# INTEGRAL POINTS ON HYPERELLIPTIC CURVES

Y. BUGEAUD, M. MIGNOTTE, S. SIKSEK, M. STOLL, SZ. TENGELY

Abstract. We give a completely explicit upper bound for integral points on (standard) affine models of hyperelliptic curves. We also explain a powerful refinement of the Mordell–Weil sieve which, combined with the upper bound, is capable of determining all the integral points. Our method is illustrated by showing that the only integral solutions to $Y^2 - Y = X^5 - X$ have $X = -1$, 0, 1, 2, 3, 30.

## 1. Introduction

Consider the hyperelliptic curve with affine model

$$C : Y^2 = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0, \tag{1}$$

with $a_0, \ldots, a_n$ rational integers, $a_n \neq 0$, $n \geq 5$, and the polynomial on the right separable. Let $H = \max\{|a_0|, \ldots, |a_n|\}$. In one of the earliest applications of his theory of lower bounds for linear forms in logarithms, Baker [1] showed that any integral point $(X, Y)$ on this affine model satisfies

$$\max(|X|, |Y|) \leq \exp\exp\exp\{(n^{10n} H)^{n^2}\}.$$

Such bounds have been improved considerably by many authors, including Sprindžuk [40], Brindza [5], Schmidt [35], Poulakis [32], Bilu [2], Bugeaud [11] and Voutier [51]. Despite the improvements, the bounds remain astronomical and often involve inexplicit constants.

In this paper we explain a new method for explicitly computing the integral points on affine models of hyperelliptic curves (1). The method falls into two distinct steps:

(i) We give a completely explicit upper bound for the size of integral solutions of (1). This upper bound combines the many refinements found in the papers of Voutier, Bugeaud, etc., together with Matveev's bounds for linear forms in logarithms [27], and a method for bounding the regulators based on a theorem of Landau [25].

(ii) The bounds obtained in (i), whilst substantially better than bounds given by earlier authors, are still astronomical. We explain a powerful variant of the Mordell–Weil sieve which, combined with the bound obtained in (i), is capable of showing that the known solutions to (1) are the only ones.

Step (i) does not demand knowledge of the class groups and unit groups of high degree number fields, merely cheaply obtainable estimates for discriminants, class numbers and regulators. Step (ii) is practical provided certain assumptions are satisfied:

*Date*: January 28, 2008.

2000 *Mathematics Subject Classification.* Primary 11G30, Secondary 11G35.

(a) Let $J$ be the Jacobian of $C$. We assume that a Mordell–Weil basis for $J(\mathbb{Q})$ is known.

(b) We assume that the canonical height $\hat{h} : J(\mathbb{Q}) \to \mathbb{R}$ is explicitly computable and that we have explicit bounds for the difference

$$\mu_1 \leq h(D) - \hat{h}(D) \leq \mu_1' \tag{2}$$

where $h$ is an appropriately normalized logarithmic height on $J$ that allows us to enumerate points $P$ in $J(\mathbb{Q})$ with $h(P) \leq B$ for a given bound $B$.

Assumptions (a) and (b) deserve a comment or two. For many families of curves of higher genus, practical descent strategies are available for estimating the rank of the Mordell–Weil group; see for example [19], [14], [31], [33], [41], [43], [44] and [34]. To provably determine the Mordell–Weil group one however needs bounds for the difference between the logarithmic and canonical heights. For Jacobians of curves of genus 2 such bounds have been determined by Stoll [42], [45], building on previous work of Flynn and Smart [21]. At present, no such bounds have been determined for Jacobians of curves of genus $\geq 3$, although work on this is in progress.

We illustrate the practicality of our approach by proving the following result.

**Theorem 1.** *The only integral solutions to the equation*

$$Y^2 - Y = X^5 - X \tag{3}$$

*are*

$$(X, Y) = (-1, 0),\ (-1, 1),\ (0, 0),\ (0, 1),\ (1, 0),\ (1, 1),\ (2, -5),$$
$$(2, 6),\ (3, -15),\ (3, 16),\ (30, -4929),\ (30, 4930).$$

The equation (3) is a special case of the family of Diophantine equations

$$Y^p - Y = X^q - X, \qquad 2 \leq p < q.$$

This family has previously been studied by Fielder and Alford [17] and by Mignotte and Pethő [28]. The (genus 1) case $p = 2$, $q = 3$ was solved by Mordell [29] who showed that the only solutions in this case are

$$(X, Y) = (0, 0), (0, 1), (\pm 1, 0), (\pm 1, 1), (2, 3), (2, -2), (6, 15), (6, -14).$$

Equation (3) is also the first problem on a list of 22 unsolved Diophantine problems [16], compiled by Evertse and Tijdeman following a recent workshop on Diophantine equations at Leiden.

To appreciate why the innocent-looking equation (3) has resisted previous attempts at solving it, let us briefly survey the available methods which apply to hyperelliptic curves and then briefly explain why they fail in this case. To determine the integral points on the affine model $C$ given by an equation (1) there are four available methods:

(I) The first is Chabauty's elegant method which in fact determines all rational points on $C$ in many cases, provided the rank of the Mordell–Weil group of its Jacobian is strictly less than the genus $g$; see for example [20], [26], [53], [46]. Chabauty's method fails if the rank of the Mordell–Weil group exceeds the genus.

(II) A second method is to use coverings, often combined with a version of Chabauty called 'Elliptic Curve Chabauty'. See [53], [22], [23], [6], [7]. This approach often requires computations of Mordell–Weil groups over

number fields (and does fail if the rank of the Mordell–Weil groups is too large).

(III) A third method is to combine Baker's approach through $S$-units with the LLL algorithm to obtain all the solutions provided that certain relevant unit groups and class groups can be computed; for a modern treatment, see [3] or [38, Section XIV.4]. This strategy often fails in practice as the number fields involved have very high degree.

(IV) A fourth approach is to apply Skolem's method to the $S$-unit equations (see [38, Section III.2]). This needs the same expensive information as the third method.

The Jacobian of (3) has rank 3 and so Chabauty's method fails. To employ Elliptic Curve Chabauty would require the computation of Mordell–Weil groups of elliptic curves without rational 2-torsion over number fields of degree 5 (which does not seem practical at present). To apply the $S$-unit approach (with either LLL or Skolem) requires the computations of the unit groups and class groups of several number fields of degree 40; a computation that seems completely impractical at present.

Our paper is arranged as follows. In Section 2 we show, after appropriate scaling, that an integral point $(x, y)$ satisfies $x - \alpha = \kappa\xi^2$ where $\alpha$ is some fixed algebraic integer, $\xi \in \mathbb{Q}(\alpha)$, and $\kappa$ is an algebraic integer belonging to a finite computable set. In Section 8 we give bounds for the size of solutions $x \in \mathbb{Z}$ to an equation of the form $x - \alpha = \kappa\xi^2$ where $\alpha$ and $\kappa$ are fixed algebraic integers. Thus, in effect, we obtain bounds for the size of solutions integral points on our affine model for (1). Sections 3–7 are preparation for Section 8: in particular Section 3 is concerned with heights; Section 4 explains how a theorem of Landau can be used to bound the regulators of number fields; Section 5 collects and refines various results on appropriate choices of systems of fundamental units; Section 6 is devoted to Matveev's bounds for linear forms in logarithms; in Section 7 we use Matveev's bounds and the results of previous sections to prove a bound on the size of solutions of unit equations; in Section 8 we deduce the bounds for $x$ alluded to above from the bounds for solutions of unit equations. Despite our best efforts, the bounds obtained for $x$ are still so large that no naive search up to those bounds is conceivable. Over the next three sections 9, 10, 11 we explain how to sieve effectively up to these bounds using the Mordell–Weil group of the Jacobian. In particular, Section 10 gives a powerful refinement of the Mordell–Weil sieve ([8], [10]) which we expect to have applications elsewhere. Finally, in Section 12 we apply the method of this paper to prove Theorem 1.

## 2. Descent

Consider the integral points on the affine model of the hyperelliptic curve (1). If the polynomial on the right-hand side is reducible then the obvious factorisation argument reduces the problem of determining the integral points on (1) to determining those on simpler hyperelliptic curves, or on genus 1 curves. The integral points on a genus 1 curve can be determined by highly successful algorithms [24], [37], [39], [47], [48], [49],[50], based on LLL and David's bound for linear forms in elliptic logarithms [15].

We therefore suppose henceforth that the polynomial on the right-hand side of (1) is irreducible; this is certainly the most difficult case. By appropriate scaling,

one transforms the problem of integral points on (1) to integral points on a model of the form

$$ay^2 = x^n + b_{n-1}x^{n-1} + \cdots + b_0, \tag{4}$$

where $a$ and the $b_i$ are integers, with $a \neq 0$. We shall work henceforth with this model of the hyperelliptic curve. Denote the polynomial on the right-hand side by $f$ and let $\alpha$ be a root of $f$. Then a standard argument shows that

$$x - \alpha = \kappa\xi^2$$

where $\kappa$, $\xi \in K = \mathbb{Q}(\alpha)$ and $\kappa$ is an **algebraic integer that comes from a finite computable set**. In this section we suppose that the Mordell–Weil group $J(\mathbb{Q})$ of the curve $C$ is known, and we show how to compute such a set of $\kappa$ using our knowledge of the Mordell–Weil group $J(\mathbb{Q})$. The method for doing this depends on whether the degree $n$ is odd or even.

2.1. **The Odd Degree Case.** Each coset of $J(\mathbb{Q})/2J(\mathbb{Q})$ has a coset representative of the form $\sum_{i=1}^{m}(P_i - \infty)$ where the set $\{P_1, \ldots, P_m\}$ is fixed under the action of Galois, and where all $y(P_i)$ are non-zero. Now write $x(P_i) = \gamma_i/d_i^2$ where $\gamma_i$ is an algebraic integer and $d_i \in \mathbb{Z}_{\geq 1}$; moreover if $P_i$, $P_j$ are conjugate then we may suppose that $d_i = d_j$ and so $\gamma_i$, $\gamma_j$ are conjugate. To such a coset representative of $J(\mathbb{Q})/2J(\mathbb{Q})$ we associate

$$\kappa = a^{(m \bmod 2)} \prod_{i=1}^{m} \left(\gamma_i - \alpha d_i^2\right).$$

**Lemma 2.1.** *Let $\mathcal{K}$ be a set of $\kappa$ associated as above to a complete set of coset representatives of $J(\mathbb{Q})/2J(\mathbb{Q})$. Then $\mathcal{K} \subset \mathcal{O}_K$ and if $(x, y)$ is an integral point on the model* (4) *then $x - \alpha = \kappa\xi^2$ for some $\kappa \in \mathcal{K}$ and $\xi \in K$.*

*Proof.* This follows trivially from the standard homomorphism

$$\theta : J(\mathbb{Q})/2J(\mathbb{Q}) \to K^*/K^{*2}$$

that is given by

$$\theta\left(\sum_{i=1}^{m}(P_i - \infty)\right) = a^m \prod_{i=1}^{m}(x(P_i) - \alpha) \pmod{K^{*2}}$$

for coset representatives $\sum(P_i - \infty)$ with $y(P_i) \neq 0$; see Section 4 of [43]. □

2.2. **The Even Degree Case.** In this case, the homomorphism $\theta$ takes values in $K^*/\mathbb{Q}^*K^{*2}$ and therefore does not provide sufficient information. However, we know that the relevant $\kappa$ must have even valuation at all prime ideals not dividing $a$ (the coefficient of $y^2$ in (4)) or the discriminant of the polynomial on the right hand side of (4). Modulo squares, this is a finite set, which can be computed explicitly, given the same kind of class and unit group information that is necessary to compute the 2-Selmer rank of $J(\mathbb{Q})$. We can then use local information at small and bad primes to restrict this set further, compare [8] and [9], where this is applied to rational points. In our case, we can restrict the local computations to $x \in \mathbb{Z}_p$ instead of $\mathbb{Q}_p$.

## 3. Heights

We fix once and for all the following notation.

| | |
|---|---|
| $K$ | a number field, |
| $\mathcal{O}_K$ | the ring of integers of $K$, |
| $M_K$ | the set of all places of $K$, |
| $M_K^0$ | the set of non-Archimedean places of $K$, |
| $M_K^\infty$ | the set of Archimedean places of $K$, |
| $v$ | a place of $K$, |
| $K_v$ | the completion of $K$ at $v$, |
| $d_v$ | the local degree $[K_v : \mathbb{Q}_v]$. |

For $v \in M_K$, we let $|\cdot|_v$ be the usual normalized valuation corresponding to $v$; in particular if $v$ is non-Archimedean and $p$ is the rational prime below $v$ then $|p|_v = p^{-1}$. Thus if $L/K$ is a field extension, and $\omega$ a place of $L$ above $v$ then $|\alpha|_\omega = |\alpha|_v$, for all $\alpha \in K$.

Define

$$\|\alpha\|_v = |\alpha|_v^{d_v}.$$

Hence for $\alpha \in K^*$, the product formula states that

$$\prod_{v \in M_K} \|\alpha\|_v = 1.$$

In particular, if $v$ is Archimedean, corresponding to a real or complex embedding $\sigma$ of $K$ then

$$|\alpha|_v = |\sigma(\alpha)| \qquad \text{and} \qquad \|\alpha\|_v = \begin{cases} |\sigma(\alpha)| & \text{if } \sigma \text{ is real} \\ |\sigma(\alpha)|^2 & \text{if } \sigma \text{ is complex.} \end{cases}$$

For $\alpha \in K$, the (absolute) logarithmic height $\mathrm{h}(\alpha)$ is given by

$$\mathrm{h}(\alpha) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} d_v \log \max\{1, |\alpha|_v\} = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} \log \max\{1, \|\alpha\|_v\}. \tag{5}$$

The absolute logarithmic height of $\alpha$ is indepedent of the field $K$ containing $\alpha$.

We shall need the following elementary properties of heights.

**Lemma 3.1.** *For any non-zero algebraic number $\alpha$, we have $\mathrm{h}(\alpha^{-1}) = \mathrm{h}(\alpha)$. For algebraic numbers $\alpha_1, \ldots, \alpha_n$, we have*

$$\mathrm{h}(\alpha_1\alpha_2\cdots\alpha_n) \le \mathrm{h}(\alpha_1)+\cdots+\mathrm{h}(\alpha_n), \qquad \mathrm{h}(\alpha_1+\cdots+\alpha_n) \le \log n+\mathrm{h}(\alpha_1)+\cdots+\mathrm{h}(\alpha_n).$$

*Proof.* The lemma is Exercise 8.8 in [36]. We do not know of a reference for the proof and so we will indicate briefly the proof of the second (more difficult) inequality. For $v \in M_K$, choose $i_v$ in $\{1, \ldots, n\}$ to satisfy $\max\{|\alpha_1|_v, \ldots, |\alpha_n|_v\} = |\alpha_{i_v}|_v$. Note that

$$|\alpha_1 + \cdots + \alpha_n|_v \le \epsilon_v |\alpha_{i_v}|_v, \qquad \text{where} \qquad \epsilon_v = \begin{cases} n & \text{if } v \text{ is Archimedean,} \\ 1 & \text{otherwise.} \end{cases}$$

Thus

$$\log\max\{1, |\alpha_1+\cdots+\alpha_n|_v\} \le \log\epsilon_v+\log\max\{1, |\alpha_{i_v}|_v\} \le \log\epsilon_v+\sum_{i=1}^{n} \log\max\{1, |\alpha_i|_v\}.$$

Observe that

$$\frac{1}{[K:\mathbb{Q}]}\sum_{v\in M_K} d_v \log \epsilon_v = \frac{\log n}{[K:\mathbb{Q}]}\sum_{v\in M_K^\infty} d_v = \log n;$$

the desired inequality follows from the definition of logarithmic height (5). □

3.1. **Height Lower Bound.** We need the following result of Voutier [52] concerning Lehmer's problem.

**Lemma 3.2.** *Let $K$ be a number field of degree $d$. Let*

$$\partial_K = \begin{cases} \frac{\log 2}{d} & \text{if } d = 1,\ 2, \\ \frac{1}{4}\left(\frac{\log\log d}{\log d}\right)^3 & \text{if } d \geq 3. \end{cases}$$

*Then, for every non-zero algebraic number $\alpha$ in $K$, which is not a root of unity,*

$$\deg(\alpha)\ \mathrm{h}(\alpha) \geq \partial_K.$$

Throughout, by the logarithm of a complex number, we mean the principal determination of the logarithm. In other words, if $x \in \mathbb{C}^*$ we express $x = re^{i\theta}$ where $r > 0$ and $-\pi < \theta \leq \pi$; we then let $\log x = \log r + i\theta$.

**Lemma 3.3.** *Let $K$ be a number field and let*

$$\partial'_K = \left(1 + \frac{\pi^2}{\partial_K^2}\right)^{1/2}.$$

*For any non-zero $\alpha$ and any place $v \in M_K$*

$$\log|\alpha|_v \leq \deg(\alpha)\,\mathrm{h}(\alpha), \qquad \log\|\alpha\|_v \leq [K:\mathbb{Q}]\,\mathrm{h}(\alpha).$$

*Moreover, if $\alpha$ is not a root of unity and $\sigma$ is a real or complex embedding of $K$ then*

$$|\log\sigma(\alpha)| \leq \partial'_K \deg(\alpha)\,\mathrm{h}(\alpha).$$

*Proof.* The first two inequalities are an immediate consequence of the definition of absolute logarithmic height. For the last, write $\sigma(\alpha) = e^{a+ib}$, with $a = \log|\sigma(\alpha)|$ and $|b| \leq \pi$, and let $d = \deg(\alpha)$. Then we have

$$|\log\sigma(\alpha)| = (a^2+b^2)^{1/2} \leq (\log^2|\sigma(\alpha)| + \pi^2)^{1/2} \leq ((d\,\mathrm{h}(\alpha))^2 + \pi^2)^{1/2}.$$

By Lemma 3.2 we have $d\,\mathrm{h}(\alpha) \geq \partial_K$, so

$$|\log\sigma(\alpha)| \leq d\ \mathrm{h}(\alpha)\ \left(1 + \frac{\pi^2}{\partial_K^2}\right)^{1/2},$$

as required. □

## 4. Bounds for Regulators

Later on we need give upper bounds for the regulators of complicated number fields of high degree. The following lemma, based on bounds of Landau [25], is an easy way to obtain reasonable bounds.

**Lemma 4.1.** *Let $K$ be a number field with degree $d = u + 2v$ where $u$ and $v$ are respectively the numbers of real and complex embeddings. Denote the absolute discriminant by $D_K$ and the regulator by $R_K$, and the number of roots of unity in $K$ by $w$. Suppose, moreover, that $L$ is a real number such that $D_K \leq L$. Let*

$$a = 2^{-v}\,\pi^{-d/2}\,\sqrt{L}.$$

*Define the function $f_K(L,s)$ by*

$$f_K(L,s) = 2^{-u}\,w\,a^s\left(\Gamma(s/2)\right)^u\left(\Gamma(s)\right)^v s^{d+1}\,(s-1)^{1-d},$$

*and let $B_K(L) = \min\{f_K(L, 2 - t/1000) \ : t = 0, 1, \ldots, 999\}$. Then $R_K < B_K(L)$.*

*Proof.* Landau [25] proved the inequality $R_K < f_K(D_K, s)$ for all $s > 1$. It is thus clear that $R_K < B_K(L)$. □

Perhaps a comment is in order. For a complicated number field of high degree it is difficult to calculate the discriminant $D_K$ exactly, though it is easy to give an upper bound $L$ for its size. It is also difficult to minimise the function $f_K(L,s)$ analytically, but we have found that the above gives an accurate enough result, which is easy to calculate on a computer.

## 5. Fundamental Units

For the number fields we are concerned with, we shall need to work with a certain system of fundamental units, given by the following lemma due to Bugeaud and Győry, which is Lemma 1 of [12].

**Lemma 5.1.** *Let $K$ be a number field of degree $d$ and let $r = r_K$ be its unit rank and $R_K$ its regulator. Define the constants*

$$c_1 = c_1(K) = \frac{(r\,!)^2}{2^{r-1}d^r}\,, \qquad c_2 = c_2(K) = c_1\left(\frac{d}{\partial_K}\right)^{r-1}, \qquad c_3 = c_3(K) = c_1\frac{d^r}{\partial_K}\,.$$

*Then $K$ admits a system $\{\varepsilon_1, \ldots, \varepsilon_r\}$ of fundamental units such that:*

(i) $$\prod_{i=1}^{r} \mathrm{h}(\varepsilon_i) \leq c_1 R_K,$$

(ii) $$\mathrm{h}(\varepsilon_i) \leq c_2 R_K, \quad 1 \leq i \leq r,$$

(iii) *Write $\mathcal{M}$ for the $r \times r$-matrix $(\log\|\varepsilon_i\|_v)$ where $v$ runs over $r$ of the Archimedean places of $K$ and $1 \leq i \leq r$. Then the absolute values of the entries of $\mathcal{M}^{-1}$ are bounded above by $c_3$.*

**Lemma 5.2.** *Let $K$ be a number field of degree $d$, and let $\{\varepsilon_1, \ldots, \varepsilon_r\}$ be a system of fundamental units as in Lemma 5.1. Define the constant $c_4 = c_4(K) = rdc_3$. Suppose $\varepsilon = \zeta\varepsilon_1^{b_1}\ldots\varepsilon_r^{b_r}$, where $\zeta$ is a root of unity in $K$. Then*

$$\max\{|b_1|, \ldots, |b_r|\} \leq c_4\,\mathrm{h}(\varepsilon).$$

*Proof.* Note that for any Archimedean place $v$ of $K$,

$$\log\|\varepsilon\|_v = \sum b_i \log\|\varepsilon_i\|_v.$$

The lemma now follows from part (iii) of Lemma 5.1, plus the fact that $\log\|\varepsilon\|_v \leq d\,\mathrm{h}(\varepsilon)$ for all $v$ given by Lemma 3.3. □

The following result is a special case of Lemma 2 of [12].

**Lemma 5.3.** *Let $K$ be a number field of unit rank $r$ and regulator $K$. Let $\alpha$ be a non-zero algebraic integer belonging to $K$. Then there exists a unit $\varepsilon$ of $K$ such that*

$$\mathrm{h}(\alpha\varepsilon) \leq c_5 R_K + \frac{\log|\mathrm{Norm}_{K/\mathbb{Q}}(\alpha)|}{[K:\mathbb{Q}]}$$

*where*

$$c_5 = c_5(K) = \frac{r^{r+1}}{2\partial_K'^{r-1}}.$$

**Lemma 5.4.** *Let $K$ be a number field, $\beta$, $\varepsilon \in K^*$ with $\varepsilon$ being a unit. Let $\sigma$ be the real or complex embedding that makes $|\sigma(\beta\varepsilon)|$ minimal. Then*

$$\mathrm{h}(\beta\varepsilon) \leq \mathrm{h}(\beta) - \log|\sigma(\beta\varepsilon)|.$$

*Proof.* As usual, write $d = [K:\mathbb{Q}]$ and $d_v = [K_v:\mathbb{Q}_v]$. Note

$$\begin{aligned}
\mathrm{h}(\beta\varepsilon) &= \mathrm{h}(1/\beta\varepsilon) \\
&= \frac{1}{d}\sum_{v\in M_K^\infty} d_v \max\{0, \log(|\beta\varepsilon|_v^{-1})\} + \frac{1}{d}\sum_{v\in M_K^0} d_v \max\{0, \log(|\beta\varepsilon|_v^{-1})\} \\
&\leq \log(|\sigma(\beta\varepsilon)|^{-1}) + \frac{1}{d}\sum_{v\in M_K^0} d_v \max\{0, \log(|\beta|_v^{-1})\} \\
&\leq -\log|\sigma(\beta\varepsilon)| + \frac{1}{d}\sum_{v\in M_K} d_v \max\{0, \log(|\beta|_v^{-1})\} \\
&\leq -\log|\sigma(\beta\varepsilon)| + \mathrm{h}(\beta),
\end{aligned}$$

as required. □

## 6. Matveev's Lower Bound for Linear Forms in Logarithms

Let $L$ be a number field and let $\sigma$ be a real or complex embedding. For $\alpha \in L^*$ we define the *modified logarithmic height of $\alpha$ with respect to $\sigma$* to be

$$\mathrm{h}_{L,\sigma}(\alpha) := \max\{[L:\mathbb{Q}]\,\mathrm{h}(\alpha)\ ,\ |\log\sigma(\alpha)|\ ,\ 0.16\}.$$

The modified height is clearly dependent on the number field; we shall need the following Lemma which gives a relation between the modified and absolute height.

**Lemma 6.1.** *Let $K \subseteq L$ be number fields and write*

$$\partial_{L/K} = \max\left\{[L:\mathbb{Q}]\ ,\ [K:\mathbb{Q}]\partial_K'\ ,\ \frac{0.16[K:\mathbb{Q}]}{\partial_K}\right\}.$$

*Then for any $\alpha \in K$ which is neither zero nor a root of unity, and any real or complex embedding $\sigma$ of $L$,*

$$\mathrm{h}_{L,\sigma}(\alpha) \leq \partial_{L/K}\,\mathrm{h}(\alpha).$$

*Proof.* By Lemma 3.3 we have

$$[K:\mathbb{Q}]\partial'_K \,\mathrm{h}(\alpha) \geq \partial'_K \deg(\alpha)\,\mathrm{h}(\alpha) \geq |\log\sigma(\alpha)|.$$

Moreover, by Lemma 3.2,

$$\frac{0.16[K:\mathbb{Q}]\,\mathrm{h}(\alpha)}{\partial_K} \geq \frac{0.16\deg(\alpha)\,\mathrm{h}(\alpha)}{\partial_K} \geq 0.16.$$

The lemma follows. □

We shall apply lower bounds on linear forms, more precisely a version of Matveev's estimates [27]. We recall that log denotes the principal determination of the logarithm.

**Lemma 6.2.** *Let $L$ be a number field of degree $d$, with $\alpha_1,\ldots,\alpha_n \in L^*$. Define a constant*

$$C(L,n) := 3\cdot 30^{n+4}\cdot(n+1)^{5.5}\,d^2\,(1+\log d).$$

*Consider the "linear form"*

$$\Lambda := \alpha_1^{b_1}\cdots\alpha_n^{b_n} - 1,$$

*where $b_1,\ldots,b_n$ are rational integers and let $B := \max\{|b_1|,\ldots,|b_n|\}$. If $\Lambda \neq 0$, and $\sigma$ is any real or complex embedding of $L$ then*

$$\log|\sigma(\Lambda)| > -C(L,n)(1+\log(nB))\prod_{j=1}^{n}\mathrm{h}_{L,\sigma}(\alpha_j).$$

*Proof.* This straightforward corollary of Matveev's estimates is Theorem 9.4 of [13]. □

## 7. Bounds for Unit Equations

Now we are ready to prove an explicit version of Lemma 4 of [11]. The proposition below allows us to replace in the final estimate the regulator of the larger field by the product of the regulators of two of its subfields. This often results in a significant improvement of the upper bound for the height. This idea is due to Voutier [51].

**Proposition 7.1.** *Let $K$ be a number field of degree $d$, which contains $K_1$ and $K_2$ as subfields. Let $R_{K_i}$ (respectively $r_i$) be the regulator (respectively the unit rank) of $K_i$. Suppose further that $\nu_1$, $\nu_2$ and $\nu_3$ are non-zero elements of $L$ with height $\leq H$ (with $H \geq 1$) and consider the unit equation*

$$\nu_1\varepsilon_1 + \nu_2\varepsilon_2 + \nu_3\varepsilon_3 = 0 \tag{6}$$

*where $\varepsilon_1$ is a unit of $K_1$, $\varepsilon_2$ a unit of $K_2$ and $\varepsilon_3$ a unit of $L$. Then, for $i = 1$ and $2$,*

$$\mathrm{h}(\nu_i\varepsilon_i/\nu_3\varepsilon_3) \leq A_2 + A_1\log\{H + \max\{\mathrm{h}(\nu_1\varepsilon_1),\mathrm{h}(\nu_2\varepsilon_2)\}\},$$

*where*

$$A_1 = 2H\cdot C(L,r_1+r_2+1)\cdot c_1(K_1)c_1(K_2)\partial_{L/L}\cdot(\partial_{L/K_1})^{r_1}\cdot(\partial_{L/K_2})^{r_2}\cdot R_{K_1}R_{K_2},$$

*and*

$$A_2 = 2H + A_1 + A_1\log\{(r_1+r_2+1)\cdot\max\{c_4(K_1),c_4(K_2),1\}\}.$$

*Proof.* Let $\{\mu_1, \dots, \mu_{r_1}\}$ and $\{\rho_1, \dots, \rho_{r_2}\}$ be respectively systems of fundamental units for $K_1$ and $K_2$ as in Lemma 5.1. Then we can write

$$\varepsilon_1 = \zeta_1 \mu_1^{b_1} \cdots \mu_{r_1}^{b_{r_1}}, \quad \varepsilon_2 = \zeta_2 \rho_1^{f_1} \cdots \rho_{r_2}^{f_{r_2}},$$

where $\zeta_1$ and $\zeta_2$ are roots of unity and $b_1, \dots, b_{r_1}$, and $f_1, \dots, f_{r_2}$ are rational integers. Set

$$B_1 = \max\{|b_1|, \dots, |b_{r_1}|\}, \quad B_2 = \max\{|f_1|, \dots, |f_{r_2}|\}, \quad B = \max\{B_1, B_2, 1\}.$$

Set $\alpha_0 = -\zeta_2\nu_2/(\zeta_1\nu_1)$ and $b_0 = 1$. By (6),

$$\frac{\nu_3\varepsilon_3}{\nu_1\varepsilon_1} = \alpha_0^{b_0} \mu_1^{-b_1} \cdots \mu_{r_1}^{-b_{r_1}} \rho_1^{f_1} \cdots \rho_{r_2}^{f_{r_2}} - 1.$$

Now choose the real or complex embedding $\sigma$ of $L$ such that $|\sigma((\nu_3\varepsilon_3)/(\nu_1\varepsilon_1))|$ is minimal. We apply Matveev's estimate (Lemma 6.2) to this "linear form", obtaining

$$\log \left| \sigma\left(\frac{\nu_3\varepsilon_3}{\nu_1\varepsilon_1}\right) \right| > -C(L,n)(1+\log(nB))\, \mathrm{h}_{L,\sigma}(\alpha_0) \prod_{j=1}^{r_1} \mathrm{h}_{L,\sigma}(\mu_j) \prod_{j=1}^{r_2} \mathrm{h}_{L,\sigma}(\rho_j),$$

where $n = r_1 + r_2 + 1$. Using Lemma 6.1 and Lemma 5.1 we obtain

$$\prod_{j=1}^{r_1} \mathrm{h}_{L,\sigma}(\mu_j) \le (\partial_{L/K_1})^{r_1} \prod_{j=1}^{r_1} \mathrm{h}(\mu_j) \le c_1(K_1)(\partial_{L/K_1})^{r_1} R_{K_1},$$

and a similar estimate for $\prod_{j=1}^{r_2} \mathrm{h}_{L,\sigma}(\rho_j)$. Moreover, again by Lemma 6.1 and Lemma 3.1, $\mathrm{h}_{L,\sigma}(\alpha_0) \le 2H\partial_{L/L}$. Thus

$$\log \left| \sigma\left(\frac{\nu_3\varepsilon_3}{\nu_1\varepsilon_1}\right) \right| > -A_1(1+\log(nB)).$$

Now applying Lemma 5.4, we obtain that

$$\mathrm{h}\left(\frac{\nu_3\varepsilon_3}{\nu_1\varepsilon_1}\right) \le \mathrm{h}\left(\frac{\nu_3}{\nu_1}\right) + A_1(1+\log(nB)) \le 2H + A_1(1+\log(nB)).$$

The proof is complete on observing, from Lemma 5.2, that

$$B \le \max\{c_4(K_1), c_4(K_2), 1)\} \max\{\mathrm{h}(\varepsilon_1), \mathrm{h}(\varepsilon_2), 1\},$$

and from Lemma 3.1, $\mathrm{h}(\nu_i\varepsilon_i) \le \mathrm{h}(\varepsilon_i) + \mathrm{h}(\nu_i) \le \mathrm{h}(\varepsilon) + H$. □

## 8. Upper Bounds for the Size of Integral Points on Hyperelliptic Curves

We shall need the following standard sort of lemma.

**Lemma 8.1.** *Let $a$, $b$, $c$, $y$ be positive numbers and suppose that*

$$y \le a + b\log(c+y).$$

*Then*

$$y \le 2b\log b + 2a + c.$$

*Proof.* Let $z = c + y$, so that $z \le (a+c) + b\log z$. Now we apply case $h = 1$ of Lemma 2.2 of [30]; this gives $z \le 2(b\log b + a + c)$, and the lemma follows. □

**Theorem 2.** *Let $\alpha$ be an algebraic integer of degree at least 3, and let $\kappa$ be a integer belonging to $K$. Let $\alpha_1$, $\alpha_2$, $\alpha_3$ be distinct conjugates of $\alpha$ and $\kappa_1$, $\kappa_2$, $\kappa_3$ be the corresponding conjugates of $\kappa$. Let*

$$K_1 = \mathbb{Q}(\alpha_1, \alpha_2, \sqrt{\kappa_1\kappa_2}), \quad K_2 = \mathbb{Q}(\alpha_1, \alpha_3, \sqrt{\kappa_1\kappa_3}), \quad K_3 = \mathbb{Q}(\alpha_2, \alpha_3, \sqrt{\kappa_2\kappa_3}),$$

*and*

$$L = \mathbb{Q}(\alpha_1, \alpha_2, \alpha_3, \sqrt{\kappa_1\kappa_2}, \sqrt{\kappa_1\kappa_3}).$$

*Let $R$ be an upper bound for the regulators of $K_1$, $K_2$ and $K_3$. Let $r$ be the maximum of the unit ranks of $K_1$, $K_2$, $K_3$. Let*

$$c_j^* = \max_{1\le i\le 3} c_j(K_i).$$

*Let*

$$N = \max_{1\le i,j\le 3} \left|\mathrm{Norm}_{\mathbb{Q}(\alpha_i,\alpha_j)/\mathbb{Q}}(\kappa_i(\alpha_i-\alpha_j))\right|^2.$$

*Let*

$$H^* = c_5^* R + \frac{\log N}{\min_{1\le i\le 3}[K_i:\mathbb{Q}]} + \mathrm{h}(\kappa).$$

*Let*

$$A_1^* = 2H^* \cdot C(L, 2r+1) \cdot (c_1^*)^2 \partial_{L/L} \cdot \left(\max_{1\le i\le 3} \partial_{L/K_1}\right)^{2r} \cdot R^2,$$

*and*

$$A_2^* = 2H^* + A_1^* + A_1^* \log\{(2r+1)\cdot \max\{c_4^*, 1\}\}.$$

*If $x \in \mathbb{Z}\backslash\{0\}$ satisfies $x - \alpha = \kappa\xi^2$ for some $\xi \in K$ then*

$$\log|x| \le 8A_1^* \log(4A_1^*) + 8A_2^* + H^* + 20\log 2 + 13\,\mathrm{h}(\kappa) + 19\,\mathrm{h}(\alpha).$$

*Proof.* Conjugating the relation $x - \alpha = \kappa\xi^2$ appropriately and taking differences we obtain

$$\alpha_1 - \alpha_2 = \kappa_2\xi_2^2 - \kappa_1\xi_1^2, \quad \alpha_3 - \alpha_1 = \kappa_1\xi_1^2 - \kappa_3\xi_3^2, \quad \alpha_2 - \alpha_3 = \kappa_3\xi_3^2 - \kappa_2\xi_2^2.$$

Let

$$\tau_1 = \kappa_1\xi_1, \qquad \tau_2 = \sqrt{\kappa_1\kappa_2}\xi_2, \qquad \tau_3 = \sqrt{\kappa_1\kappa_3}\xi_3.$$

Observe that

$$\kappa_1(\alpha_1-\alpha_2) = \tau_2^2 - \tau_1^2, \quad \kappa_1(\alpha_3-\alpha_1) = \tau_1^2 - \tau_3^2, \quad \kappa_1(\alpha_2-\alpha_3) = \tau_2^2 - \tau_3^2,$$

and

$$\tau_2 \pm \tau_1 \in K_1, \quad \tau_1 \pm \tau_3 \in K_2, \quad \tau_3 \pm \tau_2 \in \sqrt{\kappa_1/\kappa_2}K_3.$$

We claim that each $\tau_i \pm \tau_j$ can be written in the form $\nu\varepsilon$ where $\varepsilon$ is a unit in one of the $K_i$ and $\nu \in L$ is an integer satisfying $\mathrm{h}(\nu) \le H^*$. Let us show this for $\tau_2 - \tau_3$; the other cases are either similar or easier. Note that $\tau_2 - \tau_3 = \sqrt{\kappa_1/\kappa_2}\nu''$ where $\nu''$ is an integer belonging to $K_3$. Moreover, $\nu''$ divides

$$\sqrt{\frac{\kappa_2}{\kappa_1}}(\tau_3 - \tau_2)\cdot\sqrt{\frac{\kappa_2}{\kappa_1}}(\tau_3+\tau_2) = \kappa_2(\alpha_2-\alpha_3).$$

Hence $|\mathrm{Norm}_{K_3/\mathbb{Q}}(\nu'')| \le N$. By Lemma 5.3, we can write $\nu'' = \nu'\varepsilon$ where $\varepsilon \in K_3$ and

$$\mathrm{h}(\nu') \le c_5(K_3)R + \frac{\log N}{[K_3:\mathbb{Q}]}.$$

Now let $\nu = \sqrt{\kappa_1/\kappa_2}\nu'$. Thus $\tau_2 - \tau_3 = \nu\varepsilon$ where $\mathrm{h}(\nu) \le \mathrm{h}(\nu') + \mathrm{h}(\kappa) \le H^*$ proving our claim.

We apply Proposition 7.1 to the unit equation

$$(\tau_1 - \tau_2) + (\tau_3 - \tau_1) + (\tau_2 - \tau_3) = 0,$$

which is indeed of the form $\nu_1\varepsilon_1 + \nu_2\varepsilon_2 + \nu_3\varepsilon_3 = 0$ where the $\nu_i$ and $\varepsilon_i$ satisfy the conditions of that proposition with $H$ replaced by $H^*$. We obtain

$$\mathrm{h}\left(\frac{\tau_1 - \tau_2}{\tau_1 - \tau_3}\right) \le A_2^* + A_1^* \log\{H^* + \max\{\mathrm{h}(\tau_2 - \tau_3), \mathrm{h}(\tau_2 + \tau_3)\}\}.$$

Observe that

$$\begin{aligned}\mathrm{h}(\tau_2 \pm \tau_3) &\le \log 2 + \mathrm{h}(\tau_2) + \mathrm{h}(\tau_3)\\ &\le \log 2 + 2\,\mathrm{h}(\kappa) + 2\,\mathrm{h}(\xi)\\ &\le \log 2 + 3\,\mathrm{h}(\kappa) + \mathrm{h}(x - \alpha)\\ &\le 2\log 2 + 3\,\mathrm{h}(\kappa) + \mathrm{h}(\alpha) + \log|x|.\end{aligned}$$

Thus

$$\mathrm{h}\left(\frac{\tau_1 - \tau_2}{\tau_1 - \tau_3}\right) \le A_2^* + A_1^* \log(A_3^* + \log|x|),$$

where $A_3^* = H^* + 2\log 2 + 3\,\mathrm{h}(\kappa) + \mathrm{h}(\alpha)$.

We also apply Propostion 7.1 to the unit equation

$$(\tau_1 + \tau_2) + (\tau_3 - \tau_1) - (\tau_2 + \tau_3) = 0,$$

to obtain precisely the same bound for $\mathrm{h}\left(\frac{\tau_1+\tau_2}{\tau_1-\tau_3}\right)$. Using the identity

$$\left(\frac{\tau_1 - \tau_2}{\tau_1 - \tau_3}\right) \cdot \left(\frac{\tau_1 + \tau_2}{\tau_1 - \tau_3}\right) = \frac{\kappa_1(\alpha_2 - \alpha_1)}{(\tau_1 - \tau_3)^2},$$

we obtain that

$$\mathrm{h}(\tau_1 - \tau_3) \le \frac{\log 2 + \mathrm{h}(\kappa)}{2} + \mathrm{h}(\alpha) + A_2^* + A_1^* \log(A_3^* + \log|x|).$$

Now

$$\begin{aligned}\log|x| &\le \log 2 + \mathrm{h}(\alpha) + \mathrm{h}(x - \alpha_1)\\ &\le \log 2 + \mathrm{h}(\alpha) + \mathrm{h}(\kappa) + 2\,\mathrm{h}(\tau_1) \qquad (\text{using } x - \alpha_1 = \tau_1^2/\kappa_1)\\ &\le 5\log 2 + \mathrm{h}(\alpha) + \mathrm{h}(\kappa) + 2\,\mathrm{h}(\tau_1 + \tau_3) + 2\,\mathrm{h}(\tau_1 - \tau_3)\\ &\le 5\log 2 + \mathrm{h}(\alpha) + \mathrm{h}(\kappa) + 2\,\mathrm{h}\left(\frac{\kappa_1(\alpha_3 - \alpha_1)}{\tau_1 - \tau_3}\right) + 2\,\mathrm{h}(\tau_1 - \tau_3)\\ &\le 7\log 2 + 5\,\mathrm{h}(\alpha) + 3\,\mathrm{h}(\kappa) + 4\,\mathrm{h}(\tau_1 - \tau_3)\\ &\le 9\log 2 + 9\,\mathrm{h}(\alpha) + 5\,\mathrm{h}(\kappa) + 4A_2^* + 4A_1^* \log(A_3^* + \log|x|).\end{aligned}$$

The theorem follows from Lemma 8.1. □

## 9. The Mordell–Weil Sieve I

In this section we let $C/\mathbb{Q}$ be a smooth projective curve (not necessarily hyperelliptic) of genus $g \ge 2$ and we let $J$ be its Jacobian. Let $D$ be a fixed divisor on $C$ of degree 1 and let $\jmath$ be the corresponding Abel–Jacobi map:

$$\jmath : C \to J, \qquad P \mapsto [P - D].$$

Let $W$ be the image in $J$ of the known rational points on $C$. The Mordell–Weil sieve is a strategy for obtaining a very large and smooth integer $B$ such that

$$\jmath(C(\mathbb{Q})) \subseteq W + BJ(\mathbb{Q}).$$

Let $S$ be a finite set of primes, which for now we assume to be primes of good reduction for the curve $C$. The basic idea is to consider the following commutative diagram.

$$\begin{array}{ccc} C(\mathbb{Q}) & \xrightarrow{\jmath} & J(\mathbb{Q})/BJ(\mathbb{Q}) \\ \downarrow & & \downarrow \alpha \\ \displaystyle\prod_{p\in S} C(\mathbb{F}_p) & \xrightarrow{\jmath} & \displaystyle\prod_{p\in S} J(\mathbb{F}_p)/BJ(\mathbb{F}_p) \end{array}$$

The image of $C(\mathbb{Q})$ in $J(\mathbb{Q})/BJ(\mathbb{Q})$ must then be contained in the subset of $J(\mathbb{Q})/BJ(\mathbb{Q})$ of elements that map under $\alpha$ into the image of the lower horizontal map. If we find that this subset equals the image of $W$ in $J(\mathbb{Q})/BJ(\mathbb{Q})$, then we have shown that

$$\jmath(C(\mathbb{Q})) \subseteq W + BJ(\mathbb{Q})$$

as desired. Note that, at least in principle, the required computation is finite: each set $C(\mathbb{F}_p)$ is finite and can be enumerated, hence $\jmath(C(\mathbb{F}_p))$ can be determined, and we assume that we know explicit generators of $J(\mathbb{Q})$, which allows us to construct the finite set $J(\mathbb{Q})/BJ(\mathbb{Q})$. In practice, and in particular for the application we have in mind here, we will need a very large value of $B$, so this naive approach is much too inefficient. In [8] and [10], the authors describe how one can perform this computation in a more efficient way.

One obvious improvement is to replace the lower horizontal map in the diagram above by a product of maps

$$C(\mathbb{Q}_p) \xrightarrow{\jmath} G_p/BG_p$$

with suitable finite quotients $G_p$ of $J(\mathbb{Q}_p)$. We have used this to incorporate information modulo higher powers of $p$ for small primes $p$. This kind of information is often called "deep" information, as opposed to the "flat" information obtained from reduction modulo good primes.

We can always force $B$ to be divisible by any given (not too big) number. In our application we will want $B$ to kill the rational torsion subgroup of $J$.

## 10. The Mordell–Weil Sieve II

We continue with the notation of Section 9. Let $W$ be the image in $J(\mathbb{Q})$ of all the known rational points on $C$. We assume that the strategy of Section 9 is successful in yielding a large 'smooth' integer $B$ such that any point $P \in C(\mathbb{Q})$ satisfies $\jmath(P) - w \in BJ(\mathbb{Q})$ for some $w \in W$, and moreover, that $B$ kills all the torsion of $J(\mathbb{Q})$.

Let

$$\phi : \mathbb{Z}^r \to J(\mathbb{Q}), \qquad \phi(a_1, \ldots, a_r) = \sum a_i D_i,$$

so that the image of $\phi$ is simply the free part of $J(\mathbb{Q})$. Our assumption is now that

$$\jmath(C(\mathbb{Q})) \subset W + \phi(B\mathbb{Z}^n).$$

Set $L_0 = B\mathbb{Z}^n$. We explain a method of obtaining a (very long) decreasing sequence of lattices in $\mathbb{Z}^n$:

$$B\mathbb{Z}^n = L_0 \supsetneq L_1 \supsetneq L_2 \supsetneq \cdots \supsetneq L_k \tag{7}$$

such that

$$\jmath(C(\mathbb{Q})) \subset W + \phi(L_j)$$

for $j = 1, \ldots, k$.

If $q$ is a prime of good reduction for $J$ we denote by

$$\phi_q : \mathbb{Z}^r \to J(\mathbb{F}_q), \qquad \phi_q(a_1, \ldots, a_r) = \sum a_i \tilde{D}_i,$$

and so $\phi_q(\mathbf{l}) = \widetilde{\phi(\mathbf{l})}$.

**Lemma 10.1.** *Let $W$ be a finite subset of $J(\mathbb{Q})$, and let $L$ be a subgroup of $\mathbb{Z}^r$. Suppose that $\jmath(C(\mathbb{Q})) \subset W + \phi(L)$. Let $q$ be a prime of good reduction for $C$ and $J$. Let $L'$ be the kernel of the restriction $\phi_q|_L$. Let $\mathbf{l}_1, \ldots, \mathbf{l}_m$ be representatives of the* **non-zero** *cosets of $L/L'$ and suppose that $\tilde{w} + \phi_q(\mathbf{l}_i) \notin \jmath C(\mathbb{F}_q)$ for all $w \in W$ and $i = 1, \ldots, m$. Then $\jmath(C(\mathbb{Q})) \subset W + \phi(L')$.*

*Proof.* Suppose $P \in C(\mathbb{Q})$. Since $j(C(\mathbb{Q})) \subset W + \phi(L)$, we may write $\jmath(P) = w + \phi(\mathbf{l})$ for some $\mathbf{l} \in L$. Now let $\mathbf{l}_0 = \mathbf{0}$, so that $\mathbf{l}_0, \ldots, \mathbf{l}_m$ represent **all** cosets of $L/L'$. Then $\mathbf{l} = \mathbf{l}_i + \mathbf{l}'$ for some $\mathbf{l}' \in L'$ and $i = 0, \ldots, m$. However, $\phi_q(\mathbf{l}') = 0$, or in other words, $\widetilde{\phi(\mathbf{l}')} = 0$. Hence

$$\jmath(\tilde{P}) = \widetilde{\jmath(P)} = \tilde{w} + \phi_q(\mathbf{l}) = \tilde{w} + \phi_q(\mathbf{l}_i) + \phi_q(\mathbf{l}') = \tilde{w} + \phi_q(\mathbf{l}_i).$$

By hypothesis, $\tilde{w} + \phi_q(\mathbf{l}_i) \notin \jmath C(\mathbb{F}_q)$ for $i = 1, \ldots, m$, so $i = 0$ and so $\mathbf{l}_i = 0$. Hence $\jmath(P) = w + \mathbf{l}' \in W + L'$ as required. □

We obtain a very long strictly decreasing sequence of lattices as in (7) by repeated application of Lemma 10.1. However, the conditions of Lemma 10.1 are unlikely to be satisfied for a prime $q$ chosen at random. Here we give criteria that we have employed in practice to choose the primes $q$.

(I) $\gcd(B, \#J(\mathbb{F}_q)) > (\#J(\mathbb{F}_q))^{0.6}$,
(II) $L' \neq L$,
(III) $\#W \cdot (\#L/L' - 1) < 2q$,
(IV) $\tilde{w} + \phi_q(\mathbf{l}_i) \notin \jmath C(\mathbb{F}_q)$ for all $w \in W$ and $i = 1, \ldots, m$.

The criteria I–IV are listed in the order in which we check them in practice. Criterion IV is just the criterion of the lemma. Criterion II ensures that $L'$ is strictly smaller than $L$, otherwise we gain no new information. Although we would like $L'$ to be strictly smaller than $L$, we do not want the index $L/L'$ to be too large and this is reflected in Criteria I and III. Note that the number of checks required by Criterion IV (or the lemma) is $\#W \cdot (\#L/L' - 1)$. If this number is large then Criterion IV is likely to fail. Let us look at this in probabilistic terms. Assume that the genus of $C$ is 2. Then the probability that a random element of $J(\mathbb{F}_q)$ lies in the image of $C(\mathbb{F}_q)$ is about $1/q$. If $N = \#W \cdot (\#L/L' - 1)$ then the probability that Criterion IV is satisfied is about $(1 - q^{-1})^N$. Since $(1 - q^{-1})^q \sim e^{-1}$, we do not want $N$ to be too large in comparison to $q$, and this explains the choice of $2q$ in Criterion III.

We still have not justified Criterion I. The computation involved in obtaining $L'$ is a little expensive. Since we need to do this with many primes, we would

like a way of picking only primes where this computation is not wasted, and in particular $\#L/L'$ is not too large. Now at every stage of our computations, $L$ will be some element of our decreasing sequence (7) and so contained in $B\mathbb{Z}^n$. Criterion I ensures that a 'large chunk' of $L$ will be in the kernel of $\phi_q : \mathbb{Z}^n \to J(\mathbb{F}_q)$ and so that $\#L/L'$ is not too large. The exponent 0.6 in Criterion I is chosen on the basis of computational experience.

## 11. Lower Bounds for the Size of Rational Points

In this section, we suppose that the strategy of Sections 9 and 10 succeeded in showing that $\jmath(C(\mathbb{Q})) \subset W + \phi(L)$ for some lattice $L$ of huge index in $\mathbb{Z}^r$, where $W$ is the image of $J$ of the set of known rational points in $C$. In this section we provide a lower bound for the size of rational points not belonging to the set of known rational points.

**Lemma 11.1.** *Let $W$ be a finite subset of $J(\mathbb{Q})$, and let $L$ be a sublattice of $\mathbb{Z}^r$. Suppose that $\jmath(C(\mathbb{Q})) \subset W + \phi(L)$. Let $\mu_1$ be a lower bound for $h - \hat{h}$ as in (2). Let*

$$\mu_2 = \max\left\{\sqrt{\hat{h}(w)} \ : \ w \in W\right\}.$$

*Let $M$ be the height-pairing matrix for the Mordell–Weil basis $D_1, \dots, D_r$ and let $\lambda_1, \dots, \lambda_r$ be its eigenvalues. Let*

$$\mu_3 = \min\left\{\sqrt{\lambda_j} \ : \ j = 1, \dots, r\right\}.$$

*Let $m(L)$ be the Euclidean norm of the shortest non-zero vector of $L$. Then, for any $P \in C(\mathbb{Q})$, either $\jmath(P) \in W$ or*

$$h(\jmath(P)) \geq (\mu_3 m(L) - \mu_2)^2 + \mu_1.$$

Note that $m(L)$ is called the minimum of $L$ and can be computed using an algorithm of Fincke and Pohst [18].

*Proof.* Suppose that $\jmath(P) \notin W$. Then $\jmath(P) = w + \phi(\mathbf{l})$ for some non-zero element $\mathbf{l} \in L$. In particular, if $\|\cdot\|$ denotes Euclidean norm then $\|\mathbf{l}\| \geq m(L)$.

We can write $M = N\Lambda N^t$ where $N$ is orthogonal and $\Lambda$ is the diagonal matrix with diagonal entries $\lambda_i$. Let $\mathbf{x} = \mathbf{l}N$ and write $\mathbf{x} = (x_1, \dots, x_r)$. Then

$$\hat{h}(\phi(\mathbf{l})) = \mathbf{l}M\mathbf{l}^t = \mathbf{x}\Lambda\mathbf{x}^t \geq \mu_3^2\|\mathbf{x}\|^2 = \mu_3^2\|\mathbf{l}\|^2 \geq \mu_3^2 m(L)^2.$$

Now recall that $D \mapsto \sqrt{\hat{h}(D)}$ defines a norm on $J(\mathbb{Q}) \otimes \mathbb{R}$ and so by the triangle inequality

$$\sqrt{\hat{h}(\jmath(P))} \geq \sqrt{\hat{h}(\phi(\mathbf{l}))} - \sqrt{\hat{h}(w)} \geq \mu_3 m(L) - \mu_2.$$

The lemma now follows from (2). □

**Remark.** We can replace $\mu_3 m(L)$ with the minimum of $L$ with respect to the height pairing matrix. This is should lead to a very slight improvement. Since in practice our lattice $L$ has very large index, computing the minimum of $L$ with respect to the height pairing matrix may require the computation of the height pairing matrix to very great accuracy, and such a computation is inconvenient. We therefore prefer to work with the Euclidean norm on $\mathbb{Z}^r$.

TABLE 1

| coset of $J(\mathbb{Q})/2J(\mathbb{Q})$ | $\kappa$ | unit rank of $K_i$ | bound $R$ for regulator of $K_i$ | bound for $\log x$ |
|---|---|---|---|---|
| 0 | 1 | 12 | $1.8 \times 10^{26}$ | $1.0 \times 10^{263}$ |
| $D_1$ | $-2\alpha$ | 21 | $6.2 \times 10^{53}$ | $7.6 \times 10^{492}$ |
| $D_2$ | $4 - 2\alpha$ | 25 | $1.3 \times 10^{54}$ | $2.3 \times 10^{560}$ |
| $D_3$ | $-4 - 2\alpha$ | 21 | $3.7 \times 10^{55}$ | $1.6 \times 10^{498}$ |
| $D_1 + D_2$ | $-2\alpha + \alpha^2$ | 21 | $1.0 \times 10^{52}$ | $3.2 \times 10^{487}$ |
| $D_1 + D_3$ | $2\alpha + \alpha^2$ | 25 | $7.9 \times 10^{55}$ | $5.1 \times 10^{565}$ |
| $D_2 + D_3$ | $-4 + \alpha^2$ | 21 | $3.7 \times 10^{55}$ | $1.6 \times 10^{498}$ |
| $D_1 + D_2 + D_3$ | $8\alpha - 2\alpha^3$ | 25 | $7.9 \times 10^{55}$ | $5.1 \times 10^{565}$ |

## 12. Proof of Theorem 1

The equation $Y^2 - Y = X^5 - X$ is transformed into

$$C \quad : \quad 2y^2 = x^5 - 16x + 8, \tag{8}$$

via the change of variables $y = 4Y - 2$ and $x = 2X$ which preserves integrality. We shall work the model (8). Let $C$ be the smooth projective genus 2 curve with affine model given by (8), and let $J$ be its Jacobian. Using `MAGMA` [4] we know that $J(\mathbb{Q})$ is free of rank 3 with Mordell–Weil basis given by

$$D_1 = (0, 2) - \infty, \qquad D_2 = (2, 2) - \infty, \qquad D_3 = (-2, 2) - \infty.$$

The `MAGMA` programs used for this step are based on Stoll's papers [42], [43], [45].

Let $f = x^5 - 16x + 8$. Let $\alpha$ be a root of $f$. We shall choose for coset representatives of $J(\mathbb{Q})/2J(\mathbb{Q})$ the linear combinations $\sum_{i=1}^{3} n_i D_i$ with $n_i \in \{0, 1\}$. Then

$$x - \alpha = \kappa\xi^2,$$

where $\kappa \in \mathcal{K}$ and $\mathcal{K}$ is constructed as in Lemma 2.1. We tabulate the $\kappa$ corresponding to the $\sum_{i=1}^{3} n_i D_i$ in Table 1.

Next we compute the bounds for $\log x$ given by Theorem 2 for each value of $\kappa$. We implemented our bounds in `MAGMA`. Here the Galois group of $f$ is $S_5$ which implies that the fields $K_1$, $K_2$, $K_3$ corresponding to a particular $\kappa$ are isomorphic. The unit ranks of $K_i$, the bounds for their regulator as given by Lemma 4.1, and the corresponding bounds for $\log x$ are tabulated in Table 1.

A quick search reveals 17 rational points on $C$:

$$\infty,\ (-2, \pm 2),\ (0, \pm 2),\ (2, \pm 2),\ (4, \pm 22),\ (6, \pm 62),$$
$$(1/2, \pm 1/8),\ (-15/8, \pm 697/256),\ (60, \pm 9859).$$

Let $W$ denote the image of this set in $J(\mathbb{Q})$. Applying the implementation of the Mordell–Weil sieve due to Bruin and Stoll which is explained in Section 9 we obtain that $\jmath(C(\mathbb{Q})) \subseteq W + BJ(\mathbb{Q})$ where

$$B = 4449329780614748206472972686179940652515754483274306796568214048000$$
$$= 2^8 \cdot 3^4 \cdot 5^3 \cdot 7^3 \cdot 11^2 \cdot 13^2 \cdot 17^2 \cdot 19 \cdot 23 \cdot 29 \cdot 31^2 \cdot \prod_{\substack{37 \le p \le 149 \\ p \ne 107}} p\,.$$

For this computation, we used "deep" information modulo $2^9$, $3^6$, $5^4$, $7^3$, $11^3$, $13^2$, $17^2$, $19^2$, and "flat" information from all primes $p < 50000$ such that $\#J(\mathbb{F}_p)$ is 500-smooth (but keeping only information coming from the maximal 150-smooth quotient group of $J(\mathbb{F}_p)$). Recall that an integer is called *$B$-smooth* if all its prime divisors are $\leq B$. This computation took about 7 hours on a 2 GHz Intel Core 2 CPU.

We now apply the new extension of the Mordell–Weil sieve explained in Section 10. We start with $L_0 = B\mathbb{Z}^3$ where $B$ is as above. We successively apply Lemma 10.1 using all primes $q < 10^6$ which are primes of good reduction and satisfy criteria I–IV of Section 10. There are 78498 primes less than $10^6$. Of these, we discard 2, 139, 449 as they are primes of bad reduction for $C$. This leaves us with 78495 primes. Of these, Criterion I fails for 77073 of them, Criterion II fails for 220 of the remaining, Criterion III fails for 43 primes that survive Criteria I and II, and Criterion IV fails for 237 primes that survive Criteria I–III. Altogether, only 922 primes $q < 10^6$ satisfy Criteria I–IV and increase the index of $L$.

The index of the final $L$ in $\mathbb{Z}^3$ is approximately $3.32 \times 10^{3240}$. This part of the computation lasted about 37 hours on a 2.8 GHZ Dual-Core AMD Opteron.

Let $\mu_1$, $\mu_2$, $\mu_3$ be as in the notation of Lemma 11.1. Using `MAGMA` we find $\mu_1 = 2.677$, $\mu_2 = 2.612$ and $\mu_3 = 0.378$ (to 3 decimal places). The shortest vector of the final lattice $L$ is of Euclidean length approximately $1.156 \times 10^{1080}$ (it should be no surprise that this is roughly the cube root of the index of $L$ in $\mathbb{Z}^3$). By Lemma 11.1 if $P \in C(\mathbb{Q})$ is not one of the 17 known rational points then

$$h(\jmath(P)) \geq 1.9 \times 10^{2159}.$$

If $P$ is an integral point, then $h(\jmath(P)) = \log 2 + 2 \log x(P)$. Thus

$$\log x(P) \geq 0.95 \times 10^{2159}.$$

This contradicts the bounds for $\log x$ in Table 1 and shows that the integral point $P$ must be one of the 17 known rational points. This completes the proof of Theorem 1.

## References


[1] A. Baker, *Bounds for the solutions of the hyperelliptic equation*, Proc. Cambridge Philos. Soc. **65** (1969), 439–444.

[2] Yu. Bilu, *Effective analysis of integral points on algebraic curves*, Israel J. Math. **90** (1995), 235–252.

[3] Yu. F. Bilu and G. Hanrot, *Solving superelliptic Diophantine equations by Baker's method*, Compositio Mathematica **112** (1998), 273–312.

[4] W. Bosma, J. Cannon and C. Playoust: *The Magma Algebra System I: The User Language*, J. Symb. Comp. **24** (1997), 235–265. (See also `http://www.maths.usyd.edu.au/`)

[5] B. Brindza, *On $S$-integral solutions of the equation $y^m = f(x)$*, Acta. Math. Hungar. **44** (1984), 133–139.

[6] N. Bruin, *Chabauty methods and covering techniques applied to generalized Fermat equations*, Dissertation, University of Leiden, Leiden, 1999.

[7] N. Bruin, *Chabauty methods using elliptic curves*, J. reine angew. Math. **562** (2003), 27–49.

[8] N. Bruin and M. Stoll, *Deciding existence of rational points on curves: an experiment*, to appear in Experimental Math.

[9] N. Bruin and M. Stoll, *Two-cover descent on hyperelliptic curves*, in preparation.

[10] N. Bruin and M. Stoll, *The Mordell–Weil sieve: proving the non-existence of rational points on curves*, in preparation.

[11] Y. Bugeaud, *Bounds for the solutions of superelliptic equations*, Compositio Math. *107* (1997), 187–219.

[12] Y. Bugeaud and K. Győry, *Bounds for the solutions of unit equations*, Acta Arith. **74** (1996), 67–80.
[13] Y. Bugeaud, M. Mignotte and S. Siksek, *Classical and modular approaches to exponential Diophantine equations I. Fibonacci and Lucas perfect powers*, Annals of Math. **163** (2006), 969–1018.
[14] J. W. S. Cassels and E. V. Flynn, *Prolegomena to a middlebrow arithmetic of curves of genus* 2, L.M.S. lecture notes series **230**, Cambridge University Press, 1997.
[15] S. David, *Minorations de formes linéaires de logarithmes elliptiques*, Mém. Soc. Math. France **62** (1995).
[16] J.-H. Evertse and R. Tijdeman, *Some open problems about Diophantine equations*, `http://www.math.leidenuniv.nl/~evertse/07-workshop-problems.pdf`
[17] D. C. Fielder and C. O. Alford, *Observations from computer experiments on an integer equation*, in *Applications of Fibonacci numbers* **7** (Graz, 1996), pages 93–103, Kluwer Acad. Publ. Dordrecht, 1998.
[18] U. Fincke and M. Pohst, *Improved methods for calculating vectors of short length in a lattice, including complexity analysis*, Math. Comp. **44**, 463–471, 1985.
[19] E. V. Flynn, *Descent via isogeny in dimension* 2, Acta Arith. **LXVI.1** (1994), 23–43.
[20] E. V. Flynn, *A flexible method for applying Chabauty's Theorem*, Compositio Math. **105** (1997), 79–94.
[21] E. V. Flynn and N. P. Smart, *Canonical heights on the Jacobians of curves of genus* 2 *and the infinite descent*, Acta Arith. **79** (1997), no. 4, 333–352.
[22] E. V. Flynn and J. L. Wetherell, *Finding rational points on bielliptic genus 2 curves*, Manuscripta Math. **100** (1999), no. 4, 519–533.
[23] E. V. Flynn and J. L. Wetherell, *Covering collections and a challenge problem of Serre*, Acta Arith. **98** (2001), no. 2, 197–205.
[24] J. Gebel, A. Pethő, and H. G. Zimmer, *Computing integral points on elliptic curves*, Acta Arith. **68** (1994), no. 2, 171–192.
[25] E. Landau, *Verallgemeinerung eines Pólyaschen Satzes auf algebraische Zahlkörper*, Nachr. Kgl. Ges. Wiss. Göttingen, Math.-Phys. Kl. (1918), 478–488.
[26] W. McCallum and B. Poonen, *The method of Chabauty and Coleman*, preprint, 19 September 2006.
[27] E. M. Matveev, *An explicit lower bound for a homogeneous rational linear form in logarithms of algebraic numbers. II*, Izv. Ross. Acad. Nauk Ser. Mat. **64** (2000), no. 6, 125–180; English translation in Izv. Math. **64** (2000), no. 6, 1217–1269.
[28] M. Mignotte and A. Pethő, *On the Diophantine equation* $x^p - x = y^q - y$, Publ. Mat. **43** (1999), no. 1, 207–216.
[29] L. J. Mordell, *On the integer solutions of* $y(y+1) = x(x+1)(x+2)$, Pacific J. Math. **13** (1963), 1347–1351.
[30] A. Pethő and B. M. M. de Weger, *Products of prime powers in binary recurrence sequences Part I: The hyperbolic case, with applications to the Generalized Ramanujan–Nagell equation*, Math. Comp. **47** (1987), 713–727.
[31] B. Poonen and E. F. Schaefer, *Explicit descent on cyclic covers of the projective line*, J. reine angew. Math. **488** (1997), 141–188.
[32] D. Poulakis, *Solutions entières de l'équation* $y^m = f(x)$, Sém. Théor. Nombres Bordeaux **3** (1991), 187–199.
[33] E. F. Schaefer, 2*–descent on the Jacobians of hyperelliptic curves*, J. Number Theory **51** (1995), 219–232.
[34] E. F. Schaefer and J. L. Wetherell, *Computing the Selmer group of an isogeny between abelian varieties using a further isogeny to a Jacobian*, J. Number Theory **115** (2005), 158–175.
[35] W. M. Schmidt, *Integer points on curves of genus* 1, Compositio Math. **81** (1992), 33–59.
[36] J. H. Silverman, *The Arithmetic of Elliptic Curves*, GTM **106**, Springer–Verlag, 1992.
[37] N. P. Smart, *S-integral points on elliptic curves*, Proc. Camb. Phil. Soc. **116** (1994), 391–399.
[38] N. P. Smart, *The Algorithmic Resolution of Diophantine Equations*, LMS Student Texts **41**, Cambridge University Press, 1998.
[39] N. P. Smart and N. M. Stephens, *Integral points on elliptic curves over number fields*, Proc. Camb. Phil. Soc. **122** (1997), 9–16.
[40] V. G. Sprindžuk, *The arithmetic structure of integer polynomials and class numbers*, Trdu Mat. Inst. Steklov **LV** (1977), 152–174.

[41] M. Stoll, *On the arithmetic of the curves $y^2 = x^l + A$ and their Jacobians*, J. reine angew. Math. **501** (1998), 171–189.
[42] M. Stoll, *On the height constant for curves of genus two*, Acta Arith. **90** (1999), 183–201.
[43] M. Stoll, *Implementing 2-descent for Jacobians of hyperelliptic curves*, Acta Arith. **98** (2001), 245–277.
[44] M. Stoll, *On the arithmetic of the curves $y^2 = x^l + A$, II*, J. Number Theory **93** (2002), 183–206.
[45] M. Stoll, *On the height constant for curves of genus two, II*, Acta Arith. **104** (2002), 165–182.
[46] M. Stoll, *Independence of rational points on twists of a given curve*, Compositio Math. **142** (2006), 1201–1214.
[47] R. J. Stroeker and N. Tzanakis, *Solving elliptic Diophantine equations by estimating linear forms in elliptic logarithms*, Acta Arith. **67** (1994), 177–196.
[48] R. J. Stroeker and N. Tzanakis, *Computing all integer solutions of a genus* 1 *equation*, Math. Comp. **72** (2003), no. 244, 1917–1933
[49] R. J. Stroeker and B. M. M. de Weger, *Solving elliptic Diophantine equations: the general cubic case*, Acta Arith. **87** (1999), no. 4, 339–365.
[50] N. Tzanakis, *Solving elliptic Diophantine equations by estimating linear forms in elliptic logarithms. The case of quartic equations*, Acta Arith. **75** (1996), 165–190.
[51] P. M. Voutier, *An upper bound for the size of integral solutions to $Y^m = f(X)$*, J. Number Theory **53** (1995), no. 2, 247–271.
[52] P. M. Voutier, *An effective lower bound for the height of algebraic numbers*, Acta Arith. **LXXIV.1** (1996), 81–95.
[53] J. L. Wetherell, *Bounding the Number of Rational Points on Certain Curves of High Rank*, Ph.D. dissertation, University of California at Berkeley, 1997.

Yann Bugeaud and Maurice Mignotte, Université Louis Pasteur, U. F. R. de mathématiques, 7, rue René Descartes, 67084 Strasbourg Cedex, France
*E-mail address*: bugeaud@math.u-strasbg.fr
*E-mail address*: mignotte@math.u-strasbg.fr

Samir Siksek, Institute of Mathematics, University of Warwick, Coventry, CV4 7AL, United Kingdom
*E-mail address*: s.siksek@warwick.ac.uk

Michael Stoll, School of Engineering and Science, Jacobs University Bremen, P.O. Box 75 05 61, 28 725 Bremen, Germany
*E-mail address*: m.stoll@jacobs-university.de

Szabolcs Tengely, Institute of Mathematics, University of Debrecen and the Number Theory Research Group of the Hungarian Academy of Sciences, P.O.Box 12, 4010 Debrecen, Hungary
*E-mail address*: tengely@math.klte.hu