# On the Elliptic Logarithm Method for Elliptic Diophantine Equations: Reflections and an Improvement

Roel J. Stroeker and Nikos Tzanakis

CONTENTS



The elliptic logarithm method for the determination of all integral solutions of a given elliptic equation is discussed for equations with associated elliptic curve of moderately large rank. Major attention is given to the question of optimizing the choice of Mordell–Weil basis for the curves in question. A speculative argument suggests that for any curve of rank larger then 8 the calculations involved are unlikely to be feasible. The arguments are illustrated by examples of curves of rank 5, 6, 7, and 8, taken from the literature.

## 1. INTRODUCTION

The history of the elliptic diophantine equation is like a giant tree, old but very much alive, and from its many branches rich fruits can be picked. Not only is its size impressive, but also its age commands respect as the origins of the elliptic equation reach as far back as Diophantus of Alexandria. Indeed, the so-called "ascent" principle — a method, based on simple geometric and algebraic considerations, by which new rational solutions can be constructed from those already known — can be traced to Diophantus' *Arithmetica*. From there, via Fermat's famous "descent", the trail leads up to the celebrated Mordell–Weil finite basis theorem on which modern developments firmly rest. A fascinating account of the early history of this equation runs as a finely woven thread through Weil's history of number theory [1984]. See also [Bašmakova 1974; Scriba 1984] and the references cited there.

In this century Mordell initiated the search for integral solutions of elliptic equations. By Siegel's famous theorem [1929], at most finitely many integral

Mathematics Subject Classification (1991): 11D25, 11G05
Key words and phrases: diophantine equation, elliptic curve, elliptic logarithm


1058-6458/1999 $0.50 per page

solutions exist for any given elliptic equation, and because his result is ineffective the determination of all such solutions becomes a real challenge. For individual equations, the established approaches were almost always purely algebraic with the occasional geometric touch. But since Baker's famous work on linear forms in logarithms of algebraic numbers made Siegel's theorem effective, powerful diophantine approximation techniques took a prominent position in the ranks of successful solution methods.

Looking back over the history of the elliptic equation, one cannot help but be impressed by the great variety of methods and techniques that have been successfully employed to solve individual equations. Despite this, research in the field never stopped, and what is more, new life was breathed into it only recently by a new method which was developed simultaneously and independently by several researchers [Stroeker and Tzanakis 1994; Gebel, Pethő, and Zimmer 1994; Smart 1994]. This approach, which we shall henceforth refer to as the *elliptic logarithm* method — $\mathfrak{Ellog}$ for short — is a harmonious blend of algebraic, analytic and geometric ideas. For more historical comments we refer the reader to [Stroeker and Tzanakis 1994]. Quite recently the $\mathfrak{Ellog}$ method has been generalized to number fields in [Smart and Stephens 1997].

In the next section we shall give a brief description of its fundamental characteristics. In contrast to many earlier methods, $\mathfrak{Ellog}$ is generally applicable, at least in principle. Because of this, naively as it may be, it might cross one's mind that at this point the story of the elliptic equation comes to a natural and happy end. However, on reflection, it would be of interest to see whether significant improvements could be made, especially as $\mathfrak{Ellog}$ heavily relies on non-trivial computations. And, likewise, it would also be nice to know precisely where theory and practice diverge, that is to say, at which point one should expect to come up against insurmountable obstacles in the form of upper bounds for numbers, memory size and cpu-time that cannot be reduced to workable magnitude. In [Bremner et al. 1997; Stroeker and Tzanakis 1994; Stroeker 1995; Tzanakis 1996; Stroeker and de Weger 1999a] examples are given in which serious size problems are avoided, so that these questions were not really addressed.

Here we will do some modest speculation in more demanding circumstances and consider a few more challenging examples. Moreover — and this is the main contribution of this paper — we suggest an improvement of the algorithm introduced in [Stroeker and Tzanakis 1994] for Weierstraß equations and subsequently extended to quartic elliptic equations in [Tzanakis 1996], which often has a favourable impact on the final brute search work that needs to be done, not by lowering the (artificially) large upper bound resulting from the PRINCIPAL INEQUALITY [Stroeker and Tzanakis 1994, (16)], but by significantly reducing the final bound after LLL-reduction. In summary, in this paper we concentrate on the opportunities offered by the $\mathfrak{Ellog}$ method to make choices optimal where possible. In the final section, we shall illustrate these points by means of suitable examples, all taken from the existing literature. We shall also seize the opportunity to show by example that $\mathfrak{Ellog}$ does not have more trouble in dealing with quartic equations than it usually has with Weierstraß equations. This fact agrees with the claim made in the final lines of the Introduction of [Tzanakis 1996].

## 2. PRELIMINARIES

The starting point is a specific elliptic diophantine equation for which we wish to explicitly compute all rational integral solutions. This equation represents an elliptic curve $E$ over the rationals $\mathbb{Q}$, and often it has the standard Weierstraß form

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

with integral coefficients $a_1$, $a_2$, $a_3$, $a_4$, $a_6$, usually satisfying $a_1 = a_2 = a_3 = 0$. However, other models are permitted, such as the binary quartic representation

$$y^2 = f(x)$$

used in [Tzanakis 1996], where $f \in \mathbb{Z}[x]$ is monic and of degree 4, or even less common cubic models like those in [Stroeker and de Weger 1999a] and [Stroeker and de Weger 1999b] of type

$$F(u, v) = 0,$$

for any $F \in \mathbb{Z}[u, v]$ of degree 3, provided it represents an elliptic curve over $\mathbb{Q}$ with a rational point on it. Instead of giving all the details of $\mathfrak{Ellog}$ —

which may be found in [Stroeker and Tzanakis 1994] and [Tzanakis 1996] — we shall merely advance information that is strictly necessary for our purpose.

The $\mathfrak{Ellog}$ method naturally splits up into three distinct parts. In the initial stage essential characteristics of the corresponding elliptic curve are gathered to be used in the second part, like the torsion group, the fundamental period $\omega$ of the Weierstraß $\wp$-function for the minimal Weierstraß model of $E$, the rank $r$ and a basis $\{P_1, \ldots, P_r\}$ for the free part of the Mordell–Weil group. Now any point $P \in E(\mathbb{Q})$ having integral coordinates with respect to the original equation can be uniquely expressed as an integral linear combination of basis elements, allowing for torsion:

$$P = m_1 P_1 + \cdots + m_r P_r + T_0. \tag{1}$$

Here $T_0$ stands for any one of the finitely many torsion points. If torsion is trivial, then $T_0$ is absent from (1). Setting $M = \max_{i=1,\ldots,r} |m_i|$, once an absolute upper bound for $M$ is obtained, all points $P$ satisfying relation (1) can be explicitly calculated, at least in principle.

The next stage of the method forms the body of $\mathfrak{Ellog}$. An upper bound is established for a linear form in elliptic logarithms, which is closely related to (1), and which involves $M$. A basic instrument for this purpose is the group isomorphism $\varphi : E_0(\mathbb{R}) \to \mathbb{R}/\mathbb{Z} = [0, 1)$ defined by

$$P \mapsto \begin{cases} 0 & \text{if } P = O, \\ \dfrac{1}{\omega} \displaystyle\int_{x(P)}^{\infty} \frac{dx}{\sqrt{x^3 + ax + b}} \pmod 1 & \text{if } y(P) \geq 0, \\ -\varphi(-P) \pmod 1 & \text{if } y(P) < 0. \end{cases}$$

Here $E_0(\mathbb{R})$ is the infinite component of the short Weierstraß model $y^2 = x^3 + ax + b$ for $E$. The said linear form generally is of type

$$L(P) := \frac{1}{q}(u_0 + n_1 u_1 + \cdots + n_r u_r + n_0 \omega), \tag{2}$$

where the $u_i := \omega\varphi(P_i)$, for $i = 0, \ldots, r$, are known as the elliptic logarithms of the points $P_i$; the point $P_0$, if not the zero-point, is algebraic of degree at most $D = 3$, and can be easily calculated. The $q$ appearing in (2) is an explicitly known small positive integer, usually 1 or 2. Further, the integers $n_i$, for $i = 1, \ldots, r$, are explicit linear combinations of $m_1, \ldots, m_r$ with small integer coefficients; in many cases $n_i = m_i$ for all $i = 1, \ldots, r$, including those cases in which the equation to be solved is a Weierstraß equation. Non-Weierstraß equations can be found in [Tzanakis 1996, Examples 1, 5, 6, 7], where $q = 2$ or 4. Finally, the integer $n_0$ can be explicitly bounded in terms of $M$.

So the only unknowns in this linear form are the rational integers $n_i$, for $i = 0, \ldots, r$. If we put $N = \max_{i=0,\ldots,r} |n_i|$, then $N \leq \alpha M + \beta$ for some explicitly computable small positive integers $\alpha$ and $\beta$. The essential inequality referred to above looks like

$$|L(P)| \leq c_2 \exp(c_3 - c_1 M^2), \tag{3}$$

for positive constants $c_1$, $c_2$, and $c_3$. The word constant is used here to indicate independence of $M$. The computation of these constants is the main purpose of this part of the algorithm. A direct, rather automatic application of a deep result by Sinnou David [1995] leads to a lower bound for $|L(P)|$ (provided $L(P) \neq 0$), which takes the form

$$|L(P)| > \exp(-c_4 (\log N + c_5)(\log\log N + c_6)^{k+2}), \tag{4}$$

where $k = r$ if $u_0 = 0$ in (2) and $k = r+1$ otherwise. The constants $c_5$ and $c_6$ are small, but $c_4$ is the very large constant

$$c_4 = 2.9 \cdot 10^{6k+12} D^{2k+4} 4^{2(k+1)^2} \times (k+2)^{2k^2+13k+23.3} (\log \mathcal{E})^{-2k-3} \prod_{i=0}^{k} A_i, \tag{5}$$

for some small positive technical constants $A_i$ independent of $k$. Usually $\log \mathcal{E} = 1$ and $D = 1$, 2 or 3. Combining the upper and lower bounds for $|L(P)|$ and taking into account that $N \leq \alpha M + \beta$ leads to the Principal Inequality (see (8)), which gives a large upper bound $M_0$ for $M$. Note that the $c_4$ in (5) differs from the one mentioned in [Stroeker and Tzanakis 1994; Bremner et al. 1997]: the latter occurred in a preliminary version of [David 1995].

The final part of $\mathfrak{Ellog}$ is about reducing this huge upper bound $M_0$ to manageable size. In order to do this, de Weger's implementation [1989] of the LLL-algorithm is used. A brief description will suffice here. Assume, for the sake of simplicity, that the linear form (2) is homogeneous, so that $u_0$ is absent. In order to reduce $M_0$, we apply the lattice basis reduction process to the lattice spanned by the columns of the matrix

$$\mathcal{A} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ [Cu_1] & [Cu_2] & \dots & [Cu_r] & [C\omega] \end{pmatrix},$$

where $C$ is a large constant of the size of $M_0^{r+1}$, and where $[\cdot]$ denotes rounding to the nearest integer. Consider the lattice point

$$l := \mathcal{A}(n_1, \dots, n_r, n_0)^t = (n_1, \dots, n_r, \lambda)^t,$$

where $n_0$ appears in (2), which makes $\lambda$ a good approximation to $qC \cdot L(P)$. From the reduced basis we find a lower bound $d$ for the length of the shortest nonzero lattice vector. The inequality $\|l\| \geq d$ gives us a lower bound for $|L(P)|$, namely

$$|L(P)| \geq \frac{1}{qC} \times \left( \sqrt{d^2 - r(\alpha M_0 + \beta)^2} - \frac{r+1}{2}(\alpha M_0 + \beta) \right), \quad (6)$$

assuming $d$ is large enough — if not, we choose $C$ slightly larger. Together with the upper bound (3) a reduced upper bound $M_1$ for $M_0$ is obtained which is roughly $\sqrt{\log M_0}$; see (10). For a complete description, we refer to [Stroeker and Tzanakis 1994; Tzanakis 1996].

## 3. REFLECTIONS AND SPECULATIONS

In an analysis of $\mathfrak{Ellog}$ it is natural to focus on the question of choice. Since we are working with a fixed elliptic curve, invariants like rank, torsion, regulator, and the like are uniquely determined. From the PRINCIPAL INEQUALITY it is obvious that the rank $r$ plays a major role in the shaping of the upper bound for $M$. How difficult is it to compute the rank of a given elliptic curve over $\mathbb{Q}$? Even for small ranks, this may pose considerable problems if one wishes to establish the rank unconditionally. The best algorithm available seems to be John Cremona's `mwrank` [1992]. Ian Connell's `apecs` [1995] is also very useful. An example of the difficulties involved is given in [Bremner et al. 1997, Section 2]. If one does not shy away from using conjectural assumptions, like the Birch–Swinnerton-Dyer conjectures, things become easier. Even so, for $r > 8$ say, the searching for independent points could easily get out of hand. Fortunately, for most notorious elliptic equations the corresponding curve is of low rank.

Now suppose that the rank $r$ has been established. What is the effect of high rank on $\mathfrak{Ellog}$? Considering the collection of curves to which the elliptic logarithm method has been applied in [Bremner et al. 1997; Stroeker and Tzanakis 1994; Tzanakis 1996; Stroeker 1995], and also in the present paper, it is not difficult to notice the regular behaviour of the upper bound $M_0$ for $M$ obtained before LLL-reduction with respect to the size of $r$, namely

$$M_0 \sim 10^{(5r^2+15r+28)/2}. \quad (7)$$

All curves considered, even the ones of rank 6, 7, and 8, agree to this size, with the exception of the upper bounds found in [Gebel, Pethő, and Zimmer 1994], which are too small to satisfy (7). This is explained by the fact that in that paper the authors erroneously use $\wp$-values instead of elliptic logarithms, that is to say, they consider a *non-homogeneous* linear form in $r$ $\wp$-values where they should have used a *homogeneous* one in $r+1$ elliptic logarithms. The replacement of $r+1$ by $r$ has a considerable diminishing effect on the size of the upper bound $M_0$. After correction the new bounds also agree with (7). Nevertheless, being merely heuristic, this formula only serves as an indication. In order to reduce the bound $M_0$ in the third stage of $\mathfrak{Ellog}$, the elliptic logarithms need to be calculated to a precision of at least $(r+1)\log M_0$ decimal digits. Some careful extrapolation on (7) shows that this means at least 2115 digits for rank 8, which is just feasible as we shall see in Example 4 (Section 5), and no less than 3740 digits for rank 10; for curves of rank 20 and higher this bound gets completely out of reach. This seems to suggest a natural upper bound for the rank of approximately 8 or maybe a bit larger beyond which $\mathfrak{Ellog}$ is not likely to succeed at present.

Once there is no doubt about the rank, finding as many independent rational points is the next step towards a Mordell–Weil basis. The search for points could be very troublesome, since the upper bound for the canonical height of points could be rather large. In such cases, further descent techniques may be successfully applied; see [Merriman et al. 1996]. The process of infinite descent used to construct a

Mordell–Weil basis is very well described in [Siksek 1995].

So far we have not had any real choice. But having a single basis gives us immediate access to infinitely many bases. What is the effect on $\mathfrak{Ellog}$ of changing the Mordell–Weil basis? A natural choice for basis elements is those of least canonical height. This is what John Cremona's `mwrank` and Ian Connell's `apecs` do. But is such a basis also the most natural choice for $\mathfrak{Ellog}$? The answer to this question is not obvious. The element of the PRINCIPAL INEQUALITY that has the most significant effect on the size of $M_0$ (second to the rank $r$ of course) is the constant $c_1$ in (3). This can be seen as follows. Stripped from insignificant elements, the PRINCIPAL INEQUALITY essentially reduces to

$$M^2 < \frac{c_4}{c_1}\log(\alpha M + \beta)\left(\log\log(\alpha M + \beta)\right)^{k+2} \tag{8}$$

for $k = r$ or $r + 1$. Now $c_4$ really depends on $r$ only — see (5) — and is therefore more or less fixed. In view of (8) one would expect $M$ to be roughly of the order of $\sqrt{c_4}$ which is essentially $(4k)^{k^2}$; see (5). Hence, for large $k$ the estimate (7) is rather conservative. Returning to (8), the constant $c_1$ is the least eigenvalue of the positive definite Grammian height-pairing matrix $\mathcal{H} = (\mathcal{H}_{ij})_{r\times r}$, where

$$\mathcal{H}_{ij} = \tfrac{1}{2}\left(\hat{h}(P_i + P_j) - \hat{h}(P_i) - \hat{h}(P_j)\right), \tag{9}$$

and therefore depends solely on the Mordell–Weil basis $\{P_1, \ldots, P_r\}$ of $E(\mathbb{Q})$ modulo torsion. The notation $\hat{h}$ indicates the canonical, or Néron–Tate height function, so (1) implies that

$$\hat{h}(P) = \sum_{1\le i,j\le r} \mathcal{H}_{ij} m_i m_j \ge c_1 M^2.$$

Thus, our choice of a Mordell–Weil basis should reflect our wish to make $c_1$ as large as possible. It is this aspect of the $\mathfrak{Ellog}$ method we shall investigate closely in the next section.

Other choices one may have are few, and seem to have little effect on the $\mathfrak{Ellog}$ method. For example, we start with a given elliptic equation and we also need the short Weierstraß model for $E$. These are related by birational transformations, which do play a minor role in the constants $c_3$ and maybe $P_0$. Their influence however is nothing like the effect the choice of $c_1$ may have on inequality (8), although the latter effect should not be exaggerated either. Indeed, the best $M$ satisfying the PRINCIPAL INEQUALITY, say $M_0$, is always very large. Hence a doubling of the $c_1$-value results in a decrease of $M_0$ by a factor $\sqrt{2}$. This is obviously of very little significance on numbers of size $10^{100}$. It is much more important to realize that an optimal choice of Mordell–Weil basis has a favourable effect on the final upper bound for the coefficients $m_i$ after LLL-reduction has been applied a few times. This is immediately clear from the inequality

$$M_1^2 < \frac{1}{c_1}\left(\log(qc_2C)+c_3-\log\left(\sqrt{d^2-r(\alpha M_0+\beta)^2} - \frac{r+1}{2}(\alpha M_0+\beta)\right)\right); \tag{10}$$

see (6) and (3). The final upper bound for $M$ in (10) is small, usually under 20. So doubling the value of $c_1$ has a much more significant effect on the final upper bound after reduction than on the initial very large upper bound before reduction takes place. This is best illustrated by Example 4 in Section 5.

Finally, we have argued that for any rank $r > 8$ there will be grave computational problems. We add another, rather mundane argument. After applying the LLL-reduction process several times until no further improvement on the bound $M$ is obtained, we will usually end up with a final bound no less than say 6, if $r \ge 4$, but most likely larger than that for larger $r$. Then in the final search for missed integral points, there are $\frac{1}{2}t((2\times 6+1)^r - 1)$ points to be checked, where $t$ stands for torsion. This means more than 10 billion points if $r = 9$ and $t = 2$, and if no substantial number of them can be discarded beforehand this could take a very long time indeed! However, the "inequality trick" discussed in Example 4 on page 147 may reduce considerably the overall search time.

## 4. THE OPTIMAL MORDELL–WEIL BASIS

In the previous section we showed that the size of the bound on $M$ before reduction is governed by the rank $r$, which cannot be altered, and after that by the choice of Mordell–Weil basis, that is to say, by the least eigenvalue $c_1$ of the Grammian height-pairing matrix $\mathcal{H}$. Now for small $r$ it is very easy to search for and find an improved basis by taking random combinations of basis points. Of course, for $r = 1$ there is nothing to improve, but a moderately large final bound for $M$ is computationally nothing to worry about when the rank is so small. Example 1 in Section 5 is meant to illustrate this. For large values of $r$ the optimal choice for the Mordell–Weil basis is not obvious. This may be illustrated by the fact that hardly any "natural" Mordell–Weil basis of an elliptic curve of rank $\geq 5$ is $c_1$-optimal. The remaining examples in Section 5 exemplify this point. For us this fact provided ample motivation to investigate the possibility of standardizing the search for the best Mordell–Weil basis.

### 4A. An Integral Minimax Problem

As before, let $P_1, \dots, P_r$ be free generators of the Mordell–Weil group of an elliptic curve $E/\mathbb{Q}$ and let $\mathcal{H}$ be its $r \times r$ Grammian matrix, whose entries are given by (9). This matrix $\mathcal{H}$ is real, symmetric, and positive definite, and therefore all its eigenvalues are real and positive. It goes without saying that these eigenvalues depend on the choice of generators. Now the point of this discussion is to choose the set of generators so that the least eigenvalue $c_1$ of the corresponding Grammian $\mathcal{H}$ is as large as possible.

To analyze the dependency of $c_1$ on the choice of generators, we consider another set of generators $\{P_1', \dots, P_r'\}$. Then there is an $r \times r$ integer matrix $A = (a_{ij})$, such that $P_i' = \sum_{j=1}^r a_{ij} P_j$, and for the Grammian matrix corresponding with this new set of generators we have

$$\mathcal{H}' = A\mathcal{H}A^t.$$

Since $\det \mathcal{H}' = \det \mathcal{H}$, it follows that $\det A = \pm 1$, so that $A$ is unimodular. In particular, this means that the elements $a_{ij}'$ of $A^{-1}$ are also integral.

Summarizing, our problem may be formulated in the following general terms.

**Integral Minimax Problem.** *Given an $r \times r$ real, symmetric, positive definite matrix $\mathcal{H}$, find an $r \times r$ integral, unimodular matrix $A$ that maximizes the least eigenvalue of the matrix $A\mathcal{H}A^t$. In symbols, determine*

$$\max_A \min_{x \neq 0} \frac{x^t A\mathcal{H}A^t x}{x^t x},$$

*where the maximum runs over all $r \times r$ integral, unimodular matrices $A$, and the minimum runs over all nonzero vectors $x \in \mathbb{R}^r$.*

The second formulation arises because the vectors $x$ that minimize the quotient above are exactly the eigenvectors corresponding to the least eigenvalue; this is called the Rayleigh–Ritz theorem [Horn and Johnson 1985, p. 176] and is an easy consequence of the existence of an orthonomal basis of eigenvectors for $A\mathcal{H}A^t$ (for the inner product defined by $x^t x$).

It is not difficult to prove that this problem is solvable (of course the maximizing $A$ need not be unique). To do this, consider the set

$$S(\mathcal{H}) := \{c_1(A) \mid A \text{ is integral and unimodular}\},$$

where $c_1(A)$ is the least eigenvalue of $A\mathcal{H}A^t$, that is,

$$c_1(A) = \min_{x \neq 0} \frac{x^t A\mathcal{H}A^t x}{x^t x}.$$

$S(\mathcal{H})$ is nonempty, since it contains $c_1 = c_1(I)$. If $c_1$ is the largest element of $S(\mathcal{H})$, there is nothing to prove. Therefore, assume that there exists a $c_1' = c_1(A) > c_1 > 0$ and write $\mathcal{H}' = A\mathcal{H}A^t$. Then, for any $x \in \mathbb{R}^r$ with $x \neq 0$,

$$\begin{aligned} x^t \left( \frac{1}{c_1}\mathcal{H}' - I \right) x &= \frac{1}{c_1}(x^t \mathcal{H}' x - c_1 x^t x) \\ &\geq \frac{1}{c_1}(c_1' - c_1) x^t x > 0, \end{aligned}$$

so that $(1/c_1)\mathcal{H}' - I$ is a positive definite matrix. Then

$$A^{-1}\left( \frac{1}{c_1}\mathcal{H}' - I \right)(A^{-1})^t = \frac{1}{c_1}\mathcal{H} - A^{-1}(A^{-1})^t$$

is also positive definite. Therefore, the main diagonal of the latter matrix has positive elements only. From the inequalities

$$\sum_{j=1}^r (a_{ij}')^2 < \frac{1}{c_1}\mathcal{H}_{ii} \quad \text{for } i = 1, \dots, r, \tag{11}$$

bearing in mind that the $a_{ij}'$ are integers, it follows that there can be only finitely many matrices $A^{-1}$ for any given matrix $\mathcal{H}$. This proves that $S(\mathcal{H})$ is a finite, nonempty set, and the optimal $c_1$ is the maximal element of this set.

Naturally, inequality (11) could serve as a base for our algorithm. However, it seems more convenient to distinguish two separate stages in the computation of the optimal $c_1$.

Set

$$B := A^{-1}(A^{-1})^t \tag{12}$$

and observe that $B$ is integral, symmetric, positive definite, and that $\det B = 1$. Also note that the matrix $\mathcal{H}_1 := (1/c_1)\mathcal{H}$ has least eigenvalue 1.

We now reformulate the integral minimax problem in the following way.

**Integral Minimax Problem (reformulated).**

Part 1. *Given a real, symmetric, positive definite matrix $\mathcal{H}_1$ with least eigenvalue 1, find all integral, symmetric, positive definite matrices $B$ of determinant 1 and such that $\mathcal{H}_1 - B$ is also positive definite.*

Part 2. *Given an integral, symmetric, positive definite matrix $B$ of determinant 1, decide whether there exists an integral, unimodular matrix $A$ satisfying (12) and, if so, find any one such decomposition.*

In Part 2 a single decomposition (12) of $B$ suffices, since any other such decomposition of $B$ gives the same $c_1$. This can be seen as follows. Suppose

$$A_1^{-1}(A_1^{-1})^t = B = A_2^{-1}(A_2^{-1})^t,$$

for integral, unimodular matrices $A_1, A_2$. Then $Q := A_2 A_1^{-1}$ is orthogonal, and hence

$$c_1(A_2) = c_1(QA_1) = c_1(A_1),$$

because similar matrices have the same eigenvalues.

### 4B. The First Stage of the Algorithm

Let the $r \times r$ matrices $\mathcal{H}_1 = (h_{ij})$ and $B = (b_{ij})$ be defined as in Part 1 of the integral minimax problem. Since both $B$ and $\mathcal{H}_1 - B$ are positive definite, all principal minors of these matrices are positive. In particular, all $1 \times 1$ principal minors are positive, which means,

$$0 < b_{ii} < h_{ii} \tag{13}$$

for all $i = 1, \dots, r$, and also all $2 \times 2$ principal minors are positive. On the element level, the latter translates to

$$|b_{ij}| < \sqrt{b_{ii} b_{jj}} =: s_{ij}$$

and

$$|h_{ij} - b_{ij}| < \sqrt{(h_{ii} - b_{ii})(h_{jj} - b_{jj})} =: d_{ij}$$

for all $i, j = 1, \dots, r$ with $i \neq j$, which restricts $b_{ij}$ to the interval

$$\big(\max\{-s_{ij},\, h_{ij} - d_{ij}\},\ \min\{s_{ij},\, h_{ij} + d_{ij}\}\big). \tag{14}$$

This enables us to construct the matrix $B$ by successively enlarging the leading principal submatrices of $B$ by a single row and a single column, starting with $(b_{11})$, thus leading to the following description.

**Algorithm (Stage 1).**

Input: a real symmetric, positive definite matrix $\mathcal{H}_1$ with least eigenvalue 1.

1. Choose $b_{11} \in \mathbb{N}$ such that (13) holds for $i = 1$. Set $B_1 = (b_{11})$, and let $\mathfrak{B}_1$ be the set of all such $1 \times 1$ matrices $B_1$. If $\mathfrak{B}_1$ is empty, stop.
2. Suppose the finite set $\mathfrak{B}_k$ of $k \times k$ integral, symmetric and positive definite matrices $B_k$ has been constructed, for $1 \le k < r$. Now $\mathfrak{B}_{k+1}$ is the set of all possible symmetric $(k+1) \times (k+1)$ matrices $B_{k+1} = (b_{ij})$ with these properties:
   (i) the leading principal $k \times k$ submatrix of $B_{k+1}$ belongs to $\mathfrak{B}_k$,
   (ii) $b_{k+1,k+1} \in \mathbb{N}$ satisfies (13) for $i = k + 1$,
   (iii) $b_{i,k+1} \in \mathbb{Z}$ satisfies (14) for $j = k+1$ and each $i = 1, \dots, k$,
   (iv) $\det B_{k+1} > 0$.

   If $\mathfrak{B}_{k+1}$ is empty, stop. Else, if $k + 1 < r$, repeat Step 2 with $k \leftarrow k + 1$.
3. For each $B_r \in \mathfrak{B}_r$, if $\det B_r = 1$, accept $B_r$ as a possible $B$.

We stress that, when $r$ is not too small, $r \ge 6$ say, and $c_1$ is rather small compared to 1, the number of qualifying $B$ matrices could be very large, so that generating them all becomes infeasible. If we suspect this will happen, we proceed as follows. Once a few qualifying $B$ matrices are known, the process is stopped in order to check for a possible improvement of $c_1$ by means of Stage 2. If so, starting Stage 1 with this improved $c_1$-value necessarily restricts the number of qualifying $B$'s.

Another trick is to artificially enlarge $c_1$ by inserting a multiplication factor $\lambda > 1$, so as to shrink the search intervals resulting from (13) and (14). More precisely, starting from a basis with corresponding $c_1 = c$, we check whether there exists a basis with $c_1 \ge \lambda c$. Varying this factor $\lambda$, we may enlarge $c_1$ step by step, until no improvement seems likely. We then start the process all over again with $\lambda = 1$ and the best $c_1$-value so far obtained.

### 4C. The Second Stage of the Algorithm

Assume that the $r \times r$ matrices $B$ and $A$ are defined as in Part 2 of the reformulated Integral Minimax Problem (page 141). We shall say that $B$ splits if the problem formulated in Part 2 is solvable for this particular matrix $B$.

First we prove:

**Lemma.** *The matrix $B$ splits if and only if, for any real decomposition of $B$ as $R^tR$, the corresponding lattice $\Lambda_R$ generated by the columns of $R$ possesses a basis that is orthonormal with respect to the standard inner product.*

*Proof.* Since $B$ is symmetric and positive definite, real decompositions $R^tR$ of $B$ always exist. Now clearly $A^{-1}(A^{-1})^t = R^tR$ if and only if the matrix $Q := RA^t$ is orthogonal. Since $A^t$ is integral and unimodular, this is equivalent to say that there is an orthogonal matrix $Q$ such that $\Lambda_Q = \Lambda_R$, hence the result. □

Observe that $B$, being positive definite, defines a vector norm

$$\|x\|_B := \sqrt{x^tBx}, \quad \text{for } x \in \mathbb{R}^r.$$

Let $B = R^tR$ be any real decomposition of $B$. Then

$$\|x\|_B = \|Rx\|$$

for $x \in \mathbb{Z}^r$, so $\|x\|_B$ gives the length of the lattice vector $Rx \in \Lambda_R$.

Now suppose that $B$ splits, so that by the Lemma, the lattice $\Lambda_R$ has an orthonormal basis. Let this basis be given by the orthogonal matrix $Q$, so that $R = QU$ for some integral, unimodular matrix $U$. Then for any lattice vector $Rx \in \Lambda_R$ we have

$$\|x\|_B^2 = \|Rx\|^2 = \|QUx\|^2 = \|Ux\|^2 = \|y\|^2 = \sum_{i=1}^{r} y_i^2,$$

where $y = Ux$ is an integral vector. Hence, the only lattice vectors $Rx$ with $\|x\|_B \le 1$ are those with $x = 0$ or $\pm x = U^{-1}e_i$, where $e_i$ denotes the $i$-th standard basis vector of $\mathbb{R}^r$.

**Conclusion.** *If $B$ splits, the lattice $\Lambda_R$ has exactly $2r$ nonzero vectors of length $\le 1$, all of which in fact have length 1. Conversely, if $\Lambda_R$ has exactly $2r$ nonzero vectors of length $\le 1$, then either $\Lambda_R$ does not possess an orthonormal basis and therefore $B$ does not split, or $\Lambda_R$ does have such a basis the elements of which are amongst the $2r$ vectors.*

All of this enables us to use the Pari procedure $\texttt{minim}(B, 1, 2r{+}1)$. This procedure seeks vectors $x \in \mathbb{Z}^r$ with $\|x\|_B \le 1$, and returns a three-component list $u$, where $u[1]$ is the number of vectors computed, $u[2]$ is the maximum $B$-norm found, and $u[3]$ is a matrix whose columns are the vectors computed, only one being given for each pair $x, -x$ and there being at most $2r + 1$ of such pairs. No two columns of this matrix are equal, nor are they each other's additive inverse.

**Algorithm (Stage 2).**

Input: an $r \times r$ integral, symmetric, positive definite matrix $B$ with $\det B = 1$.

1. Set $u := \texttt{minim}(B, 1, 2r{+}1)$.
2. If $u[1] \ne 2r$ or $u[2] < 1$, stop, because $B$ does not split. Else set $U := u[3]$.
3. If $U^tBU$ is the identity matrix — which is equivalent to $RU$ being orthogonal — then stop and return $A = U^t$, the required integral unimodular matrix. Else stop, since $B$ does not split.

Our algorithm obviously relies on Pari's minim procedure. According to the Pari team, in earlier versions of the software sometimes this procedure happened to produce incorrect results, but we no longer have any reason to doubt its correctness.

## 5. EXAMPLES

We have gathered here some examples to illustrate the points made in previous sections. Far from being picked out to show our algorithm to advantage, these examples (apart from the first) were among the most complicated cases we could find in the literature. Some are worked out in detail; with others we make use of existing coverage in the literature.

In the computations we used a variety of machines (notably a number of Pentium PC's and a Sun Sparcstation) at different locations and over a rather long period of time. Wherever it seemed appropriate we have recorded here the machine type and the time it took to carry out the computations.

The first example is an indication of the effect a small $c_1$-value has on the final upper bound for $M$.

**Example 1.** The curve given by the Weierstraß equation

$$y^2 + y = x^3 - x$$

has rank 1, Mordell–Weil basis $\{(0,0)\}$, and the $c_1$-value, which obviously cannot be improved, is rather small, namely $c_1 = 0.0255557\ldots$. The corresponding best upper bound for $M$ is 16, which is large, as was to be expected. Details are given in [Stroeker and de Weger 1999a].

The next set of examples illustrates the fact that finding the best $c_1$-value is not automatic, at least for ranks not smaller than 5.

**Example 2.** Table 1 summarizes three examples, the first two due to Mestre [1986], with short Weierstraß forms

$$y^2 = x^3 - 1642032x + 628747920,$$
$$y^2 = x^3 - 203472x + 18487440,$$
$$y^2 = x^3 - 879984x + 319138704.$$

These three examples are studied in [Gebel, Pethő, and Zimmer 1994] without making use of an optimal Mordell–Weil basis; the authors state the ranks as 6, 5 and 5, but offer no further information as to the conditionality of these claims. All three curves have trivial torsion. It is often very difficult, if not

| | Example 2.1 | Example 2.2 | Example 2.3 |
|---|---|---|---|
| $A, B$ <br> $r$ | $1642032, 628747920$ <br> $6$ | $203472, 18487440$ <br> $5$ | $879984, 319138704$ <br> $5$ |
| $\mathcal{B}_0$ | $P_1 = [432, 108]$ <br> $P_2 = [396, 6372]$ <br> $P_3 = [360, 9180]$ <br> $P_4 = [1044, 7236]$ <br> $P_5 = [108, 21276]$ <br> $P_6 = [36, 23868]$ | $P_1 = [72, 2052]$ <br> $P_2 = [36, 3348]$ <br> $P_3 = [-36, 5076]$ <br> $P_4 = [-72, 5724]$ <br> $P_5 = [396, 108]$ | $P_1 = [540, 1188]$ <br> $P_2 = [576, 1836]$ <br> $P_3 = [468, 3132]$ <br> $P_4 = [612, 3132]$ <br> $P_5 = [432, 4428]$ |
| $c_1(\mathcal{B}_0)$ <br> $M_0$ | $0.21618\ldots$ <br> $1.09 \times 10^{144}$ | $0.40335\ldots$ <br> $2.33 \times 10^{111}$ | $0.34545\ldots$ <br> $9.85 \times 10^{112}$ |
| $M_1 \ldots M_{\text{final}}$ | $97, 17, 15, 15$ | $57, 11, 10, 10$ | $62, 11, 10, 10$ |
| $\mathcal{B}_1$ | $P_1' = [360, 9180]$ <br> $P_2' = [-1296, -24084]$ <br> $P_3' = [1060, -8900]$ <br> $P_4' = [1836, -61668]$ <br> $P_5' = [9/16, -1603611/64]$ <br> $P_6' = [36, 23868]$ | $P_1' = [36, 3348]$ <br> $P_2' = [-36, 5076]$ <br> $P_3' = [432, 3348]$ <br> $P_4' = [-216, 7236]$ <br> $P_5' = [468, 5076]$ | $P_1' = [540, 1188]$ <br> $P_2' = [468, 3132]$ <br> $P_3' = [432, 4428]$ <br> $P_4' = [-684, -24516]$ <br> $P_5' = [720, -7668]$ |
| $c_1(\mathcal{B}_1)$ | $0.53027\ldots$ | $0.46493\ldots$ | $0.49206\ldots$ |
| $M_0$ | $6.94 \times 10^{143}$ | $2.17 \times 10^{111}$ | $1.08 \times 10^{113}$ |
| $M_1 \ldots M_{\text{final}}$ | $62, 10, 9, 8, 8$ | $53, 10, 9, 9$ | $52, 10, 9, 9$ |
| $U$ | $\begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ | $\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 \\ 1 & 0 & -1 & 0 & 1 \end{pmatrix}$ | $\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}$ |

**TABLE 1.** Curves from Example 2. The Weierstraß equation of $E/\mathbb{Q}$ is $y^2 = x^3 - Ax + B$, with Mordell–Weil bases $\mathcal{B}_0 = \{P_i \mid i = 1, \ldots, r\}$, which is the basis given in [Gebel, Pethő, and Zimmer 1994], and $\mathcal{B}_1 = \{P_i' \mid i = 1, \ldots, r\}$ which is the $c_1$-improved basis. Finally, $U = (u_{ij})$ with $P_i' = \sum_{j=1}^{r} u_{ij} P_j$.

impossible, to establish the rank unconditionally for curves without points of finite order. The curve of Example 3 illustrates this point.

In Example 2.1 the first reduction step required a precision of 1250 digits. We used Pari-GP on a 75 MHz Pentium machine to execute the LLL-reduction. Together, the calculation of the $\varphi$-values and the LLL-reduction took approximately one hour runtime. The optimal $c_1$-value of Example 2.1 was obtained by first using a multiplication factor $\lambda = 2.4$. See Section 4.2 for an explanation. Total runtime for this optimization took only a few minutes. In [Gebel, Pethő, and Zimmer 1994] a different definition of canonical height is used, so that double $c_1$-values are obtained: $\lambda_1 = 2c_1$. The upper bounds $M_0$ resulting from the PRINCIPAL INEQUALITY and the LLL-reduced bounds $M_{\text{final}}$ related to the $\mathcal{B}_0$ bases of Table 1 do not agree with the corresponding bounds in [Gebel, Pethő, and Zimmer 1994], which are considerably smaller.

As we noted before in Section 3, the reason for these differences lies in the fact that the $\varphi(P_i)$-values used in [Gebel, Pethő, and Zimmer 1994] are mistaken for the elliptic logarithms $u_i$ (see for instance the six listed $u$-values on page 186 of the paper), which causes a fortunate drop by one of the number of elliptic logarithms in the linear form of elliptic logarithms. Moreover, on page 187 of the same paper, either the notation $\underline{b}_1$ is unfortunate (see the definition on page 184 of the paper), or the authors forgot to insert a factor $2^{6/2}$ which might be another cause for the fact that the final bounds are appreciably smaller than they should be (8 instead of 15 for the first curve). Comparison of the $M_0$ bounds for both bases reveals that the influence of a considerably improved $c_1$-value on the initially computed upper bound is negligible. However, after reduction, the influence is unmistakable, as can be seen from the $M_{\text{final}}$ bounds. This is most noticeable in Example 2.1, thus improving the final search effort for integral points by a factor

$$(31^6-1)/(17^6-1) \approx 36.77.$$

Since the curves of Examples 2.2 and 2.3 both have rank 5, improvements are expected to be less significant, as is indeed the case. Complete lists of integral points are provided in [Gebel, Pethő, and Zimmer 1994].

**Example 3.** In response to a request from Tzanakis, Jaap Top briefly wrote down [Top 1996] examples of some techniques for the construction of elliptic quartics with many integral points. He kindly agreed to let us use this material freely. Since we are mainly interested in curves of higher rank, we restrict our attention to the exemplary illustration of a method explained in [Mestre 1991]. Following Top, consider the 10 integers, 0, $\pm1$, $\pm2$, $\pm3$, $\pm4$, $4q$, and put

$$F(X)=(X^2-4qX)(X^2-1)(X^2-4)(X^2-9)(X^2-16).$$

Next, write $F(X) = h(X)^2 - g(X)$, where $h(X) = X^5 - 2qX^4 + \cdots$ with coefficients chosen in such a way that $g(X)$ has degree 4. Since there is no more to this than simply "completing the square", $g(X)$ is uniquely determined. This $g(X)$ has coefficients in $\mathbb{Z}[q]$ that can be explicitly worked out. Now observe that when $X = \alpha$ is any one of the 10 roots of $F(X) = 0$, then $h(\alpha)^2 = g(\alpha)$. This means that $[\alpha, h(\alpha)/2]$ is a rational point on the curve given by $Y^2 = g(X)$. In particular, for variable $q$, all these curves have at least 10 rational points. It is not unreasonable to expect that some, if not most, of these points will be independent. As it turns out, for $q = 2$ the rank is at least 4, for $q = 3$ and $q = 4$ the rank is at least 6, and for $q = 5$ the rank is at least 9.

In this example we shall consider the case $q = 3$, for which we obtain, after division by 4, the quartic model

$$y^2 = 24784x^4 + 90096x^3 + 114372x^2 + 1376352x + 7096896. \tag{15}$$

The minimal Weierstraß model for this curve is

$$y^2 + xy + y = x^3 - x^2 - 28159452x + 15511281951. \tag{16}$$

It is generally very hard to determine unconditionally the rank and a basis for the Mordell–Weil group of an elliptic curve of moderate or large rank with trivial torsion. See footnote 5 in [Siksek 1995, Example 5.2]. In fact the 2-descent which is necessary to determine the rank is greatly hampered by the absence of points of finite order. For our curve we estimate that `mwrank` would take many hundreds of days on a 75 MHz Pentium machine. However, assuming the Birch–Swinnerton-Dyer conjecture, the Taniyama–Weil conjecture, and a suitable Riemann

| | $\mathcal{B}_0$ (apecs) | $\mathcal{B}_1$ ($c_1$-optimal) |
|---|---|---|
| | $[-21/2, 14220]$<br>$[-612/59, -48170880/3481]$<br>$[-3, -1890]$<br>$[-1, -2402]$<br>$[-156/17, -3034080/289]$<br>$[-324/7, 15878304/49]$ | $[-21/2, 14220]$<br>$[-3, -1890]$<br>$[948/55, 157581648/3025]$<br>$[-1, -2402]$<br>$[-328/51, 12035240/2601]$<br>$[-36/7, 139680/49]$ |
| $c_1(\mathcal{B}_i)$ | $0.42489\ldots$ | $0.52127\ldots$ |
| $M_0$ | $5.84 \times 10^{184}$ | $5.04 \times 10^{144}$ |
| $M_1 \ldots M_{\text{final}}$ | $78, 13, 12, 12$ | $62, 11, 10, 10$ |

**TABLE 2.** Top's curve of rank 6, with equation $y^2 = 24784x^4 + 90096x^3 + 114372x^2 + 1376352x + 7096896$. Two conditional Mordell–Weil bases $\mathcal{B}_0$ and $\mathcal{B}_1$ are given. See Example 3.

conjecture, `apecs` quickly establishes the rank and a Mordell–Weil basis for our curve.

Searching for points of low canonical height with `Seek(5000)`, it took no more than 15 minutes on a 75 MHz Pentium to find a conditional basis $\mathcal{B}_0$ (see Table 2) with $c_1(\mathcal{B}_0) = 0.424899\ldots$. Although not optimal, this $c_1$-value is almost optimal, which renders the present example less exciting than Example 4. The search for the $c_1$-optimal basis $\mathcal{B}_1$ took approximately 20 minutes on the same Pentium: 178 qualifying $B$-matrices were discovered, all of which split. Most of the computation time went into the calculation of $\varphi$-values to 1500 digits precision (almost 5 hours), the first LLL-reduction step with the same precision (2 hours), and the final search, which was done on a 133 MHz Pentium (approximately 29 hours). This search revealed exactly 14 integer solutions $[x, y]$ with $y > 0$ on the quartic (15). Particulars of our calculations are listed in Table 2.

Although we did not set out to do this at first, being at it, we decided to continue and compute all integral solutions of the minimal Weierstraß equation (16) as well. We were rather surprised to find no fewer than 186 of them! For these calculations we used basis $\mathcal{B}_1$ of Table 2, properly transformed to fit Weierstraß equation (16) of course. Further, the LLL-reduction process, starting with initial upper bound $M_0 = 5.04 \times 10^{144}$ and applied several times, produced reduced bounds of $62, 11, 10, 10$. The concluding final search took no more than 1 hour on a 133 MHz Pentium. We first used inequality (3) to exclude all 6-tuples with at least one component absolutely larger than 6 (see the inequality trick on page 147). There is no point in listing all 28 integral solutions of (15) and all 186 integral solutions of (16): they can be reproduced without much effort at any moment, because, as it turns out, the absolute coefficients $|m_i|$ never exceed 2. A complete list of solutions may be obtained from Stroeker's homepage (see address at end of paper).

Here we make an important general remark about the way we carried out the final search for curves of moderately large rank. This remark particularly applies to Example 3 and the two further examples. The final search is about finding all integral $r$-tuples $(m_1, \ldots, m_r)$ satisfying $|m_i| < M$ for $i = 1, \ldots, r$ and some small value of $M$, and for which the corresponding point (1) has integral coordinates with respect to the specific model of the elliptic curve represented by our diophantine equation. At first we thought this could easily be done by Pari or `apecs`, both of which have procedures for calculating linear combinations of points (1) symbolically. We were too optimistic, which we could have foreseen, since symbolic computations are extremely costly in terms of CPU time. Although Pari is much faster than `apecs`, both demand far too much computing time to be of much practical use for our final search computations. Therefore, we turned to the Ubasic language, which combines a very large numerical precision with fast numerical computations. We wrote a search procedure in Ubasic code in which linear combinations of rational points on the curve are calculated with floating point arithmetic to, say, 50 decimal digits precision. A point is recognized by the Ubasic code as integral if its coordinates $x, y$ (with respect to the relevant equation) differ from their nearest integers by at most $10^{-20}$. In Examples 3, 4, and 5 all prospective integral points thus detected were checked by apecs and turned out to be truly integral.

**Example 4.** We consider a curve taken from the collection of Jim Buddenhagen, given by the short Weierstraß form

$$y^2 = x^3 - 20932x - 330140. \tag{17}$$

First, it is easy to show that this curve has trivial torsion. Next we used John Cremona's mwrank to determine the rank of this curve. After running for about 20 minutes on a Sun SparcStation 4 this program proved the rank to be 7 and it produced the possible basis $\mathcal{B}_0$ for the Mordell–Weil group (see Table 3).

Further, mwrank tells us that $\mathcal{B}_0$ generates a subgroup of the Mordell–Weil group of odd index $m$. However, the upper bound of 23.34 for the canonical height of possible extra generators is too large to hope for a quick settling of the Mordell–Weil basis uncertainty about $\mathcal{B}_0$. The least eigenvalue $c_1(\mathcal{B}_0)$ of the Grammian height-paring matrix happens to be extremely small, so even if we could show that $m = 1$, and consequently that $\mathcal{B}_0$ is indeed a basis, it would not be a very good basis for our purpose. So we let apecs 4.2 search for points of small canonical height with the Seek command and with its parameter set to 5000. After a few minutes this turned up $\mathcal{B}_1$, which generates the same subgroup as $\mathcal{B}_0$, because the determinants of the Grammian height-pairing matrices of $\mathcal{B}_0$ and $\mathcal{B}_1$ are equal, namely $R = 1491.0120\ldots$. This $\mathcal{B}_1$ gives a much better $c_1$-value. Again using apecs we established the following inequality between the canonical and logarithmic heights:

$$\hat{h}(P) - \tfrac{1}{2}h(P) \geq -0.69314 \quad \text{for all } P \in E(\mathbb{Q}). \tag{18}$$

Searching for points $P \in E(\mathbb{Q})$ of logarithmic height $\leq 5.5$ only takes a few seconds and produces a list of 30 points, all of which have canonical height $\geq 2.05$, except one, namely $[-18, 202]$ of canonical height $2.04754\ldots$. By (18) we deduce that any point $P$ of canonical height less than 2.0475 has logarithmic height less than 5.5. Since none was found in our search, there are no such points. Now we use Theorem 3.1 of [Siksek 1995] in a slightly adapted form, because the canonical height function Siksek uses differs from the one we use by a factor 2. This yields the following inequality for the index $m$:

$$m \leq \left( \frac{64 \cdot R}{(2 \cdot 2.0475)^7} \right)^{1/2} \leq 2.223.$$

Since $m$ is odd, $m$ can only be 1 and this proves that $\mathcal{B}_1$ is indeed a basis for the Mordell–Weil group.

Starting with $\mathcal{B}_1$, we used our algorithm described in the previous section and after a brief search we found the $c_1$-optimal basis $\mathcal{B}_2$. The corresponding $c_1$-value improves the original $c_1(\mathcal{B}_0)$ by a factor 35! It is now a matter of seconds to produce the

| | $\mathcal{B}_0$ (mwrank) | $\mathcal{B}_1$ (apecs) | $\mathcal{B}_2$ ($c_1$-optimal) |
|---|---|---|---|
| | $[1336, 48542]$<br>$[672, 17002]$<br>$[656, 16378]$<br>$[528, 11654]$<br>$[280, 3970]$<br>$[-16, 26]$<br>$[24658, 3871946]$ | $[-18, 202]$<br>$[-22, 346]$<br>$[-62, 854]$<br>$[-66, 874]$<br>$[-70, 890]$<br>$[-16, 26]$<br>$[-24, 398]$ | $[-16, 26]$<br>$[2849, 151871]$<br>$[232, 2702]$<br>$[-114, 758]$<br>$[-66, -874]$<br>$[-136, -34]$<br>$[402, -7498]$ |
| $c_1(\mathcal{B}_i)$ | $0.01785\ldots$ | $0.25397\ldots$ | $0.60346\ldots$ |
| $M_0$ | $7.95 \times 10^{180}$ | $2.10 \times 10^{180}$ | $1.37 \times 10^{180}$ |
| $M_1 \ldots M_{\text{final}}$ | $407, 69, 66, 65, 65$ | $108, 17, 15, 15$ | $69, 11, 10, 10$ |

**TABLE 3.** Buddenhagen's curve of rank 7, with Weierstraß equation $y^2 = x^3 - 20932x - 330140$. Three Mordell–Weil bases $\mathcal{B}_0$, $\mathcal{B}_1$, and $\mathcal{B}_2$ are given. See Example 4.

first upper bound $M_0$ for the coefficients $m_i$, for $i = 1, \dots, 7$. After that the reduction process produces a final bound $M_{\text{final}} = 10$. This would have been much larger, namely 65, if we had started with $\mathcal{B}_0$, the basis turned up by `mwrank`. Finally, we did a brute force search for all 7-tuples $(m_1, \dots, m_7)$ in the range $-10 \le m_i \le 10$ with $m_1 \ge 0$, generating integral points (1) with $T_0 = \mathcal{O}$. After running on a 133 MHz Pentium for about three days a grand total of $2 \times 88$ integral points $(x, y)$ were found.

As it happens, this final search could have been speeded up significantly if we had applied a simple but very effective trick, which we now describe; we call it the *inequality trick*. Note that for every 7-tuple $(m_1, \dots, m_7)$ corresponding to an integral point (1), inequality (3) must be satisfied. The trick now is based on the heuristic observation that this inequality is rarely satisfied for points (1) with at least one large coefficient $m_i$. The reason is that the elliptic logarithms $u_i$ are more or less randomly distributed — at least we see no reason to assume otherwise — so that the linear form $L(P)$ is rarely very small. In the present example it works as follows. Suppose that $(m_1, \dots, m_7)$ generates an integral point and let $|m_i| \ge 6$ for at least one $i$. Then, in view of (3), the absolute value of the linear form

$$L(P) = L(m_1 P_1 + \cdots + m_7 P_7)$$

must be bounded from above by $1.919 \cdot 10^{-7}$. Checking whether one 7-tuple $(m_1, \dots, m_7)$ in the range $-10 \le m_i \le 10$ (with $m_1 \ge 0$) satisfies this condition or not is considerably less time consuming than checking the corresponding point (1) for integrality. After 7 hours and 35 minutes, a mere 1633 out of almost a billion 7-tuples passed the inequality test and the corresponding points (1) were subsequently tested for integrality. This took no more than 1 minute CPU time and revealed no integral points. The remaining search for 7-tuples in the reduced range $-5 \le m_i \le 5$ took only 54 minutes and produced all integral points. Consequently, the initial computing time of more than three days was shrunk to less than 8 hours. An analogous trick could have been applied to Top's quartic curve of Example 3.

Again we omit a complete listing of all integral solutions as $\max_{1 \le i \le 7} |m_i| = 2$; they may be obtained from Stroeker's homepage.

**Example 5.** Among the several exciting examples of elliptic curves of large rank given in [Siksek 1995], we have selected the curve of rank 8 from Example 5.3. Its minimal Weierstraß equation is

$$y^2 + xy = x^3 - 5818216808130x \\ +5401285759982786436. \quad (19)$$

This curve is due to Kretschmer [1986], but it is Siksek who proves without assuming any conjecture that its rank is 8, and who gives an unconditional Mordell–Weil basis, called $\mathcal{B}_0$ in Table 4. Though the situation here is more advantageous than in Example 3, because the curve has 2-torsion, it was no easy matter to establish without any condition whatsoever that $\mathcal{B}_0$ is a Mordell–Weil basis modulo torsion. It came as no surprise to us that this basis is not $c_1$-optimal. Running our $c_1$-optimal program with multiplication factor $\lambda = 1$ for this basis soon made us realize that the number of qualifying $B$-matrices is way too large. So, repeating the process of using different values for $\lambda$ and stopping as soon as we found a few good $B$'s, after a few steps turned up the free basis $\mathcal{B}_1$ of Table 4 with improved $c_1$-value. Starting with this basis, and using $\lambda = 1$, it took approximately 5 hours on a 75 MHz Pentium to find 10 qualifying $B$'s, 9 of which split. This produced the $c_1$-optimal free basis $\mathcal{B}_2$.

The first reduction step was executed in 2500 digits precision, which took approximately 34 hours CPU time on a Sun Sparcstation 4; this included the time needed to calculate the $\varphi$-values to the same precision. The final search for integral points

$$P = \sum_{i=1}^{8} m_i P_i + \varepsilon Q,$$

where $\{P_1, \dots, P_8\}$ is the $c_1$-optimal free basis, $Q = [1402932, -701466]$ generates the torsion group, $\varepsilon \in \{0, 1\}$, and $m_i \in \mathbb{Z}$ with $|m_i| \le 8$ for $i = 1, \dots, 8$, would have taken a few weeks of computing time on a 133 MHz Pentium had we not applied the inequality trick mentioned in Example 4.

In the present example this trick works as follows. Assume $(m_1, \dots, m_8)$ generates an integral point and let $|m_i| \ge 5$ for at least one $i$. Then for this point (1) (where $T_0$ is either $\mathcal{O}$ or $Q$) inequality (3) must be satisfied with $8.09 \cdot 10^{-9}$ in the right-hand side. All 8-tuples in the range $-8 \le m_i \le 8$

| | $\mathcal{B}_0$ (Siksek) | $\mathcal{B}_1$ ($c_1$-improved) | $\mathcal{B}_2$ ($c_1$-optimal) |
|---|---|---|---|
| | $[1410240, -29977314]$<br>$[1704648, -661672482]$<br>$[1421184, -55353570]$<br>$\left[\frac{51952344}{25}, -\frac{189069355038}{125}\right]$<br>$[4740024, 9180268266]$<br>$[975216, 808674546]$<br>$[7028688, -17659711842]$<br>$[-2623596, -1613325930]$ | $[1365048, 51389034]$<br>$[1437384, 88804830]$<br>$[1284264, -218219910]$<br>$[1410240, -29977314]$<br>$\left[\frac{110804779056}{78961}, -\frac{139162465484094}{22188041}\right]$<br>$\left[-\frac{120141159}{64}, -\frac{1594807366953}{512}\right]$<br>$[1404150, 9858594]$<br>$[1368480, -45144546]$ | $[-2520768, 2013726114]$<br>$[1410240, -29977314]$<br>$[1145136, 489626526]$<br>$\left[\frac{917950008}{361}, \frac{18200617327182}{6859}\right]$<br>$\left[\frac{2288462304}{1849}, -\frac{24701908414938}{79507}\right]$<br>$\left[\frac{5313903}{4}, \frac{1021799877}{8}\right]$<br>$[1368480, -45144546]$<br>$\left[\frac{637573719058}{998001}, \frac{1390151499263611822}{997002999}\right]$ |
| $c_1(\mathcal{B}_i)$ | $0.13586\ldots$ | $1.17637\ldots$ | $1.20392\ldots$ |
| $M_0$ | $2.59 \times 10^{225}$ | $8.77 \times 10^{224}$ | $8.67 \times 10^{224}$ |
| $M_1 \ldots M_{\text{final}}$ | $176, 27, 25, 25$ | $60, 9, 8, 8$ | $59, 9, 8, 8$ |

**TABLE 4.** Siksek's curve of rank 8, with Weierstraß equation $y^2+xy = x^3-5818216808130x+5401285759982786436$. Three Mordell–Weil bases mod torsion are given: $\mathcal{B}_0$, $\mathcal{B}_1$, and $\mathcal{B}_2$. See Example 5.

were checked; this took almost 33 hours and of all relevant 8-tuples only 14277 with $T_0 = \mathcal{O}$ and 14130 with $T_0 = Q$ satisfied the required inequality. A few minutes CPU time proved that no such 8-tuple corresponds to an integral point.

It remained to search for all integral points among those corresponding to 8-tuples in the range $-4 \leq m_i \leq 4$. This took 3 hours and 40 minutes in case $T_0 = \mathcal{O}$ and produced $2 \times 21$ integral points, and 5 hours and 43 minutes were needed for the case $T_0 = Q$ which found another $2 \times 13 + 1$ integral points. The factor 2 is explained by the fact that for any point $(x, y)$ on (19), $(x, -x-y)$ is also a point different from $(x, y)$, except when $(x, y) = Q$. All integral solutions satisfy $\max_{1 \leq i \leq 8} |m_i| = 1$; for a complete listing see the first author's homepage.

## ACKNOWLEDGEMENT

This paper was subjected to a high standard of scrutiny: we received no less than three referee reports, a first for us. The referees' suggestions helped improve the paper, and we are grateful for them.

## REFERENCES

[Bašmakova 1974] I. G. Bašmakova, *Diophant und diophantische Gleichungen*, Birkhäuser, Basel, 1974. Translated from the Russian. English translation: *Diophantus and Diophantine equations*, *Dolciani Math. Expositions* **20**, MAA, Washington, 1997.

[Bremner et al. 1997] A. Bremner, R. J. Stroeker, and N. Tzanakis, "On sums of consecutive squares", *J. Number Theory* **62**:1 (1997), 39–70.

[Connell 1995] I. Connell, *Apecs* (a Maple program for arithmetic of plane elliptic curves), 1995. See ftp://ftp.math.mcgill.ca/pub/apecs/.

[Cremona 1992] J. E. Cremona, *Algorithms for modular elliptic curves*, Cambridge University Press, Cambridge, 1992. See ftp://euclid.ex.ac.uk/pub/cremona/progs/.

[David 1995] S. David, *Minorations de formes linéaires de logarithmes elliptiques*, Mém. Soc. Math. France **62**, Soc. math. France, Paris, 1995.

[Gebel, Pethő, and Zimmer 1994] J. Gebel, A. Pethő, and H. G. Zimmer, "Computing integral points on elliptic curves", *Acta Arith.* **68**:2 (1994), 171–192.

[Horn and Johnson 1985] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge, 1985. Reprinted with corrections 1990.

[Kretschmer 1986] T. J. Kretschmer, "Construction of elliptic curves with large rank", *Math. Comp.* **46**:174 (1986), 627–635.

[Merriman et al. 1996] J. R. Merriman, S. Siksek, and N. P. Smart, "Explicit 4-descents on an elliptic curve", *Acta Arith.* **77**:4 (1996), 385–404.

[Mestre 1986] J.-F. Mestre, "Formules explicites et minorations de conducteurs de variétés algébriques", *Compositio Math.* **58**:2 (1986), 209–232.

[Mestre 1991] J.-F. Mestre, "Courbes elliptiques de rang $\geq 12$ sur $\mathbb{Q}(t)$", *C. R. Acad. Sci. Paris Sér. I Math.* **313**:4 (1991), 171–174.

[Scriba 1984] C. J. Scriba, *Zur Geschichte der Bestimmung rationaler Punkte auf elliptischen Kurven: Das Problem von Behā-Eddīn 'Amūlī*, Berichte aus den Sitzungen der Joachim Jungius–Gesellschaft der Wissenschaften e.V., Hamburg **1/6**, Vandenhoeck & Ruprecht, Göttingen, 1984.

[Siegel 1929] C. L. Siegel, "Über einige Anwendungen Diophantischer Approximationen", *Abh. Preuss. Akad. Wiss. Phys. Math. Kl.* (1929), 1–41. Reprinted as pp. 209–266 of his *Gesammelte Abhandlungen I*, Springer, Berlin, 1966.

[Siksek 1995] S. Siksek, "Infinite descent on elliptic curves", *Rocky Mountain J. Math.* **25**:4 (1995), 1501–1538.

[Smart 1994] N. P. Smart, "$S$-integral points on elliptic curves", *Math. Proc. Cambridge Philos. Soc.* **116**:3 (1994), 391–399.

[Smart and Stephens 1997] N. P. Smart and N. M. Stephens, "Integral points on elliptic curves over number fields", *Math. Proc. Cambridge Philos. Soc.* **122**:1 (1997), 9–16.

[Stroeker 1995] R. J. Stroeker, "On the sum of consecutive cubes being a perfect square", *Compositio Math.* **97**:1-2 (1995), 295–307.

[Stroeker and de Weger 1999a] R. J. Stroeker and B. M. M. de Weger, "Elliptic binomial Diophantine equations", *Math. Comp.* **68** (1999), 1257–1281.

[Stroeker and de Weger 1999b] R. J. Stroeker and B. M. M. de Weger, "Solving elliptic Diophantine equations: the general cubic case", *Acta Arith.* **87**:4 (1999), 339–365.

[Stroeker and Tzanakis 1994] R. J. Stroeker and N. Tzanakis, "Solving elliptic Diophantine equations by estimating linear forms in elliptic logarithms", *Acta Arith.* **67**:2 (1994), 177–196.

[Top 1996] J. Top, "Examples of elliptic quartics with many integral points", 1996. unpublished.

[Tzanakis 1996] N. Tzanakis, "Solving elliptic Diophantine equations by estimating linear forms in elliptic logarithms: The case of quartic equations", *Acta Arith.* **75**:2 (1996), 165–190.

[de Weger 1989] B. M. M. de Weger, *Algorithms for Diophantine equations*, CWI Tract **65**, Stichting Mathematisch Centrum, Amsterdam, 1989.

[Weil 1984] A. Weil, *Number theory: An approach through history from Hammurapi to Legendre*, Birkhäuser, Boston, 1984.

Roel J. Stroeker, Econometrisch Instituut, Erasmus Universiteit, Postbus 1738, 3000 DR Rotterdam, The Netherlands (stroeker@few.eur.nl, http://www.few.eur.nl/few/people/stroeker)

Nikos Tzanakis, Department of Mathematics, University of Crete, Iraklion, Greece (tzanakis@math.uch.gr)

Received September 8, 1997; accepted in revised from June 11, 1998