Y) \in E(K_v)} \frac{\max(|f(X)|_v, |g(X)|_v)}{\max(1, |X|_v)^4}$$

for non-archimedean v. We note by Lemma 2.1 that if the reduction of the curve $E(k_v)$ is nonsingular, then $\varepsilon_v = 1$. Hence, we can assume that E has bad reduction at v and calculate the infimum over the points of $E(K_v)$ which have singular reduction modulo v. To do this, we define the following sequence of sets:

We define U_i for $i = 1, 2, \ldots,$ to be the set of all $X \pmod{\pi^{2i}}$ satisfying

(1) $f(X) \equiv 0 \pmod{\pi^{2i}}$,

(2) $g(X) \equiv 0 \pmod{\pi^{2i-1}}$, and

(3) there exists $X_0 \in K_v$ such that $X \equiv X_0 \pmod{\pi^{2i}}$ and $f(X_0) \in K_v^2$.

And we define V_i for $i = 1, 2, \ldots$, to be the set of all $X \pmod{\pi^{2i}}$ satisfying

(1) $f(X) \equiv g(X) \equiv 0 \pmod{\pi^{2i}}$,

(2) there exists $X_0 \in K_v$ such that $X \equiv X_0 \pmod{\pi^{2i}}$ and $f(X_0) \in K_v^2$.

Lemma 2.6. (1) *Suppose* $v(2) = 0$. *If* $i \geq 1$ *and* $U_i \neq \varnothing$, *then* $V_i = U_i$ *and* $\pi^{2i} | \Delta$.

(2) *Suppose* $v(2) = e > 0$. *If* $U_i \neq \varnothing$ *or* $V_i \neq \varnothing$, *then* $\pi^{2i} \mid 4\Delta$.

Proof. We recall the identity and the congruence we used in the proof of Lemma 2.3

$$4g(X) = (6X^2 + b_2X + b_4)^2 - (8X + b_2)f(X). \tag{16}$$

$$[48X^2 + 8b_2X + (-b_2^2 + 32b_4)](6X^2 + b_2X + b_4)^2 \equiv -4\Delta \pmod{f(X)}. \tag{17}$$

It follows from the first that if $v(2) = 0$ and $X \in U_i$, then

$$(6X^2 + b_2X + b_4)^2 \equiv 0 \pmod{\pi^{2i}}$$

and so $\pi^{2i} \mid g(X)$ and $X \in V_i$. Further, by the congruence, $\pi^{2i} \mid \Delta$, and this completes the proof of the first part. The proof of the second part is similar. □

Corollary 2.1. *If* $v(2) = 0$ *and* $U_1 = \varnothing$, *then* $\varepsilon_v = 1$. *If* $U_j \neq \varnothing$ *and* $U_{j+1} = \varnothing$, *then* $\varepsilon_v = |\pi|_v^{-2j}$.

Hence if $v(2) = 0$, then we compute (U_i) explicitly for $i = 1, 2, \ldots$, until we reach the empty set. Then the value of ε_v is given by the above corollary. Here in calculating the (U_i) it is needed to be

able to test, given $X \pmod{\pi^{2i}}$, if there exists $X_0 \in K_v$ such that $X \equiv X_0 \pmod{\pi^{2i}}$ and $f(X_0) \in K_v^2$. For this, we use a suitable generalization of the algorithm in Lemmas 6 and 7 of [**1**]. This is given in [**19**].

Corollary 2.2. *Suppose that* $v(2) \neq 0$.

(1) *If* $U_1 = \varnothing$, *then* $\varepsilon_v = 1$.

(2) *If* $U_j \neq \varnothing$ *and* $V_j = \varnothing$, *then* $\varepsilon_v = |\pi|_v^{-(2j-1)}$.

(3) *If* $V_j \neq \varnothing$ *and* $U_{j+1} = \varnothing$, *then* $\varepsilon_v = |\pi|_v^{-2j}$.

Hence, if $v(2) \neq 0$, then we compute (U_j) and (V_j) explicitly until one of them is empty. Then we compute ε_v from the above corollary.

2.5. The height modulo torsion. As will be seen in the examples, curves where the bound obtained by Theorem 2.1 is small tend to be those where the Tamagawa indices are trivial at the larger primes which divide the discriminant. This is often not the case where the torsion group is nontrivial. However, the following theorem will show us how to exploit the torsion group in order to reduce the bound obtained.

Theorem 2.2. *Under the notation and hypotheses of Theorem* 2.1, *let* $v_1, \ldots, v_n$ *be the* (*finitely many*) *valuations in* M_K *where the quantities* $\mu_v \log(\varepsilon_v)$ *are nonzero. Suppose* (*for some* $m \leq n$) *that* $v_1, \ldots, v_m$ *are non-archimedean valuations such that* E *is minimal at each of them, and there exists a subgroup* $H \leq \operatorname{Tor}(E(K))$ *such that* H *surjects onto* $E(K_{v_i})/E^0(K_{v_i})$ (*via the natural map*) *for* $1 \leq i \leq m$. *Then for each* $P \in E(K)$, *there exists* $T \in H$, *such that*

$$h(P+T) - \hat{h}(P) \leq \frac{1}{[K:\mathbf{Q}]}\left(\frac{|H|-1}{|H|}\left(\sum_{i=1}^{m} \mu_v n_v \log(\varepsilon_v)\right)\right. \\ \left. + \frac{1}{[K:\mathbf{Q}]}\left(\sum_{i=m+1}^{n} \mu_v n_v \log(\varepsilon_v)\right)\right). \tag{18}$$

Proof. Let

$$H = \{T_1, \ldots, T_k\}.$$

Given any $P \in E(K)$, and $1 \leq i \leq m$, we must have that exactly one of $P + T_j$ has good reduction at v_i. Hence, using Theorem 2.1, we get that

(19)
$$\begin{aligned}\sum_{j=1}^{k} h(P+T_j) - \hat{h}(P) &= \sum_{j=1}^{k} h(P+T_j) - \hat{h}(P+T_j) \\ &\leq \frac{1}{[K:\mathbf{Q}]}\left(\sum_{i=1}^{n} \mu_{v_i} n_{v_i} \sum_{j=1}^{k} \log(\varepsilon(v_i, P+T_j))\right) \\ &\leq \frac{k-1}{[K:\mathbf{Q}]}\left(\sum_{i=1}^{m} \mu_{v_i} n_{v_i} \varepsilon_{v_i}\right) \\ &\quad + \frac{k}{[K:\mathbf{Q}]}\left(\sum_{i=m+1}^{n} \mu_{v_i} n_{v_i} \varepsilon_{v_i}\right).\end{aligned}$$

Hence, for one of the T_j, we must have that

$$\begin{aligned}k(h(P+T_j) - \hat{h}(P)) \leq \frac{k-1}{[K:\mathbf{Q}]}&\left(\sum_{i=1}^{m} \mu_{v_i} n_{v_i} \varepsilon_{v_i}\right) \\ &+ \frac{k}{[K:\mathbf{Q}]}\left(\sum_{i=m+1}^{n} \mu_{v_i} n_{v_i} \varepsilon_{v_i}\right)\end{aligned}$$

which gives us the statement of the theorem. □

2.6. Examples.

Example 2.1.

$$E : Y^2 = X^3 - 73705X - 7526231. \tag{20}$$

We find that the equation is minimal and that its discriminant is

$$\Delta = 1155136043932048 = 2^4 \times 199 \times 362793983647$$

as a product of prime factors. Hence, the Tamagawa's indices will be 1, except possibly at 2, and so from Definition 2.1, all the $\mu_p = 0$ except possibly for $p = 2$, or $p = \infty$. Using `Pari/GP` we find that the

Tamagawa index at 2 is 3. Hence, $\mu_2 = \mu_\infty = 1/3$. To use Theorem 2.1, it remains to calculate ε_2 and ε_∞.

We find that

$$f = 4x^3 - 294820x - 30104924 = 4(x^3 - 73705x - 7526231)$$

and

$$g = x^4 + 147410x^2 + 60209848x + 5432427025.$$

Now if $g \equiv 0 \pmod 2$, then x is odd. But clearly, if x is odd, then $|f|_2 = 1/4$, and $|g|_2 \leq 1/4$. Moreover, $(-137, -1) \in E(\mathbf{Q}) \subseteq E(\mathbf{Q}_2)$ and $|f(-137)| = |g(-137)| = 1/4$. Hence, $\varepsilon_2 = 4$.

In computing ε_∞, we find $D_\infty = \varnothing$ and

$$D'_\infty = [-0.007299, -0.005691] \cup [0, 0.003198].$$

Using Lemma 2.4, we find $\varepsilon_\infty = 2.939442$. Applying Lemma 2.1 we get

$$h(P) - \hat{h}(P) \leq 0.8215047 \tag{21}$$

for all $P \in E(\mathbf{Q})$.

Here we note that Silverman's Theorem 1.2 gives a bound

$$h(P) - \hat{h}(P) \leq 13.0242.$$

Example 2.2. We begin with a curve of Mestre (quoted on page 234 of [**21**])

$$E : Y^2 + Y = X^3 - 6349808647X + 193146346911036. \tag{22}$$

The discriminant of this curve is

$$\Delta = 60259 \times 550469 \times 11241887 \times 722983930261$$

as a product of primes. Since it is not divisible by any squares, we must have that all constants $\mu_p = 0$ for all finite primes p. By definition, $\mu_\infty = 1/3$ and it remains to determine ε_∞. Hence, we write D_∞, and D'_∞ as unions of intervals as described on page 9:

$$D_\infty = [-1, 1]$$

and

$$D'_\infty = [-1, -1.08780 \times 10^{-5}] \cup [0, 2.02512 \times 10^{-5}] \cup [2.35024 \times 10^{-5}, 1].$$

Hence we find that $d_\infty \approx 4 \times 10^{19}$ and $d'_\infty = 0.1289169$. So $\varepsilon_\infty = 7.75693$, and using Theorem 2.1, we get

$$h(P) - \hat{h}(P) \leq \mu_\infty \log(\varepsilon_\infty) = 0.68286 \tag{23}$$

for all points $P \in E(\mathbf{Q})$. We note here that Silverman's Theorem 1.2 gives an upper bound of 21.7782 instead of 0.68286.

It is apparent in the last two examples that the reason why the bound for $h(P) - \hat{h}(P)$ is so small is that all or almost all of the Tamagawa indices were 1. Here is an example where this is not the case:

Example 2.3. We compute the bound for the following curve which is given by Thomas Kretschmer in [**14**, p. 633]

$$Y^2 + XY = X^3 - 5818216808130X + 5401285759982786436. \tag{24}$$

The model given here is the minimal and the discriminant is

$$\Delta = 2^6 \times 3^8 \times 7^2 \times 11^2 \times 29^2 \times 31^2 \times 41^2 \times 47^2 \times 277891391058913.$$

We compute the following table.

TABLE 1.

p	c_p	μ_p	ε_p
2	6	1/3	2^6
3	8	21/64	3^8
7	2	1/4	7^2
11	2	1/4	11^2
29	2	1/4	29^2
31	2	1/4	31^2
41	2	1/4	41^2
47	2	1/4	47^2
∞	–	1/3	518.48024

Hence, we get

$$h(P) - \hat{h}(P) \leq 15.70819.$$

In comparison, Silverman's bound is 27.5866.

Here we note that although our bound is much smaller than Silverman's, it is still somewhat larger for the purpose of the infinite descent (see the continuation of this example on page 33). However, we note that the reduction of the point of order 2

$$Q = [1402932, -701466]$$

is singular at the primes 7, 11, 29, 31, 41, 47. Hence, using Theorem 2.2 we get that for all points $P \in E(\mathbf{Q})$ there is a $T \in \{0, Q\}$ such that

$$h(P + T) - \hat{h}(P) \leq 11.03099.$$

3. The canonical height and results from the geometry of numbers. It is worth recalling at the outset of this section, that in the case when the elliptic curve E has rank 1 over the number field K, the infinite descent can be performed in a much easier way than that described in the introduction. This is well known: suppose $P \in E(K)$ has infinite order, and let us say that P generates $E(K)/2E(K)$. Then,

modulo torsion, $P = nQ$ where $n \geq 1$, and Q generates the free part of $E(K)$. Since P generates $E(K)/2E(K)$, n cannot be even and hence $n = 1$ or $n \geq 3$. If $n \geq 3$, then

$$\hat{h}(Q) \leq \frac{1}{9}\hat{h}(P)$$

and so, if P is not the generator of the free part of $E(K)$, we will find a generator in a much smaller region than that given by Zagier's Theorem 1.1.

In this section we develop a general technique for the infinite descent which is analogous to the reduction of the bound for the rank 1 case given above. It is here that we shall employ the language of lattices. Following [**11**] we define $\hat{E}(K) = E(K)/\mathrm{Tor}\,(E(K))$, where $\mathrm{Tor}\,(E(K))$ is the torsion of $E(K)$. Suppose that $P_1, \ldots, P_r$ generate a sublattice of $\hat{E}(K)$ of full rank (for example, $P_1, \ldots, P_r$ could be a basis of $\hat{E}(K)/m\hat{E}(K)$ for some $m \geq 1$). Suppose that this sublattice had index n. If $n = 1$, then of course, $P_1, \ldots, P_r$ is a basis for $\hat{E}(K)$, and we can easily recover a basis for $E(K)$. We will define the height pairing matrix of $P_1, \ldots, P_r$ as follows:

$$H(P_1, \ldots, P_r) = (\langle P_i, P_j \rangle)_{i,j=1,\ldots,r} \tag{25}$$

where for all P, Q in $E(K)$

$$\langle P, Q \rangle =: \frac{1}{2}(\hat{h}(P+Q) - \hat{h}(P) - \hat{h}(Q)). \tag{26}$$

Let $R(P_1, \ldots, P_r)$ be the determinant of the height matrix $H(P_1, \ldots, P_r)$. If R is the regulator of $E(K)$, it follows that

$$R = \frac{1}{n^2} R(P_1, \ldots, P_r). \tag{27}$$

We recall that the regulator is roughly of the same order of magnitude as the product of the canonical heights of some basis for $\hat{E}(K)$. (See, for example, the proof of Manin's theorem in [**11**]). Hence, if the index n was very large we would expect (by virtue of (27)) there to be points of $\hat{E}(K) - \{0\}$ of very small canonical height. We make this idea precise. Roughly it tells us that if there are no points of $\hat{E}(K) - \{0\}$ of height

smaller than some lower bound, then we can get an upper bound for the index n and hence reduce the infinite descent to checking the index of $P_1, \ldots, P_r$ in $\hat{E}(K)$. We make use of the following lemma from the geometry of numbers.

Lemma 3.1 (Hermite, Minkowski and others). *Suppose*

$$f(\mathbf{x}) = \sum_{i,j=1}^{r} f_{ij} x_i x_j \tag{28}$$

where (f_{ij}) is a symmetric positive definite matrix with determinant

$$D = \det(f_{ij}) > 0. \tag{29}$$

Then there exists a positive constant γ_r such that

$$\inf_{\mathbf{m} \neq 0 \text{ integral}} f(\mathbf{m}) \leq \gamma_r D^{1/r}. \tag{30}$$

Moreover, we can take

$$\begin{aligned} &\gamma_1^1 = 1, \quad \gamma_2^2 = \frac{4}{3}, \quad \gamma_3^3 = 2, \quad \gamma_4^4 = 4, \\ &\gamma_5^5 = 8, \quad \gamma_6^6 = \frac{64}{3}, \quad \gamma_7^7 = 64, \quad \gamma_8^8 = 2^8 \end{aligned} \tag{31}$$

and for $r \geq 9$,

$$\gamma_r = \left(\frac{4}{\pi}\right) \Gamma\left(\frac{r}{2} + 1\right)^{2/r}. \tag{32}$$

Proof. The lemma with constant $\gamma_r = (4/3)^{(r-1)/2}$ was originally due to Hermite. The formula (30) with γ_r given for all r by (32) is the formula for the 'first Minima' in Minkowski's second theorem (see [**6**, p. 260] and [**18**, p. 26] for the formula). The constants $\gamma_1, \ldots, \gamma_8$ given above are, for $1 \leq r \leq 8$, the smallest constants which make the lemma valid (see [**7**, p. 332]).

I'm unaware if the smallest possible values of γ_r have been determined for any $r \geq 9$. □

Lemma 3.2. *Let E be an elliptic curve defined over a number field K. Let R be the regulator of $E(K)$. If the rank r is greater than or equal to 1, then there exists a point Q in $E(K)$ of infinite order such that*

$$\hat{h}(Q) \leq \gamma_r R^{1/r}. \tag{33}$$

Proof. Suppose that $Q_1, \ldots, Q_r$ is a basis for $\hat{E}(K)$. If $Q = \sum_{i=1}^{r} m_i Q_i$, then

$$\hat{h}(Q) = \sum_{i,j=1}^{r} m_i m_j \langle Q_i, Q_j \rangle. \tag{34}$$

Recall that the height pairing matrix $H(Q_1, \ldots, Q_r) = (\langle Q_i, Q_j \rangle)$ is symmetric positive definite, and its determinant is R, the regulator of $E(K)$. It follows from lemma (30) that there exists an $\mathbf{m} \neq \mathbf{0}$ integral such that

$$\hat{h}(Q) = \left(\sum_{i,j=1}^{r} m_i m_j \langle Q_i, Q_j \rangle \right) \leq \gamma_r R^{1/r}. \tag{35}$$

Since $Q_1, \ldots, Q_r$ is a basis for $\hat{E}(K)$ and $\mathbf{m} \neq \mathbf{0}$, Q must have infinite order, and the lemma now follows. □

We now combine the above with the observation (27) to deduce the following theorem.

Theorem 3.1. *Let E be an elliptic curve defined over a number field K. Suppose that $E(K)$ contains no point Q of infinite order with canonical height $\hat{h}(Q) \leq \lambda$ where λ is some positive real number. Suppose that $P_1, \ldots, P_r$ generate a sublattice of $\hat{E}(K)$ of full rank $r \geq 1$. Then the index n of the span of $P_1, \ldots, P_r$ in $\hat{E}(K)$ satisfies*

$$n \leq R(P_1, \ldots, P_r)^{1/2} (\gamma_r / \lambda)^{r/2} \tag{36}$$

where $R(P_1, \ldots, P_r)$ is the determinant of the height pairing matrix and

$$\begin{aligned} &\gamma_1^1 = 1, \quad \gamma_2^2 = \frac{4}{3}, \quad \gamma_3^3 = 2, \quad \gamma_4^4 = 4, \\ &\gamma_5^5 = 8, \quad \gamma_6^6 = \frac{64}{3}, \quad \gamma_7^7 = 64, \quad \gamma_8^8 = 2^8 \end{aligned} \tag{37}$$

and for $r \geq 9$,

$$\gamma_r = (4/\pi)\Gamma(r/2+1)^{2/r}. \tag{38}$$

Proof. By Lemma 3.2, if R is the regulator of $E(K)$, then there exists Q in $E(K)$ of infinite order such that

$$\hat{h}(Q) \leq \gamma_r R^{1/r}.$$

It follows that

$$\lambda \leq \gamma_r R^{1/r}.$$

But $R = (1/n^2)R(P_1, \ldots, P_r)$. Hence

$$\lambda^r \leq \frac{\gamma_r^r R(P_1, \ldots, P_r)}{n^2}.$$

Rearranging, we get the required inequality

$$n \leq R(P_1, \ldots, P_r)^{1/2}(\gamma_r/\lambda)^{r/2}. \qquad \square$$

4. A sub-lattice enlargement procedure. Suppose we are given $P_1, \ldots, P_r$ which is a basis for a sublattice of $\hat{E}(K)$ of full rank. By the methods of the previous section, we can establish an upper bound for n, the index of this sublattice in $\hat{E}(K)$. If $n < 2$, then it is clear that $P_1, \ldots, P_r$ is a basis for $\hat{E}(K)$ and the infinite descent is finished.

Suppose now that the method of the previous section gave us a bound $n \leq \alpha$ for some $\alpha \geq 2$. Here it is necessary to check, for each prime $p \leq \alpha$ whether or not the index n is divisible by p. Equivalently, we must determine if there exist $a_1, \ldots, a_r \in \mathbf{Z}$, not all divisible by p, such that

$$\sum a_i P_i = pQ \tag{39}$$

for some $Q \in \hat{E}(K)$.

It is clear that in checking this we can assume that $|a_i| \leq p/2$. This leaves us with a finite number of equations of type (39) to solve. We

explain how these may be solved later. However, as these equations can be many, it is useful to start with some sieving. In practice, we have found the sieving described below to be very effective.

4.1. Sieving. In the notation above, given a prime $p \leq \alpha$, we let $P_{r+1}, \ldots, P_{r+s}$ be a basis for $\mathrm{Tor}\,(E(K))/p\mathrm{Tor}\,(E(K))$, where $\mathrm{Tor}\,(E(K))$ is the torsion subgroup of $E(K)$ (and so typically $s = 0$). We let

$$V_p = \Bigg\{ \bar{\mathbf{a}} \in \mathbf{F}_p^{r+s} : \text{ if } \mathbf{a} \in \mathbf{Z}^{r+s} \text{ and } \mathbf{a} \equiv \bar{\mathbf{a}} \pmod{p} \text{ then } \sum_{i=1}^{r+s} a_i P_i \in pE(K) \Bigg\}.$$

It is clear that V_p is an $\mathbf{F}_p$-linear subspace of $\mathbf{F}_p^{r+s}$ and that the index n is divisible by p if and only if $V_p \neq \{0\}$.

Suppose that $v \in M_K^0$ is a prime such that:

(1) E has good reduction at v,

(2) $|E(k_v)|$ is divisible by p but not by p^2.

Write $|E(k_v)| = lp$ where p does not divide l.

We let π be a uniformizer at v and compute $P_i' \equiv lP_i \pmod{\pi}$. If $P_i' \equiv 0 \pmod{\pi}$ for $i = 1, \ldots, r+s$, then the sieving modulo π will give us nothing and we should start with another $v \in M_K$ satisfying the two conditions above. However, suppose, say that P_1' is not $0 \pmod{\pi}$. We note that the subgroup $lE(k_v)$ of $E(k_v)$ is cyclic of order p and contains $P_1', \ldots, P_{r+s}'$; in particular, $P_1' \pmod{\pi}$ generates $lE(k_v)$. By computing all the multiples of $P_1' \pmod{\pi}$, we determine m_i such that $P_i' \equiv m_i P_i' \pmod{\pi}$. Hence, if $(\bar{a}_1, \ldots, \bar{a}_{r+s}) \in V_p$, we must have that

$$\sum m_i \bar{a}_i = 0 \tag{40}$$

in $\mathbf{F}_p$. This gives us a relation that must be satisfied by the vectors in V_p. If we were to compute $r+s$ independent relations by this method, then $V_p = \{0\}$, and the index would not be divisible by p.

At the very least, our hope is that by sieving modulo a few of these primes π, we have reduced V_p to being in a much smaller subspace of

$\mathbf{F}_p^{r+s}$, and so we have considerably reduced the number of equations of type (39) to be checked.

Our method of sieving has an obvious gap, which is to find $v \in M_K$, for which $|E(k_v)|$ is divisible by p but not p^2. At least the second assumption is not always attainable (for example, if Tor $(E(K))$ had a subgroup of order p^2). So we note that the assumption that p^2 does not divide $|E(k_v)|$ can be easily circumvented after determining the structure of the p-Sylow subgroup of $E(k_v)$, as the reader may readily verify. However, the assumption that p divides $|E(k_v)|$ is essential to the idea of the sieving.

If primes $v \in M_K$ satisfying the conditions above exist, we hope to uncover some by computing sufficiently many $|E(k_v)|$. For $K = \mathbf{Q}$, there exist efficient methods of computing $|E(\mathbf{F}_q)|$ for primes q, and judging from [**8**, pp. 396–398], these have become very impressive.

4.2. Solving the equation $P = pQ$. If the sieving described above has not been entirely successful in proving that $V_p = \{0\}$, then it will leave us with a subspace V_p' of $\mathbf{F}_p^{r+s}$, containing V_p (V_p' is simply the set of all solutions to the equations (40)). Here it is useful to take a projective subset of V_p', which we denote by S_p; we will let S_p be a subset of $\mathbf{Z}^{r+s}\backslash\{0\}$ with the following properties

(1) if $(b_1, \dots, b_{r+s}) \in S_p$, then $|b_i| \leq (p-1)/2$ unless $p = 2$ in which case $b_i = 0$ or 1,

(2) for every $(\bar{a}_1, \dots, \bar{a}_{r+s}) \in V_p/\{0\}$, there exists exactly one $(b_1, \dots, b_{r+s}) \in S_p$ such that $(\bar{a}_1, \dots, \bar{a}_{r+s}) \equiv \beta(b_1, \dots, b_{r+s}) \pmod{p}$ for some $\beta \in \mathbf{F}_p$.

It is clear that all that remains is to check, for all $(b_1, \dots, b_{r+s}) \in S_p$, if

$$\sum_{i=1}^{r+s} b_i P_i = pQ \tag{41}$$

for some $Q \in E(K)$.

For each $(b_1, \dots, b_{r+s}) \in S_p$, the equation (41) has exactly p^2 solutions in $E(\mathbf{C})$, and it is not at all difficult to find these p^2 possible $Q = (x, y) \in E(\mathbf{C})$ with $x, y \in \mathbf{C}$ computed as accurately as is desired using elliptic logarithms (see [**8**]). This leaves us with the problem of

deciding, given a sufficiently accurate computation of $x, y \in \mathbf{C}$, whether or not these are in our number field K. We make use of the following lemma.

Lemma 4.1. *Suppose the elliptic curve E is given by Weierstrass equation (3) with $a_1, \ldots, a_6 \in \mathcal{O}_K$, and suppose that $P = nQ$, where $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ are in $E(K)\backslash\{0\}$. If $v \in M_K$ and $v(x_2) < 0$, then $v(x_1) \leq v(x_2)$.*

Moreover, if $c \in \mathcal{O}_K$ is such that $cx_1 \in \mathcal{O}_K$, then $cx_2 \in \mathcal{O}_K$.

Proof. Let E' be the minimal Weierstrass equation at v, and let $(x', y') \in E'(K_v)$ correspond to coordinates $(x, y) \in E(K_v)$. Then by [**21**, p. 172] there exist $u, r, t, s \in \mathcal{O}_v$ such that

$$x = u^2 x' + r, \qquad y = u^3 y' + u^2 s x' + t.$$

If $v(x) < 0$, then $v(x') = v(x) - 2v(u)$, where $v(u) \geq 0$. Hence, it is sufficient to assume that $v(x_2') < 0$ and show that $v(x_1') \leq v(x_2')$.

Let $v(x_2') = -2m$, where $m \in \mathbf{Z}$ (as is well known, $v(x_2') < 0$ implies that $3v(x_2') = 2v(y_2')$ and hence that $v(x_2')$ is even). Then the subset

$$E'_m(K_v) = \{(x', y') \in E'(K_v) : v(x') \leq -2m\} \cup \{0\}$$

is a subgroup of $E'(K_v)$ (see, for example, [**21**, p. 187]). Hence, $P' \in E'_m(K_v)$ and $v(x_1') \leq -2m = v(x_2')$.

This concludes the proof of the first part of the lemma. The second part is now obvious. □

Hence, given $(b_1, \ldots, b_{r+s}) \in S_p$, we calculate $P = (x_1, y_1) = \sum b_i P_i$, and find $c \in \mathcal{O}_K$ such that $cx_1 \in \mathcal{O}_K$. If $P = pQ$, with $Q = (x_2, y_2) \in \mathcal{O}_K$, then $cx_2 \in \mathcal{O}_K$ by the above lemma. So if we compute the p^2 values x_2 accurately enough[3] we can determine if any of the cx_2 is expressible as a $\mathbf{Z}$-linear combination of any $\mathbf{Z}$-basis for $\mathcal{O}_K$, using an LLL-based algorithm such as the one given on page 100 of [**8**]. (Of course, if $K = \mathbf{Q}$, then we can be much more down to earth. We simply calculate the $x_2 s$ accurately enough to see if any of cx_2 is an integer to many decimal places.) If any cx_2 seems to equal an element

$a \in \mathcal{O}_K$, then we can substitute a/c for x in the equation for E and ask if there is a solution $y \in K$.

If we have found that none of the equations (41) is soluble with $Q \in E(K)$, then we have proven that the index is not divisible by p, and we can proceed to the next prime until we reach α, our upper bound for the index. However, if we find that $\sum b_i P_i = pQ$ with $Q \in E(K)$, then there is a $1 \le j \le r$ such that p does not divide b_j. Here we replace P_j by Q. The index of the sublattice generated by the new $P_1, \dots, P_r$ in $\hat{E}(K)$ is $\le \alpha/p$. In any case, we continue until we get to show that the index is 1.

5. Examples.

Example 5.1. Here we return to Example 2.1

$$E : Y^2 = X^3 - 73705X - 7526231.$$

We recall that we established

$$h(P) - \hat{h}(P) \le 0.8215 \tag{42}$$

for all $P \in E(\mathbf{Q})$. It is easy to show that this curve has no torsion. Using Cremona's program `mwrank`, we found that the 2-part of the Tate-Shafarevich group is trivial, that the rank is 4, and that a basis for $E(\mathbf{Q})/2E(\mathbf{Q})$ is

$$P_1 = (-137, -1), \qquad P_2 = (-157, -419),$$
$$P_3 = (-175, -113), \qquad P_4 = (413, -5699);$$

this the program did in approximately 1.5 minutes.

The determinant of the height pairing matrix of $P_1, \dots, P_4$ is 248.987. We search for points of logarithmic height ≤ 5 using Cremona's program `findinf`. The search takes a few seconds and turns up only one point: $P_1 = (-137, -1)$. This has canonical height 4.41996. We note that had there been any point of canonical height ≤ 4.1, then its logarithmic height would have been $\le 4.1 + 0.8215 < 5$ and would have been uncovered by the search. Hence there are no points of canonical height ≤ 4.1. Using Theorem 3.1, we find that the index of the span of $P_1, \dots, P_4$ is ≤ 1.88. Hence we have found the Mordell-Weil group.

Next we compare our method to that outlined in the introduction. We recall that if $(X, Y) \in E(\mathbf{Q})$, then we can write $X = x/z^2$ where $x, z \in \mathbf{Z}$. Hence, to search up to logarithmic height 5, our search region on x, z is

$$-148 \leq x \leq 148, \qquad 1 \leq z \leq 12.$$

We note that, had we used Zagier's 1.1 on page 2, we would be required to enumerate all the points on $E(\mathbf{Q})$ of canonical height ≤ 13.5831. If we combine this with our estimate (42) above, we must list all points with logarithmic height 14.4046. The corresponding search region is

$$-1802346 \leq x \leq 1802346, \qquad 1 \leq z \leq 1321.$$

To search this region is possible using a well-written program such as `findinf` mentioned above, but this would take a few hours on a work station.

Moreover, we note that if we had to use Silverman's bound on the difference $h(P) - \hat{h}(P)$ as well as Zagier's lemma, we would have to search for all points on $E(\mathbf{Q})$ with logarithmic height ≤ 26.6073. Then the search region would be

$$-359255618029 \leq x \leq 359255618029, \qquad 1 \leq z \leq 599379.$$

Finally, at the suggestion of Dr. Cremona, we compute the following table to give another illustration of how reasonable our bound of 0.8215 is.

TABLE 2.

P	$h(P)$	$\hat{h}(P)$	$h(P) - \hat{h}(P)$
P_1	4.9199809	4.4199587	0.50002214
P_2	5.0562458	4.4416097	0.61463607
P_3	5.1647859	4.4605122	0.70427372
P_4	6.0234476	5.8817481	0.14169942

Example 5.2. We return here to Mestre's curve:

$$E : Y^2 + Y = X^3 - 6349808647X + 193146346911036. \tag{43}$$

We recall that in Example 2.2 we proved that

$$h(P) - \hat{h}(P) \leq 0.682862 \tag{44}$$

for all points $P \in E(\mathbf{Q})$. Mestre (see [**15**]) has shown that this curve has rank at least 12 and has given 12 independent points (Mestre in fact gave a nonminimal model of the curve, and the equation (43) which we will work with is the minimal model). Moreover, he has shown that the standard conjectures[4] imply that the rank is 12. Here we will not take on the task of determining a rank unconditionally[5]; we will simply assume that the rank is 12, and obtain a basis from the points given by Mestre. Here is a list of the points that Mestre gave (after applying the change of variable which takes the points onto our minimal model (43)):

$$\begin{aligned} P_1 &= [49421, 200114], & P_2 &= [49493, 333458], \\ P_3 &= [49513, 362258], & P_4 &= [49632, 502899], \\ P_5 &= [49667, 538049], & P_6 &= [49797, 654674], \\ P_7 &= [49899, 735713], & P_8 &= [50012, 818375], \\ P_9 &= [50165, 921837], & P_{10} &= [50215, 954017], \\ P_{11} &= [50823, 1305633], & P_{12} &= [51108, 1454591]. \end{aligned}$$

Here we proceeded with the sieving first. We used `Pari/GP`, which calculates $|E(\mathbf{F}_q)|$ for prime q using the Shanks-Mestre algorithm (see [**8**, p. 397]). We found that it took roughly 1 second to compute $|E(\mathbf{F}_q)|$ for the first 200 primes q (i.e., for all the primes ≤ 1223). We wrote a program which does the following: for each prime $2 \leq p \leq 11$, it lists all the primes $q \leq 1223$ for which $|E(\mathbf{F}_q)|$ is divisible by p but not p^2 as recommended by our sieving algorithm in Section 4.1. Next, for each prime q satisfying these conditions, it computes a relation modulo p, which must be satisfied by the vectors in V_p as defined in Section 4.1 using the idea described there; if it finds 12 independent relations, then the rank of V_p is 0 and the index is not divisible by p. For each of the primes p, the program continues computing relations until the rank of the relations is 12 or until there are no more primes $q \leq 1223$ satisfying the conditions described. The program took roughly 25 seconds to run and output that for all the primes $p \leq 11$ the rank of relations found is 12 except for $p = 2$ where the rank was 10. We note that there are

47 primes q in the above range satisfying the criterion that 2 divides $|E(\mathbf{F}_q)|$ but 4 does not. Hence, it seems very probable that the index is divisible by 2. Calculating the kernel of the relations obtained, we get that

$$V_2' = \operatorname{span}\{(1,0,0,1,1,0,1,0,1,1,0,0), \\ (1,0,0,1,1,0,0,1,0,0,0,1)\} \pmod 2.$$

Hence we want to test if any of the three points $P_1+P_4+P_5+P_7+P_9+P_{10}$, $P_1+P_4+P_5+P_8+P_{12}$, $P_7+P_8+P_9+P_{10}+P_{12}$ is 2-divisible in $E(\mathbf{Q})$. Using `Pari/GP`, we calculate the periods of E and the 2-division points[6] of the first two points. We get for each one a division point which is integral to 50 decimal places. We checked that these give us integral points on the curve. We replace our old P_7 and P_8 with these two new points:

$$P_7 = [38756, -2294721], \qquad P_8 = [208314, 88938858],$$

thus gaining index 4.

We repeat the sieving for $p = 2$. This time the rank of relations obtained for $p = 2$ is 11. We find that if the index is still divisible by 2, then $P_3 + P_5 + P_6 + P_8 + P_{10} + P_{11} + P_{12}$ must be 2-divisible in $E(\mathbf{Q})$. Here none of the 2-division points of this were integral, and we used Lemma 4.1 to recover a rational 2-division point. This becomes our new P_3:

$$P_3 = \left[\frac{2739835340}{5041}, \frac{141949849330392}{357911}\right].$$

Repeating the sieve described for $p = 2$, we find that the rank of relations obtained is 12, and hence the index of the span of our new $P_1, \ldots, P_{12}$ is not divisible by 2. Moreover, this index is not divisible by any prime $3 \le p \le 11$ since the index of the span of the original points was not.

We return to the sieving again. We calculate $|E(\mathbf{F}_q)|$ for the first 2500 primes q (i.e., all the primes $q \le 22307$), and we extend our range for the prime p to all the primes ≤ 200. It took `Pari/GP` roughly 25 seconds to compute all the $|E(\mathbf{F}_q)|$ for all the primes $q \le 22307$. Our

program this time took about 10 minutes to stop. In each case, the rank of relations computed was 12 except for $p = 167, 179, 191$ where the ranks were respectively 8, 10, 10. Hence if the index of the span of our new $P_1, \ldots, P_{12}$ is not 1, then it must be ≥ 167.

The determinant of the height matrix of $P_1, \ldots, P_{12}$ is

$$R(P_1, \ldots, P_{12}) = 586593208.77747$$

and computing γ_{12} we get 3.81181 according to formula (38). Hence, Theorem 3.1 gives us that if there are no rational points on E with canonical height $\leq \lambda$ then the index of the span of $P_1, \ldots, P_{12}$ in $E(\mathbf{Q})$ satisfies:

$$n \leq \frac{74295365.4988}{\lambda^6}.$$

Using this inequality we find that if there were no points of canonical height ≤ 8.73, then the index would be ≤ 166.9 and we would be finished. Using the inequality (44) we see the need to find all points of logarithmic height ≤ 9.41. We used Cremona's program `findinf` and found none in the range of canonical height ≤ 8.73 (the program took roughly 5 minutes to list all the points of logarithmic height ≤ 9.41). Hence, the points listed below form a basis assuming that the rank (as predicted by the Birch and Swinnerton-Dyer conjecture) is 12:

$$\begin{aligned}
P_1 &= [49421, 200114], & P_2 &= [49493, 333458],\\
P_3 &= \left[\frac{2739835340}{5041}, \frac{141949849330392}{357911}\right], & P_4 &= [49632, 502899],\\
P_5 &= [49667, 538049], & P_6 &= [49797, 654674],\\
P_7 &= [38756, -2294721], & P_8 &= [208314, 88938858],\\
P_9 &= [50165, 921837], & P_{10} &= [50215, 954017],\\
P_{11} &= [50823, 1305633], & P_{12} &= [51108, 1454591].
\end{aligned}$$

Example 5.3. Here we return to the curve

$$Y^2 + XY = X^3 - 5818216808130X + 5401285759982786436. \tag{45}$$

In [**14**], Kretschmer gave this as a curve of (exact) rank 8 with torsion of order 2, but he did not give the points he found on the curve. We

used Cremona's program `mwrank` and it gave a basis for $E(\mathbf{Q})/2E(\mathbf{Q})$:

$$\begin{aligned}
P_1 &= [1410240, -29977314], & P_2 &= [1704648, -661672482],\\
P_3 &= [1421184, -55353570], & P_4 &= \left[\frac{259761720}{125}, -\frac{189069355038}{125}\right],\\
P_5 &= [4740024, 9180268266], & P_6 &= [975216, 808674546],\\
P_7 &= [7028688, -17659711842], & P_8 &= \left[\frac{3418038804}{289}, \frac{195936026213238}{4913}\right],\\
Q &= [1402932, -701466],
\end{aligned}$$

where $P_1, \ldots, P_8$ are of infinite order and Q is a point of order 2. Here it is easy to show that there are no other torsion points. It remains to compute the infinite descent.

Of course, the index of the span of the points above is not divisible by 2 since the points are independent modulo $2E(\mathbf{Q})$. Sieving (as in the above example) with roughly 200 primes (here we excluded all the primes of bad reduction), we were able to show that the index of the span of the given points is not divisible by 5, 7, 11, 13 and detected a possibly 3-divisible linear combination of the points. We found

$$P_4 - P_5 - P_6 - P_7 + P_8 = 3[-2623596, -1613325930]$$

and hence, replacing P_8 by

$$P_8 = [-2623596, -1613325930]$$

we reduce the index by a factor of 3. Repeating the sieving we found that the new index is not divisible by 3. Now we continued the sieving using 15000 primes q and our program proved that the index is not divisible by any prime p less than 500 (this took roughly 30 minutes).

The determinant of the height pairing matrix of the new $P_1, \ldots, P_8$ is 184808.298. Using Theorem 3.1 it is now sufficient to show that there are no points of canonical height ≤ 1.96 whence it would follow that the index is 1. Here we recall that we proved that

$$h(P) - \hat{h}(P) \leq 15.70819,$$

and so that to check that there are no points of canonical height ≤ 1.96, using this it would be necessary to uncover all the points of logarithmic

height ≤ 17.67. We expect that this computation would take roughly 10 days. However, we also proved that for any point P there is a point T which is either 0 or Q such that

$$h(P+T) - \hat{h}(P) \leq 11.03099. \tag{46}$$

Now it is sufficient to enumerate all the points of logarithmic height ≤ 13 and check that none have canonical height ≤ 1.96. We did this in roughly 45 minutes using `findinf`. Hence, it follows that

$$\begin{aligned}
P_1 &= [1410240, -29977314], & P_2 &= [1704648, -661672482],\\
P_3 &= [1421184, -55353570], & P_4 &= \left[\frac{259761720}{125}, -\frac{189069355038}{125}\right],\\
P_5 &= [4740024, 9180268266], & P_6 &= [975216, 808674546],\\
P_7 &= [7028688, -17659711842], & P_8 &= [-2623596, -1613325930]\\
Q &= [1402932, -701466],
\end{aligned}$$

are a basis for $E(\mathbf{Q})$.

Finally, we would like to point out that we were able to obtain the bound (46) using the fact that the torsion group surjects onto $E(\mathbf{Q}_p)/E^0(\mathbf{Q}_p)$ for most of the primes where the Tamagawa index is not 1. Since this will not be the case for most curves we would like to illustrate a third method which can be used to complete the infinite descent when the bound for $h(P) - \hat{h}(P)$ is too large. We note that for all the non-archimedean primes except 2 and 3, the Tamagawa index is either 1 or 2 (see Table 1). In any case, if $P \in E(\mathbf{Q})$ was of infinite order and had canonical height ≤ 1.96, then $2P$ will have canonical height ≤ 7.84 and will have good reduction at all the non-archimedean primes except possibly at 2 or 3. Hence, in the notation of Theorem 2.1, we have

$$\varepsilon(p, 2P) = 1$$

for all primes $p \neq 2, 3, \infty$ and

$$\varepsilon(p, 2P) \leq \varepsilon_p$$

for $p = 2, 3, \infty$. Using the values of ε_p given in Table 1 for the primes $p = 2, 3, \infty$ and Theorem 2.1, we get

$$h(2P) - \hat{h}(2P) \leq 6.39956.$$

Hence, to uncover $2P$ we need to find all the points of logarithmic height ≤ 14.24, and this would not take much longer than the search we have already done. Finally, we would have to test each point found with canonical height ≤ 7.84 to see if it is twice a point.

6. Concluding remarks. Regarding the bound for $h - \hat{h}$, it would be useful to develop better methods for calculating the constants ε_v. Here we suspect that for v non-archimedean, a case-by-case method (reminiscent of Tate's algorithm) exists, and would be best for machine implementation of the algorithm.

We would also like to take this opportunity to point out that it would be of great value if the 2-descent algorithm in [**1**] was improved or extended to cope with elliptic curves defined over number fields.

Acknowledgment. I am greatly indebted to J. Cremona for suggesting this problem to me and guiding my work on it, and also for his detailed reading of this paper and for making many valuable comments and corrections. I am also grateful to N. Smart for asking questions which stimulated many of the developments in this paper, and to an anonymous referee for pointing out many corrections.

ENDNOTES

1. Here it is appropriate to make two comments:

(a) We quoted only one half of Silverman's theorem which is given in [**20**]. Silverman also gives a lower bound for $h - \hat{h}$ but this shall not concern us as it is not needed for the infinite descent.

(b) Silverman's bounds for $h(P) - \hat{h}(P)$ hold for all points P on E defined over any extension of the ground number field K. Our Theorem 2.1 gives a bound for $h(P) - \hat{h}(P)$ for all points $P \in E(K)$. Thus, a bound derived by our method for points on an elliptic curve over a certain number field will not always hold for points defined over extensions of that field.

2. `mwrank` and `findinf` are available by anonymous ftp from `euclid.exeter.ac.uk` (144.173.8.2) in directory pub/cremona. There are executable binaries for both Sun (sparc) and Silicon Graphics (Irix 4) machines.

3. Here, if K has a real embedding, then it is useful to replace K with a real conjugate field at the beginning of the computation, and so reject all the values of x_2 which are not real (taking into account that in floating-point arithmetic over $\mathbf{C}$, a real number is one with a very small imaginary part!).

4. The Birch and Swinnerton-Dyer conjecture, the Taniyama-Weil conjecture, and a suitable Riemann hypothesis.

5. Here `mwrank` would take too long. In the absence of 2-torsion, `mwrank` uses the algorithm for 2-descent described in [**1**] and in [**9**, pp. 68–76]. In this algorithm the size of the search region for the homogeneous spaces is roughly proportional to the square root of the discriminant of the elliptic curve. In cases where the discriminant is very large, such as that for Mestre's curve above, the algorithm is no longer practical. Unfortunately there does not seem to be any unconditional algorithm suited for determining Mordell-Weil groups of curves of large discriminant and no torsion.

6. Our use of terminology here is unconventional. Normally the term '2-division point' denotes a point of order 2. Where as we say Q is a 2-division point of P to mean $P = 2Q$.

REFERENCES

1. B.J. Birch and H.P.F. Swinnerton-Dyer, *Notes on elliptic curves* I, J. Reine Angew. Math. **212** (1963), 7–25.

2. J.P. Buhler, B.H. Gross and D.B. Zagier, *On the conjecture of Birch and Swinnerton-Dyer for an elliptic curve of rank* 3, Math. Comp. **44** (1985), 473–481.

3. A. Bremner, *On the equation* $Y^2 = X(X^2 + p)$, in *Number theory and applications* (R.A. Mollin, ed.), Kluwer, Dordrecht, 3–23, 1989.

4. A. Bremner and J.W.S. Cassels, *On the equation* $Y^2 = X(X^2 + p)$, Math. Comp. **42** (1984), 257–264.

5. J.W.S. Cassels, *Lectures on elliptic curves*, LMS Student Texts, Cambridge University Press, 1991.

6. ———, *Rational quadratic forms*, LMS Monographs, Academic Press, London, 1978.

7. ———, *Introduction to the geometry of numbers*, Springer-Verlag, 1959.

8. H. Cohen, *A course in computational algebraic number theory*, GTM 138, Springer-Verlag, 1993.

9. J.E. Cremona, *Algorithms for modular elliptic curves*, Cambridge University Press, 1992.

10. V.A. Dem'janenko, *An estimate of the remainder term in Tate's formula*, Mat Zametki **3** (1968), 271–278, in Russian.

11. J. Gebel and H.G. Zimmer, *Computing the Mordell-Weil group of an elliptic curve over* Q, in *Elliptic curves and related topics* (H. Kisilevsky and M. Ram. Murty, ed.), CRM Proceedings and Lecture Notes Volume 4, Amer. Math. Soc., 1994.

12. J. Gebel, A. Pethő and H.G. Zimmer, *Computing integral points on elliptic curves*, Acta Arith., to appear.

13. D. Husemoller, *Elliptic curves*, Springer-Verlag, 1987.

14. T.J. Kretschmer, *Construction of elliptic curves with large rank*, Math. Comp. **46** (1986), 627–635.

15. J.-F. Metre, *Construction d'une courbe elliptique de rang* ≥ 12, C.R. Acad. Sci. Paris **295** (1982), 643–644.

16. C. Batut, D. Bernardi, H. Cohen and M. Olivier, *User's guide to PARI-GP* (version 1.38.62), 1994.

17. H.A. Priestley, *Introduction to complex analysis*, Oxford University Press, 1985.

18. C.L. Siegel, *Lectures on the geometry of numbers*, Springer-Verlag, 1988.

19. S. Siksek, *Descents on curves of genus* 1, Ph.D. thesis, Exeter University, 1995.

20. J.H. Silverman, *The difference between the Weil height and the canonical height on elliptic curves*, Math. Comp. **55** (1990), 723–743.

21. ———, *The arithmetic of elliptic curves*, GTM 106, Springer-Verlag, 1986.

22. ———, *Computing heights on elliptic curves*, Math. Comp. **51** (1988), 339–358.

23. N.P. Smart, *S-integral points on elliptic curves*, Proc. Camb. Phil. Soc. **116** (1994), 391–399.

24. N.P. Smart and N.M. Stephens, *Integral points on elliptic curves over number fields*, to appear.

25. R.J. Stroeker and J. Top, *On the equation* $Y^2 = (X + p)(X^2 + p^2)$, Rocky Mountain J. Math. **24** (1994), 1135–1161.

26. R.J. Stroeker and N. Tzanakis, *Solving elliptic diophantine equations by estimating linear forms in elliptic logarithms*, Acta Arith. **67** (1994), 177–196.

27. H.G. Zimmer, *On the difference of the Weil height and the Neron-Tate height*, Math. Z. **147** (1976), 35–51.

Department of Mathematics, University of Exeter, Exeter, EX4 4QE, UK
E-mail address: ssiksek@maths.exeter.ac.uk