MATHEMATICS OF COMPUTATION
Volume 67, Number 221, January 1998, Pages 413–438
S 0025-5718(98)00900-4

ON THE DIOPHANTINE EQUATION $|ax^n - by^n| = 1$

MICHAEL A. BENNETT AND BENJAMIN M. M. DE WEGER

Abstract. If a, b and n are positive integers with $b \geq a$ and $n \geq 3$, then the equation of the title possesses at most one solution in positive integers x and y, with the possible exceptions of (a, b, n) satisfying $b = a + 1$, $2 \leq a \leq \min\{0.3n, 83\}$ and $17 \leq n \leq 347$. The proof of this result relies on a variety of diophantine approximation techniques including those of rational approximation to hypergeometric functions, the theory of linear forms in logarithms and recent computational methods related to lattice-basis reduction. Additionally, we compare and contrast a number of these last mentioned techniques.

1. Introduction

In 1909, Thue [Th] used a result on rational approximation to algebraic numbers to show that if $F(x, y)$ is an irreducible binary form (in $\mathbb{Z}[x, y]$) of degree at least 3, and m a nonzero integer, then the equation

$$F(x, y) = m \tag{1}$$

has at most finitely many solutions in integers x and y. This fundamental relationship between homogeneous (and related) diophantine equations and diophantine approximation has been exploited in subsequent years in bounding the number of solutions of given equations and even the size of such solutions. The equation

$$|ax^n - by^n| = 1 \tag{2}$$

where a, b and n are nonzero integers and $n \geq 3$, is, in a certain sense, the simplest case of (1), and has been frequently studied both from the viewpoint of diophantine approximation and from a more algebraic perspective. In particular, Delone [De] and Nagell [N] independently showed that if $n = 3$, then equation (2) has at most one solution in positive integers x and y, corresponding (if it exists) to the fundamental unit of $\mathbb{Q}(\sqrt[3]{a/b})$ (see also [Lj1]). Later, paralleling this (primarily algebraic) approach, Ljunggren [Lj1] (see also [DF] and [Ta]) proved a like result for equations of the form

$$|ax^4 - by^4| = 1$$

(i.e. that they too possess at most one solution in positive integers for each pair (a, b) of nonzero integers).

Received by the editor July 22, 1996 and, in revised form, October 7, 1996.
1991 *Mathematics Subject Classification.* Primary 11D41; Secondary 11Y50.
De Weger's research was supported by the Netherlands Mathematical Research Foundation SWON with financial aid from the Netherlands Organization for Scientific Research NWO.

In 1937, by extending Thue's method and constructing explicit rational function (Padé) approximations to $\sqrt[n]{1-z}$, Siegel [Si] deduced that if c is a given real number, then the inequality

$$|ax^n - by^n| \leq c \tag{3}$$

has at most one (positive) integral solution (x, y), provided

$$|ab|^{n/2-1} \geq 4 \left(n \prod_{p|n} p^{\frac{1}{p-1}} \right)^n c^{2n-2}. \tag{4}$$

By refining this approach, Domar [Do] was able to prove that (2) has at most two solutions in positive integers under the restriction that $n \geq 5$, and that, if $a = 1$, equation (2) possesses at most one positive solution, except possibly when $b = 2$ or $n = 5$ or 6 and $b = 2^n \pm 1$. In the special case when $b = 2$ in the above equation, Darmon and Merel [DM] have shown that no solutions exist with $xy > 1$, as a consequence of a much more general result extending Wiles' remarkable work on the Shimura-Taniyama-Weil conjecture. With a fundamental improvement of Siegel's gap principle (to prevent potential solutions to inequality (3) from being too close together in size), Evertse [Ev1] significantly relaxed condition (4) (see also Mueller [Mu] for a rather different treatment, closer to Thue's original approach). For a more detailed historical perspective of results on equation (2), the reader is directed to [Mo] and [R].

In this paper, we combine the Thue-Siegel machinery (as used by Evertse) with recent explicit bounds for rational approximation to algebraic numbers due to the first author [Be2] (see also [Be1]), new estimates for linear forms in the logarithms of two algebraic numbers due to Laurent, Mignotte and Nesterenko [LMN], somewhat older estimates for linear forms in the logarithms of several algebraic numbers due to Baker and Wüstholz [BW] and techniques from computational diophantine approximation. We prove

Theorem 1.1. *If a, b and n are integers with $b > a \geq 1$ and $n \geq 3$, then the equation* (2) *has at most one solution in positive integers (x, y), except possibly for the cases where $b = a + 1$, $2 \leq a \leq \min\{0.3n, 83\}$ and $17 \leq n \leq 347$.*

It should be noted that this approach combining the techniques of linear forms in logarithms (the Gel′fond-Baker method, and, to be more precise in our usage in Section 4, the Schneider-Waldschmidt method) with irrationality measures derived from consideration of hypergeometric functions (the Thue-Siegel-Baker method) has been utilized previously on similar problems, by, for example, Shorey [Sh] and Shorey and Tijdeman [ST]. Additionally, Mignotte [Mi] has recently applied the aforementioned bounds for linear forms in two logarithms to deduce a number of results of a flavour reminiscent of the above (including that (2) has exactly one positive solution for $b = a + 1$ and $n > 600$). The advantage of Theorem 1.1 is that it provides a very explicit bound upon both a and n and treats small values of n. Further, the cases omitted above (which by Domar's theorem may possess at most two such solutions) may each be "effectively solved" via the theory of linear forms in (several) logarithms (see e.g. [Ba3]). In reality, however, as one may observe from a perusal of Section 3, this appears to be a rather difficult computational problem.

It was conjectured by Siegel and proved by Mueller and Schmidt [MS] that the number of solutions to the general Thue equation (1) depends only upon m and the number of monomials present in the form $F(x, y)$. In this regard, equation (2) is a minimal case. It appears that the techniques of this paper are not particularly well suited to generalization to nonbinomial forms (unless $n = 3$; see [Ev3]).

This paper is organized as follows. In Section 2 we apply arguments based on rational function approximation to hypergeometric functions (à la Thue-Siegel) to prove Theorem 1.1 for "small" n relative to $\max\{a, b\}$. In Section 3, we treat a number of special cases with $5 \leq n \leq 13$ where the aforementioned techniques fail to apply. Here we use linear forms in several logarithms of algebraic numbers and tools from computational diophantine approximation. Additionally, we compare and contrast the efficiency of certain of these methods, relative to this problem. Finally, in Section 4, we state a lower bound for linear forms in the logarithms of pairs of algebraic numbers, due to Laurent, Mignotte and Nesterenko [LMN], and use it to finish the proof of Theorem 1.1.

2. The method of Thue-Siegel

In [Ev1], refining Siegel's result in [Si], Evertse proved

Theorem 2.1. *If a, b and n are positive integers with $n \geq 3$ and c is a positive real number, then there is at most one positive integral solution (x, y) to the inequality*

$$|ax^n - by^n| \leq c$$

with

$$\max\{|ax^n|, |by^n|\} > \beta_n c^{\alpha_n},$$

where β_n and α_n are effectively computable positive constants satisfying $\beta_3 = 1152.2$, $\beta_4 = 98.53$ and $\beta_n < n^2$ for $n \geq 5$.

While techniques from [Be1] and [Be2] enable us to sharpen this somewhat, the above formulation is adequate for our purposes (as, for that matter, is an earlier sharpening of Siegel's result due to Hyrrö [H], at least for $n \geq 7$). For details of the proof of Theorem 2.1, which utilizes Padé approximants to $\sqrt[n]{1-z}$ (à la Siegel) together with an iterated gap principle, the reader is directed to [Ev1] (note: the corresponding result in [Ev2] is significantly weaker if $n = 3$).

Let us take $c = 1$ in the above theorem and assume, without loss of generality, that $b > a \geq 1$. By the aforementioned results of Delone, Nagell and Ljunggren, we may assume that $n \geq 5$ so that

$$2^n > n^2 > \beta_n$$

in Theorem 2.1. In these cases, we therefore have at most one solution to (2) with $\max\{|x|, |y|\} > 1$. It follows that we may restrict our attention to equations with solution $(x, y) = (1, 1)$, namely those of the form

$$|ax^n - (a+1)y^n| = 1 \tag{5}$$

where a and n are positive integers with $n \geq 3$ (and, from Darmon and Merel, $a \geq 2$). Suppose that (x, y) is a positive solution to (5). Then we have

$$\left| \sqrt[n]{1 + \frac{1}{a}} - \frac{x}{y} \right| < \frac{1}{any^n} \tag{6}$$

so that x/y is an exceptionally good rational approximation to $\sqrt[n]{1+\frac{1}{a}}$. To eliminate this possibility for $x > y > 1$, at least with a handful of exceptions, we appeal to the following special cases of two results of the first author (see [Be2]). Define

$$\mu_n = \prod_{p|n} p^{\frac{1}{p-1}}.$$

We have

Theorem 2.2. *For integer n, define the constant $c(n)$ by*

n	$c(n)$	n	$c(n)$	n	$c(n)$	n	$c(n)$	n	$c(n)$
3	2.03	11	1.67	23	1.53	41	1.45	59	1.40
4	1.62	13	1.65	29	1.51	43	1.43	61	1.39
5	1.84	17	1.58	31	1.51	47	1.44	67	1.38
7	1.76	19	1.56	37	1.46	53	1.40	71	1.36

Suppose that a, n, p and q are positive integers with n occurring in the above table. If, further, we have that

$$\left(\sqrt{a}+\sqrt{a+1}\right)^{2(n-2)} > \left(\frac{n\mu_n}{c(n)}\right)^n,$$

then we can conclude that

$$\left|\sqrt[n]{1+\frac{1}{a}} - \frac{p}{q}\right| > a^{-1}(10^{10}\ q)^{-\lambda_1}$$

with

$$\lambda_1 = 1 + \frac{\log\left(\frac{n\mu_n}{c(n)}\left(\sqrt{a}+\sqrt{a+1}\right)^2\right)}{\log\left(\frac{c(n)}{n\mu_n}\left(\sqrt{a}+\sqrt{a+1}\right)^2\right)}.$$

And we have

Theorem 2.3. *If a, n, p and q are positive integers with $n \geq 3$ and*

$$\left(\sqrt{a}+\sqrt{a+1}\right)^{2(n-2)} > (n\mu_n)^n,$$

then

$$\left|\sqrt[n]{1+\frac{1}{a}} - \frac{p}{q}\right| > (8n\mu_n a)^{-1} q^{-\lambda_2}$$

with

$$\lambda_2 = 1 + \frac{\log\left(n\mu_n\left(\sqrt{a}+\sqrt{a+1}\right)^2\right)}{\log\left(\frac{1}{n\mu_n}\left(\sqrt{a}+\sqrt{a+1}\right)^2\right)}.$$

Both of these results follow from consideration of Padé approximants to the binomial function and differ, in essense, in that the former includes information about the p-adic valuations of binomial coefficients appearing in the approximating polynomials (see also [Ba1], [Ba2] and [Ch] for a detailed discussion of this approach).

If we combine these two theorems with (6), we derive bounds upon solutions (x, y) to (5), of the form

$$y < \left(10^{10\lambda_1}/n\right)^{\frac{1}{n-\lambda_1}} \tag{7}$$

provided $\lambda_1 < n$, or

$$y < (8\mu_n)^{\frac{1}{n-\lambda_2}} \tag{8}$$

if $\lambda_2 < n$. We use these inequalities to prove the following two lemmas which summarize our refinements of Theorem 2.1 in the situation related to equation (5).

Lemma 2.4. *If a and n are positive integers with $n \geq 3$ such that equation (5) has more than a single positive solution, then $a < 0.3n$.*

Lemma 2.5. *If a and n are positive integers with $3 \leq n \leq 16$, then the only solution to equation (5) in positive integers is given by $x = y = 1$.*

To obtain these results, we further require:

Lemma 2.6. *If (x, y) is a positive solution to equation (5), then either $x = y = 1$ or $\min\{x, y\} > an$.*

Proof of Lemma 2.6. If $x \leq y$ and $y > 1$, then

$$|ax^n - (a+1)y^n| \geq y^n > 1$$

so that if at least one of x or y exceeds 1, we may suppose that $x \geq y + 1$, whereby

$$ax^n - (a+1)y^n \geq a(y+1)^n - (a+1)y^n.$$

By the binomial theorem, this equals

$$any^{n-1} - y^n + a\sum_{k=2}^{n} \binom{n}{k} y^k$$

and since $a\sum_{k=2}^{n}\binom{n}{k}y^k > 1$, we require that $any^{n-1} < y^n$, whence $y = \min\{x, y\} > an$. $\square$

Noting that, if $n = 3$ or 4, Lemma 2.5 follows from the work of Delone, Nagell and Ljunggren, it is clearly sufficient to prove Lemmas 2.4 and 2.5 for prime values of $n \geq 5$. First, let us suppose that $n \geq 79$ is prime. Then if $a \geq 0.3n$ and (x, y) is a positive solution to (5), we have from Theorem 2.3 and (8) that

$$y < \left(8n^{1/(n-1)}\right)^{1/(n-\lambda_2)}$$

where

$$\lambda_2 < 1 + \frac{\log\left(n^{n/(n-1)}(1.2n)\right)}{\log\left(n^{-n/(n-1)}(1.2n)\right)} < 16\log n + 3 < n - 6.$$

Thus, $y < 9^{1/6} < 2$, so that $x = y = 1$. Now, if $n = 59, 61, 67, 71$ or 73, we apply Theorem 2.3 and (8) to deduce that $y < an$ for $a \geq 18, 19, 20, 21$ and 22 respectively. Together with Lemma 2.6 this implies that $x = y = 1$ for $a \geq 0.3n$ in these cases as well. For smaller primes, we apply (7) and (8) to find bounds upon solutions (x, y) to (5) of the form $y < c_0$ or better (where (?) indicates which of inequalities (7) or (8) is used), as in the following table.

n	a	c_0	(?)	n	a	c_0	(?)	a	n	c_0	(?)
5	$3 \le a \le 6$	10^{59}	(7)	17	$a \ge 8$	$17a$	(8)	37	$a = 12$	10^6	(7)
5	$a \ge 7$	$5a$	(8)	19	$6 \le a \le 7$	10^{14}	(7)	37	$a \ge 13$	$37a$	(8)
7	$3 \le a \le 5$	10^{43}	(7)	19	$a \ge 8$	$19a$	(8)	41	$a = 13$	10^6	(7)
7	$a \ge 6$	$7a$	(8)	23	$7 \le a \le 8$	10^{12}	(7)	41	$a \ge 14$	$41a$	(8)
11	$3 \le a \le 6$	10^{158}	(7)	23	$a \ge 9$	$23a$	(8)	43	$a = 13$	10^6	(7)
11	$a \ge 7$	$11a$	(8)	29	$9 \le a \le 10$	10^8	(7)	43	$a \ge 14$	$43a$	(8)
13	$4 \le a \le 6$	10^{27}	(7)	29	$a \ge 11$	$29a$	(8)	47	$a \ge 15$	$47a$	(8)
13	$a \ge 7$	$13a$	(8)	31	$a = 10$	10^7	(7)	53	$a = 16$	10^5	(7)
17	$6 \le a \le 7$	10^{12}	(7)	31	$a \ge 11$	$31a$	(8)	53	$a \ge 17$	$53a$	(8)

By virtue of Lemma 2.6 and the above table, to complete the proofs of Lemmas 2.4 and 2.5 we need only consider the 27 cases above where we fail to obtain an upper bound of the form $y < an$ upon possible solutions to (5), as well as the pairs (a, n) satisfying $a = 2$ (for $n = 5, 7$ and 11) or $2 \le a \le 3$ (for $n = 13$). These latter cases will be dealt with in detail in Section 3. In the former situation, we observe from (6) that a positive solution to (5) corresponds to a convergent in the continued fraction expansion to $\sqrt[n]{1+\frac{1}{a}}$. For such a convergent p_i/q_i, we have (see e.g. [Le])

$$\left| \sqrt[n]{1+\frac{1}{a}} - \frac{p_i}{q_i} \right| > \frac{1}{(a_{i+1}+2)q_i^2}$$

where a_{i+1} is the $(i+1)$st partial quotient in the aforementioned continued fraction expansion. It therefore follows from (6) that a solution (x, y) to (5) (with $x/y = p_i/q_i$) induces a partial quotient a_{i+1} satisfying

$$a_{i+1} \ge anq_i^{n-2} - 1. \tag{9}$$

For each of the 27 pairs (a, n) under consideration, we compute the initial terms in the continued fraction expansion to $\sqrt[n]{1+\frac{1}{a}}$ and verify in each case that none of the first five convergents yields a solution to (5) other than with $x = y = 1$. Since we always find that $q_5 \ge 151$ (where equality is obtained for $(a, n) = (3, 11)$), inequality (9) implies that we require a partial quotient exceeding 10^7 in order to contradict Theorem 1.1. The previously derived upper bounds upon the denominators of the convergents allow us to further restrict our attention to, at most, the first 314 partial quotients in each expansion (corresponding, again, to $(a, n) = (3, 11)$ where we find precisely 313 convergents with $q_i < 10^{158}$). Since the largest partial quotient we find in the ranges under consideration is $a_{308} = 3397$ (for $(a, n) = (5, 11)$), we conclude as stated.

3. Some heavier computations

3.1. **Introduction.** In this section we complete the proof of Lemma 2.5 through

Theorem 3.1. *The diophantine equation*

$$ax^n - (a+1)y^n = 1 \tag{10}$$

with $(n,a) = (5,2), (7,2), (11,2), (13,2)$ *or* $(13,3)$ *has only* $x = y = -1$ *as solution in rational integers* x, y.

We obtain this result by the essentially routine method for solving Thue equations, following Tzanakis and de Weger [TW] (see also [dW]), using the lower bound for linear forms in logarithms of algebraic numbers from Baker and Wüstholz [BW] and a new variant of the computational diophantine approximation method, combining ideas from Bilu and Hanrot [BH] and Mignotte and de Weger [MW].

3.2. **The relevant field data.** Using Pari 1.39.03 on a workstation and a Pentium 75 personal computer we computed the following data on the relevant algebraic number fields.

3.2.1. *The case* $(n,a) = (5,2)$. Let θ be a root of $t^5 - 48$, and put $\mathbb{K} = \mathbb{Q}(\theta)$. The discriminant is $2^4 3^4 5^5$, an integral basis is

$$\left\{1, \theta, \tfrac{1}{2}\theta^2, \tfrac{1}{4}\theta^3, \tfrac{1}{8}\theta^4\right\},$$

the class group is trivial, the regulator is $49.089947\ldots$, a system of fundamental units (both of norm 1) is ϵ_1, ϵ_2, where the following table gives their coefficients in terms of the integral basis:

	1	θ	$\frac{1}{2}\theta^2$	$\frac{1}{4}\theta^3$	$\frac{1}{8}\theta^4$
ϵ_1	1	1	2	2	1
ϵ_2	1	-29	-17	15	23

and the rational prime 2 ramifies as $(2) = (\rho)^5$, where

$$\rho = 2 + \theta + \frac{1}{2}\theta^2 + \frac{1}{4}\theta^3 + \frac{1}{8}\theta^4 = \frac{2}{\theta - 2}$$

is a prime of norm 2 (in fact, the only one up to multiplication by units).

3.2.2. *The case* $(n,a) = (7,2)$. Let θ be a root of $t^7 - 192$, and put $\mathbb{K} = \mathbb{Q}(\theta)$. The discriminant is $-2^6 3^6 7^7$, an integral basis is

$$\left\{1, \theta, \frac{1}{2}\theta^2, \frac{1}{4}\theta^3, \frac{1}{8}\theta^4, \frac{1}{16}\theta^5, \frac{1}{32}\theta^6\right\},$$

the class group is trivial, the regulator is $765.90150\ldots$, a system of fundamental units (all of norm 1) is $\epsilon_1, \epsilon_2, \epsilon_3$, where the following table gives their coefficients in terms of the integral basis:

	1	θ	$\frac{1}{2}\theta^2$	$\frac{1}{4}\theta^3$	$\frac{1}{8}\theta^4$	$\frac{1}{16}\theta^5$	$\frac{1}{32}\theta^6$
ϵ_1	7	4	2	4	3	2	3
ϵ_2	-5	2	0	-2	2	-3	3
ϵ_3	7	0	-1	-1	1	3	3

and the rational prime 2 ramifies as $(2) = (\rho)^7$, where

$$\rho = 2 + \theta + \tfrac{1}{2}\theta^2 + \tfrac{1}{4}\theta^3 + \tfrac{1}{8}\theta^4 + \tfrac{1}{16}\theta^5 + \tfrac{1}{32}\theta^6 = \frac{2}{\theta - 2}$$

is a prime of norm 2 (in fact, the only one up to multiplication by units).

3.2.3. *The case* $(n, a) = (11, 2)$. Let θ be a root of $t^{11} - 3072$, and put $\mathbb{K} = \mathbb{Q}(\theta)$. The discriminant is $-2^{10}3^{10}11^{11}$, an integral basis is

$$\left\{1, \theta, \frac{1}{2}\theta^2, \frac{1}{4}\theta^3, \frac{1}{8}\theta^4, \frac{1}{16}\theta^5, \frac{1}{32}\theta^6, \frac{1}{64}\theta^7, \frac{1}{128}\theta^8, \frac{1}{256}\theta^9, \frac{1}{512}\theta^{10}\right\},$$

the class group is trivial, the regulator is $410432.22\ldots$, a system of fundamental units (all of norm 1) is $\epsilon_1, \ldots, \epsilon_5$, where the following table gives their coefficients in terms of the integral basis:

	1	θ	$\frac{1}{2}\theta^2$	$\frac{1}{4}\theta^3$	$\frac{1}{8}\theta^4$	$\frac{1}{16}\theta^5$	$\frac{1}{32}\theta^6$	$\frac{1}{64}\theta^7$	$\frac{1}{128}\theta^8$	$\frac{1}{256}\theta^9$	$\frac{1}{512}\theta^{10}$
ϵ_1	-5	-1	2	2	0	-2	-1	1	2	1	-1
ϵ_2	1	2	0	1	0	-1	-1	-1	0	-1	1
ϵ_3	-17	-1	-9	1	15	-3	-4	-3	-7	10	6
ϵ_4	-5	-6	-5	0	2	2	4	5	2	-1	-1
ϵ_5	-35	-24	-21	-10	3	15	21	20	15	6	-4

and the rational prime 2 ramifies as $(2) = (\rho)^{11}$, where

$$\rho = 2 + \theta + \frac{1}{2}\theta^2 + \frac{1}{4}\theta^3 + \frac{1}{8}\theta^4 + \frac{1}{16}\theta^5 + \frac{1}{32}\theta^6 + \frac{1}{64}\theta^7 + \frac{1}{128}\theta^8 + \frac{1}{256}\theta^9 + \frac{1}{512}\theta^{10} = \frac{2}{\theta - 2}$$

is a prime of norm 2 (in fact, the only one up to multiplication by units).

3.2.4. *The case* $(n, a) = (13, 2)$. Let θ be a root of $t^{13} - 12288$, and put $\mathbb{K} = \mathbb{Q}(\theta)$. The discriminant is $2^{12}3^{12}13^{13}$, an integral basis is

$$\left\{1, \theta, \frac{1}{2}\theta^2, \frac{1}{4}\theta^3, \frac{1}{8}\theta^4, \frac{1}{16}\theta^5, \frac{1}{32}\theta^6, \frac{1}{64}\theta^7, \frac{1}{128}\theta^8, \frac{1}{256}\theta^9, \frac{1}{512}\theta^{10}, \frac{1}{1024}\theta^{11}, \frac{1}{2048}\theta^{12}\right\},$$

the class group is trivial, the regulator is $12465830.\ldots$, a system of fundamental units (all of norm 1) is $\epsilon_1, \ldots, \epsilon_6$, where the following table gives their coefficients in terms of the integral basis:

	1	θ	$\frac{1}{2}\theta^2$	$\frac{1}{4}\theta^3$	$\frac{1}{8}\theta^4$	$\frac{1}{16}\theta^5$	$\frac{1}{32}\theta^6$	$\frac{1}{64}\theta^7$	$\frac{1}{128}\theta^8$	$\frac{1}{256}\theta^9$	$\frac{1}{512}\theta^{10}$	$\frac{1}{1024}\theta^{11}$	$\frac{1}{2048}\theta^{12}$
ϵ_1	7	4	5	3	3	4	3	2	3	3	2	3	3
ϵ_2	1	-1	0	1	0	2	0	0	1	-1	0	-1	-1
ϵ_3	55	-8	-21	22	5	-24	12	15	-21	0	19	-13	-8
ϵ_4	13	22	0	-20	-5	17	8	-15	-12	11	14	-7	-14
ϵ_5	-41	13	-6	14	-13	7	-2	4	-4	-4	8	-3	4
ϵ_6	-743	-970	122	941	105	-859	-300	737	454	-586	-563	419	627

and the rational prime 2 ramifies as $(2) = (\rho)^{13}$, where

$$\rho = 2 + \theta + \frac{1}{2}\theta^2 + \frac{1}{4}\theta^3 + \frac{1}{8}\theta^4 + \frac{1}{16}\theta^5 + \frac{1}{32}\theta^6 + \frac{1}{64}\theta^7 + \frac{1}{128}\theta^8 + \frac{1}{256}\theta^9 + \frac{1}{512}\theta^{10} + \frac{1}{1024}\theta^{11} + \frac{1}{2048}\theta^{12} = \frac{2}{\theta - 2}$$

is a prime of norm 2 (in fact, the only one up to multiplication by units).

3.2.5. *The case* $(n, a) = (13, 3)$. Let θ be a root of $t^{13} - 1458$, and put $\mathbb{K} = \mathbb{Q}(\theta)$. The discriminant is $2^{12}3^{12}13^{13}$, an integral basis is

$$\left\{1, \theta, \theta^2, \frac{1}{3}\theta^3, \frac{1}{3}\theta^4, \frac{1}{9}\theta^5, \frac{1}{9}\theta^6, \frac{1}{27}\theta^7, \frac{1}{27}\theta^8, \frac{1}{81}\theta^9, \frac{1}{81}\theta^{10}, \frac{1}{243}\theta^{11}, \frac{1}{243}\theta^{12}\right\},$$

the class group is trivial, the regulator is $12555373.\ldots$, a system of fundamental units (all of norm 1) is $\epsilon_1, \ldots, \epsilon_6$, where the following table gives their coefficients in terms of the integral basis:

	1	θ	θ^2	$\frac{1}{3}\theta^3$	$\frac{1}{3}\theta^4$	$\frac{1}{9}\theta^5$	$\frac{1}{9}\theta^6$	$\frac{1}{27}\theta^7$	$\frac{1}{27}\theta^8$	$\frac{1}{81}\theta^9$	$\frac{1}{81}\theta^{10}$	$\frac{1}{243}\theta^{11}$	$\frac{1}{243}\theta^{12}$
ϵ_1	1	1	0	1	0	1	0	1	0	1	0	1	0
ϵ_2	−5	4	1	4	2	−2	4	5	1	4	0	6	2
ϵ_3	25	4	−11	3	4	−19	−2	15	−4	−5	11	2	−5
ϵ_4	13	−6	0	14	−8	4	6	−12	4	2	−2	4	−1
ϵ_5	−53	71	−43	95	−63	101	−67	86	−49	53	−17	13	14
ϵ_6	409	−316	−178	238	206	−148	−221	52	224	40	−212	−130	191

and the rational prime 3 ramifies as $(3) = (\rho)^{13}$, where

$$\rho = 3 + 2\theta + \theta^2 + \frac{2}{3}\theta^3 + \frac{1}{3}\theta^4 + \frac{2}{9}\theta^5 + \frac{1}{9}\theta^6 + \frac{2}{27}\theta^7 + \frac{1}{27}\theta^8 + \frac{2}{81}\theta^9 + \frac{1}{81}\theta^{10} + \frac{2}{243}\theta^{11} + \frac{1}{243}\theta^{12} = \frac{3}{\theta^2 - 3}$$

is a prime of norm 3 (in fact, the only one up to multiplication by units).

3.3. **Upper bounds.** Each of our five fields $\mathbb{K}$ has one real embedding and $2r = n - 1$ non-real embeddings. Here r is the unit rank of the field $\mathbb{K}$. We number the conjugates as follows:

$$\begin{array}{rcll} \theta_1 & \in & \mathbb{R}, & \\ \theta_j & = & \theta_1\, e^{2\pi i j/n} & \text{for} \quad j = 2, 3, \ldots, r+1, \\ \theta_j & = & \bar{\theta}_{j-r} & \text{for} \quad j = r+2, r+3, \ldots, n \end{array}$$

and correspondingly for ϵ_i, ρ, etc. Here the bar denotes complex conjugation. In the cases with $a = 2$ we put $\phi = \theta$ and in the cases with $a = 3$ we put $\phi = \theta^2$, so that in all six cases $N_{\mathbb{K}/\mathbb{Q}}(\phi) = a^{n-1}(a+1)$.

For a solution $x, y \in \mathbb{Z}$ of equation (10) we write

$$\beta = ax - \phi y.$$

Then by (10) we have

$$N_{\mathbb{K}/\mathbb{Q}}(\beta) = (ax)^n - \left(N_{\mathbb{K}/\mathbb{Q}}(\phi)\right) y^n = a^{n-1}\left(ax^n - (a+1)y^n\right) = a^{n-1}, \tag{11}$$

hence there are $n_1, n_2, \ldots, n_r \in \mathbb{Z}$ such that

$$\beta = ax - \phi y = \rho^{n-1}\epsilon_1^{n_1}\epsilon_2^{n_2}\cdots\epsilon_r^{n_r}. \tag{12}$$

(Note that, a priori, $\beta = -\rho^{n-1}\epsilon_1^{n_1}\epsilon_2^{n_2}\cdots\epsilon_r^{n_r}$ is also possible, but since all norms are positive, this case does not occur.)

By way of example, for the known solutions with $x = y = -1$ we have

$$\begin{array}{lllll}
\beta & = & -2+\theta & = & \rho^4\epsilon_1^{-2}\epsilon_2 & \text{in the case} & (n,a)=(5,2), \\
\beta & = & -2+\theta & = & \rho^6\epsilon_2^2\epsilon_3^{-1} & \text{in the case} & (n,a)=(7,2), \\
\beta & = & -2+\theta & = & \rho^{10}\epsilon_2^2\epsilon_3^2\epsilon_4^{-1}\epsilon_5^{-1} & \text{in the case} & (n,a)=(11,2), \\
\beta & = & -2+\theta & = & \rho^{12}\epsilon_1^{-1}\epsilon_2^5\epsilon_3^2\epsilon_4^{-2}\epsilon_5^{-1}\epsilon_6 & \text{in the case} & (n,a)=(13,2), \\
\beta & = & -3+\theta^2 & = & \rho^{12}\epsilon_1^{-6}\epsilon_2^{-2}\epsilon_3^{-2}\epsilon_4\epsilon_6^2 & \text{in the case} & (n,a)=(13,3).
\end{array}$$

We start with showing that if $|y|$ is large enough, then $|\beta_1|$ is extremely small, whereas its conjugates $|\beta_j|$ with $j = 2, 3, \dots, n$ are relatively large. Put

$$b_j = |\mathrm{Im}\,\phi_j| \quad \text{for} \quad j = 2, 3, \dots, n, \qquad c_1 = a^{n-1} \Big/ \prod_{j=2}^{n} b_j \, .$$

Lemma 3.2. *We have*

$$\begin{array}{lll}
|\beta_1| & \leq & c_1|y|^{-(n-1)}, \\
|\beta_j| & \geq & b_j|y| \qquad \textit{for} \quad j = 2, 3, \dots, n.
\end{array}$$

Proof of Lemma 3.2. For $j = 2, 3, \dots, n$ we have

$$|\beta_j| = |ax - \phi_j y| \geq |\mathrm{Im}\,(ax - \phi_j y)| = |\mathrm{Im}\,\phi_j|\,|y| = b_j|y|.$$

Equation (11) now at once leads to

$$|\beta_1| = a^{n-1} \Big/ \prod_{j=2}^{n} |\beta_j| \leq a^{n-1} \Big/ \prod_{j=2}^{n} b_j|y| = c_1|y|^{-(n-1)}. \quad \Box$$

Notice that $b_j = b_{j-r}$ for $j = r+2, r+3, \dots, n$, because the jth and the $(j+r)$th conjugates are each others complex conjugates. It follows that we are only interested in the b_j for $j = 2, 3, \dots, r+1$.

From the equation $\beta = ax - \phi y$ we now take three conjugates, the real one and two complex conjugated ones, and eliminate x and y from these three equations. For each $j = 2, 3, \dots, r+1$ we thus derive the so-called Siegel identity

$$(\phi_j - \bar{\phi}_j)\beta_1 + (\bar{\phi}_j - \phi_1)\beta_j + (\phi_1 - \phi_j)\bar{\beta}_j = 0,$$

which we write as

$$\frac{\phi_1 - \bar{\phi}_j}{\phi_1 - \phi_j}\frac{\beta_j}{\bar{\beta}_j} - 1 = \frac{\phi_j - \bar{\phi}_j}{\phi_1 - \phi_j}\frac{\beta_1}{\bar{\beta}_j}. \tag{13}$$

Lemma 3.2 implies that the right hand side of this equation is extremely small in absolute value. Notice that $\frac{\phi_1 - \bar{\phi}_j}{\phi_1 - \phi_j}\frac{\beta_j}{\bar{\beta}_j}$ is on the unit circle, so if we put

$$\Lambda_j = -i\mathrm{Log}\,\frac{\phi_1 - \bar{\phi}_j}{\phi_1 - \phi_j}\frac{\beta_j}{\bar{\beta}_j},$$

then $\Lambda_j \in \mathbb{R}$. Here Log denotes the principal branch of the complex logarithm, with imaginary part in $(-\pi, \pi]$. We find that

$$e^{i\Lambda_j} - 1 = \frac{\phi_1 - \bar{\phi}_j}{\phi_1 - \phi_j}\frac{\beta_j}{\bar{\beta}_j} - 1$$

is extremely close to zero by Lemma 3.2, hence so is Λ_j for all $j = 2, 3, \ldots, r+1$. To be precise, for $j = 2, 3, \ldots, r+1$ put

$$d_j = 2^{n+1} \arcsin\left(2^{-(n+1)} \left|\frac{\phi_j - \bar{\phi}_j}{\phi_1 - \phi_j}\right| \frac{c_1}{b_j}\right).$$

Then we have the following result.

Lemma 3.3. *If $|y| \geq 2$ then for $j = 2, 3, \ldots, r+1$ we have*

$$|\Lambda_j| < d_j |y|^{-n}.$$

Proof of Lemma 3.3. Put $\delta_j = \left|\frac{\phi_j - \bar{\phi}_j}{\phi_1 - \phi_j}\right| \frac{c_1}{b_j}$. Lemma 3.2 and equation (13) yield $\left|e^{i\Lambda_j} - 1\right| < \delta_j |y|^{-n}$. By $|y| \geq 2$ we find $\left|e^{i\Lambda_j} - 1\right| < 2^{-n}\delta_j$, hence by [dW, Lemma 2.3] we find $|\Lambda_j| \leq \frac{d_j}{\delta_j}|e^{i\Lambda_j} - 1|$, and the result follows. □

We will now derive from equation (12) useful estimates relating $|y|$ to the exponents n_i. Define

$$U = \begin{pmatrix} \log|\epsilon_{1,2}| & \cdots & \log|\epsilon_{r,2}| \\ \vdots & \ddots & \vdots \\ \log|\epsilon_{1,r+1}| & \cdots & \log|\epsilon_{r,r+1}| \end{pmatrix},$$

which, as a matrix with determinant 2^{-r} times the regulator, is necessarily invertible. We obtain by (12) that

$$\begin{pmatrix} n_1 \\ \vdots \\ n_r \end{pmatrix} = U^{-1} \begin{pmatrix} \log|\beta_2/\rho_2^{n-1}| \\ \vdots \\ \log|\beta_{r+1}/\rho_{r+1}^{n-1}| \end{pmatrix}. \tag{14}$$

Let us further define

$$U^{-1} = \begin{pmatrix} u_{1,2} & \cdots & u_{1,r+1} \\ \vdots & \ddots & \vdots \\ u_{r,2} & \cdots & u_{r,r+1} \end{pmatrix},$$

and, for $k = 1, 2, \ldots, r$, set

$$\xi_k = \sum_{j=2}^{r+1} u_{k,j} \log\left|\frac{\phi_1 - \phi_j}{\rho_j^{n-1}}\right|,$$

$$\eta_k = \sum_{j=2}^{r+1} u_{k,j}.$$

An interesting observation of Bilu and Hanrot [BH] is that n_k is extremely close to $\xi_k + \eta_k \log|y|$ for $k = 1, 2, \ldots, r$. We make this precise in the following lemma. For $k = 1, 2, \ldots, r$ let

$$e_k = 2^n \left|\log\left(1 - 2^{-n} \frac{c_1}{\min_{j=2,3,\ldots,r+1} |\phi_1 - \phi_j|}\right)\right| \sum_{j=2}^{r+1} |u_{k,j}|.$$

Lemma 3.4. *If $|y| \geq 2$, then for $k = 1, 2, \ldots, r$ we have*

$$|n_k - (\xi_k + \eta_k \log|y|)| < e_k |y|^{-n}.$$

Proof of Lemma 3.4. Note that β_j is almost equal to $y(\phi_1 - \phi_j)$, namely

$$\beta_j = ax - \phi_j y = ax - \phi_1 y + y(\phi_1 - \phi_j) = \beta_1 + y(\phi_1 - \phi_j),$$

where β_1 is extremely small by Lemma 3.2. It follows that

$$\begin{aligned} |n_k - (\xi_k + \eta_k \log|y|)| &= \left| \sum_{j=2}^{r+1} u_{k,j} \log \left| \frac{\beta_j}{y(\phi_1 - \phi_j)} \right| \right| \\ &\leq \sum_{j=2}^{r+1} |u_{k,j}| \left| \log \left| 1 + \frac{\beta_1}{y(\phi_1 - \phi_j)} \right| \right|. \end{aligned}$$

Put $\delta = 2^{-n} \dfrac{c_1}{\min\limits_{j=2,3,\dots,r+1} |\phi_1 - \phi_j|}$. Then by Lemma 3.2 we have $\left|\frac{\beta_1}{y(\phi_1-\phi_j)}\right| \leq 2^n \delta |y|^{-n}$ and by $|y| \geq 2$ it follows that this is at most δ. By [dW, Lemma 2.2] we find $\left|\log\left|1 + \frac{\beta_1}{y(\phi_1-\phi_j)}\right|\right| \leq \frac{|\log(1-\delta)|}{\delta} \left|\frac{\beta_1}{y(\phi_1-\phi_j)}\right|$, and the result follows at once. □

We readily compute the following numerical values for the η_k and the ξ_k.

(n,a)	$(5,2)$	$(7,2)$	$(11,2)$	$(13,2)$	$(13,3)$
$\eta_1 =$	$-0.58282108\dots$	$-0.21603507\dots$	$0.15960187\dots$	$-0.41401237\dots$	$-1.3802655\dots$
$\eta_2 =$	$0.38913837\dots$	$0.52412775\dots$	$0.74643940\dots$	$1.1392655\dots$	$-0.52417382\dots$
$\eta_3 =$	$-$	$-0.31443098\dots$	$0.48137366\dots$	$0.49790031\dots$	$-0.44133730\dots$
$\eta_4 =$	$-$	$-$	$-0.097411241\dots$	$-0.23147615\dots$	$0.082129902\dots$
$\eta_5 =$	$-$	$-$	$-0.089777304\dots$	$-0.029527061\dots$	$0.23665553\dots$
$\eta_6 =$	$-$	$-$	$-$	$-0.0067848669\dots$	$0.45157277\dots$
$\xi_1 =$	$-2.0228582\dots$	$-0.0058037263\dots$	$0.0027937681\dots$	$-1.0062166\dots$	$-6.0149492\dots$
$\xi_2 =$	$1.0150570\dots$	$2.0146875\dots$	$2.0131235\dots$	$5.0171804\dots$	$-2.0056462\dots$
$\xi_3 =$	$-$	$-1.0087836\dots$	$2.0086302\dots$	$2.0075213\dots$	$-2.0047788\dots$
$\xi_4 =$	$-$	$-$	$-1.0018188\dots$	$-2.0035830\dots$	$1.0009323\dots$
$\xi_5 =$	$-$	$-$	$-1.0017135\dots$	$-1.0005334\dots$	$0.0025110131\dots$
$\xi_6 =$	$-$	$-$	$-$	$1.0002169\dots$	$2.0048679\dots$

Note the remarkable fact that all ξ_k are almost integers, and that these integers are just the exponents n_j corresponding to the known solutions with $x = y = -1$.

Define $N = \max\limits_{k=1,2,\dots,r} |n_k|$. An easy consequence of Lemma 3.4 is the following result, estimating $|y|$ in terms of N. Choose k_0 such that $|\eta_k|$ is maximal for $k = k_0$ and let

$$\begin{aligned} f_0 &= \tfrac{n}{|\eta_{k_0}|}, \\ g_0 &= \exp\left(n\left(\tfrac{\xi_{k_0}}{\eta_{k_0}} + \tfrac{2^{-n} e_{k_0}}{|\eta_{k_0}|}\right)\right). \end{aligned}$$

Further, define

$$\begin{aligned} Y_0 \;=\; \max\Bigg\{ 2, &\max_k \left\lceil \exp\left(-\frac{\xi_k}{\eta_k} + \frac{2^{-n} e_k - 1}{|\eta_k|} \right) \right\rceil, \\ &\max_{k \neq k_0} \left\lceil \exp\left(-\frac{\xi_{k_0} - s_k \xi_k}{\eta_{k_0} - s_k \eta_k} + \frac{2^{-n}(e_{k_0} + e_k) - 1}{|\eta_{k_0} - s_k \eta_k|} \right) \right\rceil \Bigg\}, \end{aligned}$$

where s_k is the sign of $\eta_{k_0}\eta_k$.

Lemma 3.5. *If $|y| \geq Y_0$, then we have*

$$\log|y| > \frac{f_0}{n} N - \frac{1}{n} \log g_0.$$

Proof of Lemma 3.5. Since $|y| \geq 2$, Lemma 3.4 yields $|n_k - (\xi_k + \eta_k \log|y|)| < 2^{-n} e_k$ for all k. Now the condition $\log|y| \geq -\frac{\xi_k}{\eta_k} + \frac{2^{-n}e_k - 1}{|\eta_k|}$ guarantees that n_k and η_k have the same sign, and the condition $\log|y| \geq -\frac{\xi_{k_0} - s_k \xi_k}{\eta_{k_0} - s_k \eta_k} + \frac{2^{-n}(e_{k_0} + e_k) - 1}{|\eta_{k_0} - s_k \eta_k|}$ guarantees that $|n_{k_0}| \geq |n_k|$, so that $N = |n_{k_0}|$. The result therefore follows easily from Lemma 3.4 applied with $k = k_0$. □

Note that in all our cases we found $Y_0 = 2$.

Now we can combine Lemmas 3.3 and 3.5, to find an upper bound for $|\Lambda_j|$ in terms of N. Put

$$f_j = d_j g_0 \quad \text{for} \quad j = 2, 3, \ldots, r+1.$$

Lemma 3.6. *If $|y| \geq Y_0$, then for $j = 2, 3, \ldots, r+1$ we have*

$$|\Lambda_j| < f_j \exp(-f_0 N).$$

Proof of Lemma 3.6. This is immediate from Lemmas 3.3 and 3.5. □

On the other hand, using equation (12) we can write Λ_j as a linear form in logarithms of algebraic numbers, viz.

$$\Lambda_j = -i \left(\operatorname{Log} \alpha_{0,j} + \sum_{k=1}^{r} n_k \operatorname{Log} \alpha_{k,j} + n_{0,j} \operatorname{Log}(-1) \right),$$

where

$$\alpha_{0,j} = \frac{\phi_1 - \bar{\phi}_j}{\phi_1 - \phi_j} \frac{\rho_j^{n-1}}{\bar{\rho}_j^{n-1}}, \qquad \alpha_{k,j} = \frac{\epsilon_{k,j}}{\bar{\epsilon}_{k,j}} \quad \text{for} \quad k = 1, 2, \ldots, r$$

and $n_{0,j}$ is an even integer, appearing because $\operatorname{Log} z_1 z_2 = \operatorname{Log} z_1 + \operatorname{Log} z_2$ holds only modulo $2\pi i$, and all $-i\operatorname{Log}$'s, including Λ_j itself, are in $(-\pi, \pi]$. Transcendence theory tells us that Λ_j cannot be too near to zero. Specifically, we apply the recent explicit and very sharp result of Baker and Wüstholz [BW]. The algebraic numbers $\alpha_{k,j}$ occurring inside the logarithms of the linear forms Λ_j are all in the field $\mathbb{Q}(\theta_1, \theta_j, \bar{\theta}_j)$, which is of degree at most (in fact, in our cases, equal to) $d = n(n-1)(n-2)$. The number of terms in the linear forms is $r+2$, at least a priori. We can however win a little bit here, by noting that in fact there is a multiplicative relation between, on the one hand, $\alpha_{0,j}$ and, on the other hand, $\alpha_{1,j}, \alpha_{2,j}, \ldots, \alpha_{r,j}$ (we found this relation in the reduction step described below, when we observed that certain numbers occurring were very close to rational numbers with denominator $2n$). In fact, we found that in all three cases

$$\alpha_{0,j}^n = -\left(\prod_{k=1}^{r} \alpha_{k,j}^{\nu_k} \right)^{n-1},$$

where ν_k is as in the table below (here $\nu = \max_k |\nu_k|$).

(n,a)	$(5,2)$	$(7,2)$	$(11,2)$	$(13,2)$	$(13,3)$
$\nu_1 =$	2	0	0	1	6
$\nu_2 =$	-1	-2	-2	-5	2
$\nu_3 =$	–	1	-2	-2	2
$\nu_4 =$	–	–	1	2	-1
$\nu_5 =$	–	–	1	1	0
$\nu_6 =$	–	–	–	-1	-2
$\nu =$	2	2	2	5	6

Note the remarkable fact that these numbers ν_k are exactly the negatives of the exponents n_j occurring for the known solution $x = y = -1$.

This shows that we can rewrite Λ_j as a form with one fewer term, considerably reducing the upper bound and permitting a somewhat simpler reduction procedure. Indeed, we write

$$n\Lambda_j = -i\left(\sum_{k=1}^{r} n'_k \mathrm{Log}\,\alpha_{k,j} + m'_j \mathrm{Log}\,(-1)\right),$$

where we have $n'_k = nn_k + (n-1)\nu_k$, and m'_j an (odd) integer. With $N' = \max_{k=1,2,\ldots,r} |n'_k|$ we have $N' \leq nN + (n-1)\nu$. Further we have to estimate $|m'_j|$ in terms of N. Note that

$$|m'_j| = \frac{1}{\pi}\left|in\Lambda_j - \sum_{k=1}^{r} n'_k \mathrm{Log}\,\alpha_{k,j}\right| \leq n + rN' \leq nrN + (n-1)r\nu + n.$$

As a result we have linear forms with only $m = r+1$ terms.

The result of Baker and Wüstholz, in our situation, implies the inequality

$$|\Lambda_j| > \exp\left(-C' \log\max\{N', |m'_j|\}\right), \tag{15}$$

where

$$C' = C(m,d)\prod_{k=0}^{r} h'(\alpha_{k,j})h'(-1)$$

for h' a certain height function and

$$C(m,d) = 18(m+1)!m^{m+1}(32d)^{m+2}\log(2md).$$

It's mainly these numbers that determine the sizes of the upper bounds to be derived.

We need to compute upper bounds for the heights of the algebraic numbers. We note that in our cases the height function $h'(\alpha)$ used by Baker and Wüstholz for our $\alpha_{k,j}$'s happens to coincide with the absolute logarithmic Weil height $h(\alpha)$. For an algebraic integer α it is given by

$$h(\alpha) = \frac{1}{[\mathbb{Q}(\alpha):\mathbb{Q}]}\log\prod_{\sigma}\max\{1, |\sigma(\alpha)|\},$$

where σ runs over the embeddings of $\mathbb{Q}(\alpha)$ into $\mathbb{C}$, and for a quotient of algebraic integers α/β the logarithmic Weil height can be estimated by

$$h(\alpha/\beta) \leq h(\alpha) + h(\beta).$$

In this way we found upper bounds for $h'(\alpha_{k,j})$ (that are obviously independent from j) and C' (note that $h(-1) = 0$, but $h'(-1) = \frac{1}{d}\pi$).

Now we can prove the main result of this section. Define

$$N_0 = \begin{cases} 1.7681627 \times 10^{22} & \text{in the case} \quad (n,a) = (5,2), \\ 3.4856031 \times 10^{30} & \text{in the case} \quad (n,a) = (7,2), \\ 1.4191886 \times 10^{48} & \text{in the case} \quad (n,a) = (11,2), \\ 4.1035085 \times 10^{57} & \text{in the case} \quad (n,a) = (13,2), \\ 8.6956453 \times 10^{57} & \text{in the case} \quad (n,a) = (13,3). \end{cases}$$

Lemma 3.7. *We have $|y| \leq 1$ or $N < N_0$.*

Proof of Lemma 3.7. Combining inequality (15) and Lemma 3.6, we find

$$N < \frac{1}{f_0} \log \min_j f_j + \frac{C'}{f_0} \log\left(nrN + r(n-1)\nu + n\right),$$

from which we derive at once the absolute upper bounds N_0 for N given above, by working out all the constants $b_j, c_1, d_j, e_j, k_0, f_0, g_0, f_j, h'(\alpha_{k,j}), C'$ for $j = 2, 3, \ldots, r+1$ and $k = 1, 2, \ldots, r$. □

3.4. **Reduction of the upper bounds: classical method.** In view of Lemma 3.7 there remains only a finite computation to complete the proof of Theorem 3.1. We describe four different techniques to carry out this task and compare their efficiency. Three of these methods have already been described in the literature and the fourth one, the only one that we'll present in full detail, is a variant combining ideas of the other methods.

The classical method, described by Tzanakis and de Weger [TW], tries to solve the following problem. For a given $j \in \{2, 3, \ldots, r+1\}$, and for $k = 1, 2, \ldots, r$ write

$$\phi_{k,j} = -i \operatorname{Log} \alpha_{k,j}.$$

These are known real numbers, that can be computed to the desired accuracy. We now have

$$\begin{cases} n\Lambda_j &= \displaystyle\sum_{k=1}^{r} n'_k \phi_{k,j} + m'_j \pi, \\ N' &= \displaystyle\max_{k=1,2,\ldots,r} |n'_k|, \end{cases}$$

and we want to determine the solutions $n'_1, n'_2, \ldots, n'_r, m'_j$ in $\mathbb{Z}$ of

$$\begin{cases} |\Lambda_j| &< f_j \exp(-f_0 N), \\ N &< N_0, \\ N' &\leq nN + (n-1)\nu, \\ |m'_j| &\leq nrN + (n-1)r\nu + n. \end{cases} \tag{16}$$

Here the constants f_0, f_j and N_0 are given above in Lemmas 3.5, 3.6 and 3.7.

The classical approach, as outlined in [TW], makes use of only one linear form at a time, i.e. a Λ_j for only one j. As Λ_j is a homogeneous linear form with $r+1$ terms, we define a lattice Γ of dimension $r+1$, consisting of the $\mathbb{Z}$-linear combinations of the columns of the matrix

$$\mathcal{A} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ \Phi_{1,j} & \Phi_{2,j} & \cdots & \Phi_{r,j} & \Psi \end{pmatrix}.$$

Here we take

$$\Phi_{k,j} = [C\phi_{k,j}], \qquad \Psi = [C\pi],$$

where C is a large positive number, somewhat larger than N_0^{r+1}, and $[\cdot]$ means rounding to an integer. This means that we have to compute the numbers $\phi_{k,j}$ and π to somewhat more than $(r+1)\log N_0/\log 10$ decimal digits.

By the LLL-algorithm [LLL] we can compute reduced bases of the lattices. In practice we use the functions `lllintpartial` and `lllint` of Pari-1.39.03. A reduced basis enables us to find a lower bound for the length of any vector pointing to a non-zero lattice point. By [LLL], this lower bound, denoted by $d(\Gamma)$, is given by $2^{-r/2}|\mathbf{b}_1|$, where $\mathbf{b}_1$ is the first basis vector of the reduced basis. Heuristically reasoning we expect that this length is of order $(\det\Gamma)^{1/\dim\Gamma}$, which, by our choice of C, should be somewhat larger than N_0. If this is not the case, the parameter C should be enlarged a bit.

For a solution $(n_1', n_2', \ldots, n_r', m_j')$ of (16) we put $\mathbf{x} = (n_1', n_2', \ldots, n_r', m_j')^\top$, and look at the lattice point $\mathcal{A}\mathbf{x}$, since it can be expected to be near the origin. Namely, if we put

$$\lambda_j = \sum_{k=1}^{r} n_k'\Phi_{k,j} + m_j'\Psi,$$

which is the last coordinate of $\mathcal{A}\mathbf{x}$, then λ_j is approximately $Cn\Lambda_j$, with a rounding error of the size of N_0. On the other hand, this number λ_j cannot be very small, since we have an upper bound N_0 for the other coordinates of $\mathcal{A}\mathbf{x}$, and a lower bound $d(\Gamma)$ for its length, which is a bit larger than N_0. So we obtain an explicit lower bound for $|\Lambda_j|$, of the size of N_0/C, and thus by Lemma 3.6 a reduced upper bound N_1 for N, that one expects to be of order size of $\frac{1}{f_0}\log C/N_0 \approx r\frac{1}{f_0}\log N_0$.

Subsequent reduction steps can be made, with N_0 replaced by N_1, and C adapted accordingly.

3.5. **Reduction of the upper bounds: using more linear forms simultaneously.** The main technique of the paper [MW], which was inspired by ideas of Yu. Bilu, is to use the linear forms Λ_j for $j = 2, 3, \ldots r+1$ simultaneously to solve (14). This works as follows.

We now define a lattice Γ of dimension $2r$, incorporating all linear forms, namely consisting of the $\mathbb{Z}$-linear combinations of the columns of the matrix

$$\mathcal{A} = \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ \Phi_{1,2} & \cdots & \Phi_{r,2} & \Psi & \cdots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ \Phi_{1,r+1} & \cdots & \Phi_{r,r+1} & 0 & \cdots & \Psi \end{pmatrix}.$$

Here we take $\Phi_{k,j}$ and Ψ as above, but now it suffices to take a smaller value for C, namely somewhat larger than N_0^2. The reason is that $(\det \Gamma)^{1/\dim \Gamma}$ is of the same order as the expected minimal length of any non-zero vector pointing to a lattice point, and we wish this to be essentially the same size as the upper bound N_0. So this means that we have to compute the numbers $\phi_{k,j}$ and π this time only to somewhat more than $2 \log N_0 / \log 10$ decimal digits. Of course we pay a price, namely we now have to do the LLL-algorithm for a lattice of dimension $2r$ instead of $r+1$.

For a solution $(n_1', n_2', \ldots, n_r', m_2', m_3', \ldots, m_{r+1}')$ of (16) we put

$$\mathbf{x} = (n_1', n_2', \ldots, n_r', m_2', m_3', \ldots, m_{r+1}')^\top,$$

and look at the lattice point $\mathcal{A}\mathbf{x}$, since it can be expected to be near to the origin. Namely, with

$$\lambda_j = \sum_{k=1}^{r} n_k' \Phi_{k,j} + m_j' \Psi,$$

we have, just as before, that λ_j is approximately $Cn\Lambda_j$ and cannot be particularly small. As above we obtain a reduced upper bound N_1 for N that one expects to be of the size of $\frac{1}{f_0} \log C/N_0$, which this time is $\approx \frac{1}{f_0} \log N_0$.

3.6. **Reduction of the upper bounds: using only two-dimensional lattices.** There is a third method of reducing the bound, introduced by Bilu and Hanrot in their ground-breaking paper [BH]. In their paper they give two examples of Thue equations of very high degree, with one of the corresponding algebraic number fields being totally real, and the other one, like in our situation, has only one real embedding and all other embeddings are non-real. Let's see how their method works in this setting.

Bilu and Hanrot do not work with the linear forms Λ_j, but rather start directly from Lemma 3.4. Note that the linear forms Λ_j have not completely left the scene: they still have been used to derive the upper bounds N_0. They do not, however, figure in the reduction step.

We now take $k_1, k_2 \in \{1, 2, \ldots, r\}$ with $k_1 \neq k_2$. Eliminating $\log|y|$ from the two inequalities given in Lemma 3.4

$$\begin{aligned} |n_{k_1} - (\xi_{k_1} + \eta_{k_1} \log|y|)| &\leq e_{k_1} |y|^{-n}, \\ |n_{k_2} - (\xi_{k_2} + \eta_{k_2} \log|y|)| &\leq e_{k_2} |y|^{-n} \end{aligned}$$

and setting

$$\gamma_{k_1,k_2} = \xi_{k_1}\eta_{k_2} - \xi_{k_2}\eta_{k_1}, \qquad g_{k_1,k_2} = e_{k_1}|\eta_{k_2}| + e_{k_2}|\eta_{k_1}|$$

and

$$\Lambda_{k_1,k_2} = \eta_{k_2} n_{k_1} - \eta_{k_1} n_{k_2} - \gamma_{k_1,k_2},$$

we thus have

$$|\Lambda_{k_1,k_2}| < g_{k_1,k_2} |y|^{-n}. \tag{17}$$

Now we can apply the good old Baker-Davenport Lemma [BD], or, equivalently, lattice base reduction in 2-dimensional lattices. This means that we do not have to use the full power of the LLL-algorithm, but only the simple euclidean algorithm, i.e. continued fraction expansions. For convenience we still use the Pari routine `lllint`, which in this 2-dimensional case is equivalent to the euclidean algorithm. Notice that the nice thing here is that the unknowns m'_j play no role at all.

We define a 2-dimensional lattice Γ, consisting of the $\mathbb{Z}$-linear combinations of the columns of the matrix

$$\mathcal{A} = \begin{pmatrix} 1 & 0 \\ H_{k_2} & H_{k_1} \end{pmatrix}$$

and we define the point

$$\mathbf{y} = \begin{pmatrix} 0 \\ K_{k_1,k_2} \end{pmatrix}$$

where we take $H_k = [C\eta_k]$ for $k = k_1, k_2$, and $K_{k_1,k_2} = [C\gamma_{k_1,k_2}]$. It suffices to take C somewhat larger than N_0^2 to make sure that the distances in the lattice are of the size of the upper bound N_0.

By the euclidean algorithm we can compute reduced bases of the lattices for the case $(k_1, k_2) = (1, 2)$. From these bases it is easy to find the distance $d(\Gamma, \mathbf{y})$ between the point $\mathbf{y}$ and the nearest lattice point. For a solution (n_{k_1}, n_{k_2}) of (17) we put $\mathbf{x} = (n_{k_1}, -n_{k_2})^\top$ and consider the lattice point $\mathcal{A}\mathbf{x}$ which we expect to be near $\mathbf{y}$. As in the previous two subsections we obtain a reduced upper bound N_1 for N that one expects to be of the size of $\frac{1}{f_0} \log C/N_0$, this time again $\approx \frac{1}{f_0} \log N_0$.

3.7. **The fourth variant.** We finish with a new variant, based on combining the ideas of the second and third methods, of Mignotte and de Weger [MW] and Bilu and Hanrot [BH]. This time we make use of as many inequalities of the type (17) as possible. Namely, we take $k_1 = 1$ fixed, and let k_2 run through $\{2, 3, \ldots, r\}$. So this is better than the original method of Bilu and Hanrot only if $r \geq 3$, i.e. in our cases only if $n \geq 7$. For this variant we will present full details for our five Thue equations (10).

This time we define a lattice Γ of dimension r, consisting of the $\mathbb{Z}$-linear combinations of the columns of the matrix

$$\mathcal{A} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ H_2 & H_1 & 0 & \cdots & 0 \\ H_3 & 0 & H_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ H_r & 0 & 0 & \cdots & H_1 \end{pmatrix}$$

and we define the point

$$\mathbf{y} = \begin{pmatrix} 0 \\ K_{1,2} \\ \vdots \\ K_{1,r} \end{pmatrix}$$

where we take H_k and K_{k_1,k_2} as above. Now it suffices to take C somewhat larger than $N_0^{1+1/(r-1)}$ only, to make sure that the distances in the lattice are of the size of the upper bound N_0. So with this variant the needed precision is smallest.

In practice we took C as follows:

$$\begin{aligned} C &= 10^{50} \quad \text{in the cases with} \quad n = 5, 7, \\ C &= 10^{65} \quad \text{in the case with} \quad n = 11, \\ C &= 10^{72} \quad \text{in the cases with} \quad n = 13. \end{aligned}$$

We give the values of $f_0, g_0, g_{1,k}$ for $k = 2, 3, \dots, r$ below.

(n,a)	$(5,2)$	$(7,2)$	$(11,2)$	$(13,2)$	$(13,3)$
$f_0 >$	8.5789620	13.355522	14.736628	11.410859	9.4184775
$g_0 <$	3.9739194×10^7	5.8915406×10^{11}	1.0334327×10^{13}	1.0394378×10^{25}	4.0214207×10^{24}
$g_{1,2} <$	0.41752317	1.0391846	19.907771	209.81907	1.8561913
$g_{1,3} <$	–	0.62342017	12.838385	91.698532	1.5454583
$g_{1,4} <$	–	–	3.7738454	47.821515	0.54077517
$g_{1,5} <$	–	–	3.2535583	20.563430	0.84497512
$g_{1,6} <$	–	–	–	14.006386	1.5813005

By the LLL-algorithm (again using `lllintpartial` and `lllint`) we compute reduced bases for these lattices. From these bases it is easy to find the distance $d(\Gamma, \mathbf{y})$ between the point $\mathbf{y}$ and the nearest lattice point. Here are the results of our computations.

$$d(\Gamma, \mathbf{y}) \geq \begin{cases} 8.1456649 \times 10^{24} & \text{in the case} \quad (n,a) = (5,2), \\ 5.3812421 \times 10^{32} & \text{in the case} \quad (n,a) = (7,2), \\ 1.1134613 \times 10^{51} & \text{in the case} \quad (n,a) = (11,2), \\ 3.2622236 \times 10^{60} & \text{in the case} \quad (n,a) = (13,2), \\ 5.4466343 \times 10^{59} & \text{in the case} \quad (n,a) = (13,3). \end{cases}$$

The computation times for the LLL-algorithm were

$$\begin{cases} \ll 1 \text{ sec.} & \text{in the case} \quad (n,a) = (5,2), \\ \ll 1 \text{ sec.} & \text{in the case} \quad (n,a) = (7,2), \\ 3.02 \text{ sec.} & \text{in the case} \quad (n,a) = (11,2), \\ 7.97 \text{ sec.} & \text{in the case} \quad (n,a) = (13,2), \\ 7.41 \text{ sec.} & \text{in the case} \quad (n,a) = (13,3). \end{cases}$$

For a solution $(n_1, n_2, \dots, n_r)$ of the inequalities (17) we put

$$\mathbf{x} = (n_1, -n_2, -n_3, \dots, -n_r)^\top$$

and consider the lattice point $\mathcal{A}\mathbf{x}$, which one anticipates will be near $\mathbf{y}$. Setting

$$\lambda_{1,k} = n_1 H_k - n_k H_1 - K_{1,k},$$

we have

$$|\lambda_{1,k} - C\Lambda_{1,k}| \leq 1 + 2N_0.$$

On the other hand,

$$d(\Gamma, \mathbf{y})^2 \leq |\mathcal{A}\mathbf{x} - \mathbf{y}|^2 = n_1^2 + \sum_{k=2}^{r} \lambda_{1,k}^2 \leq N_0^2 + (r-1)\left(\max_{k=2,3,\dots,r} |\lambda_{1,k}|\right)^2,$$

so we obtain

$$|\Lambda_{1,k}| > \frac{1}{C}\left(\sqrt{\frac{1}{r-1}d(\Gamma,\mathbf{y})^2 - \frac{1}{r-1}N_0^2} - (1+2N_0)\right).$$

Combined with (17) and Lemma 3.5 we thus find a reduced upper bound N_1 for N, viz.

$$N_1 = \left\lfloor \frac{1}{f_0} \log \frac{C g_0 \max_{k=2,3,\dots,r} g_{1,k}}{\sqrt{\frac{1}{r-1}d(\Gamma,\mathbf{y})^2 - \frac{1}{r-1}N_0^2} - (1+2N_0)} \right\rfloor.$$

This can in general, once again, be expected to be roughly $\frac{1}{f_0}\log C/N_0$, which now is $\approx \frac{1}{r-1}\frac{1}{f_0}\log N_0$. This is better than in any of the other three methods. In practice we find

$$N_1 = \begin{cases} 8 & \text{in the case} \quad (n,a) = (5,2), \\ 5 & \text{in the case} \quad (n,a) = (7,2), \\ 4 & \text{in the case} \quad (n,a) = (11,2), \\ 8 & \text{in the case} \quad (n,a) = (13,2), \\ 9 & \text{in the case} \quad (n,a) = (13,3). \end{cases}$$

3.8. **Comparing the variants.** Below we give for the case $(n,a) = (13,3)$, with $N_0 = 8.6956453 \times 10^{57}$ and $r = 6$, for the four reduction methods described above, the parameter C (controlling the size of the numbers to be dealt with), the reduced upper bound N_1 reached in one reduction step, the dimension of the lattice and the computation time used by Pari-1.39.03 on a Pentium 75 personal computer for computing the reduced lattice bases.

method	dim	theory		practice		reduction time
		$C \approx$	$N_1 \approx$	$C =$	$N_1 =$	
Tzanakis & dW	$r+1$	N_0^{r+1}	$r\frac{1}{f_0}\log N_0$	10^{430}	96	15.20 sec.
Mignotte & dW	$2r$	N_0^2	$\frac{1}{f_0}\log N_0$	10^{125}	22	14 min. 41.06 sec.
Bilu & Hanrot	2	N_0^2	$\frac{1}{f_0}\log N_0$	10^{125}	21	$\ll 1$ sec.
fourth variant	r	$N_0^{1+\frac{1}{r-1}}$	$\frac{1}{r-1}\frac{1}{f_0}\log N_0$	10^{72}	9	7.41 sec.

So the fourth variant needs the smallest precision and obtains the best result in one reduction step, whereas the Bilu-Hanrot method shows the fastest lattice basis reduction. We believe that this is typical.

3.9. **Finding the small solutions.** This is trivial. A good idea is to use Lemma 3.4, because it leaves really only small ranges to check, so that it is by no means necessary to try all the possible r-tuples $(n_1, n_2, \dots, n_r)$ with $N \le N_1$. In this way we find that in all three cases there are no solutions with $|y| \ge 2$. This completes the proof of Theorem 3.1.

4. Linear forms in two logarithms

To show that equation (5) has precisely one positive solution for "large" n, we will refer to the following result of Laurent, Mignotte and Nesterenko [LMN] (where, as noted in [Mi], the conditions upon a_1 and a_2 may be relaxed to those stated here).

Theorem 4.1. *Let α_1 and α_2 be two positive real algebraic numbers. Consider*

$$\Lambda = b_2 \log \alpha_2 - b_1 \log \alpha_1,$$

where b_1 and b_2 are positive rational integers. Put $D = [\mathbb{Q}(\alpha_1, \alpha_2) : \mathbb{Q}]$ and suppose that $\log \alpha_1$ and $\log \alpha_2$ are linearly independent over $\mathbb{Q}$. For any $\rho > 1$, take

$$\begin{aligned} h &\ge \max\left\{ \frac{D}{2}, 5\lambda, D\left(\log\left(\frac{b_1}{a_2} + \frac{b_2}{a_1} \right) + \log \lambda + 1.56 \right) \right\}, \\ a_i &\ge (\rho - 1)\left|\log \alpha_i\right| + 2Dh(\alpha_i) \qquad (i = 1, 2), \\ a_1 + a_2 &\ge 4 \max\{1, \lambda\}, \\ \frac{1}{a_1} + \frac{1}{a_2} &\le \min\{1, \lambda^{-1}\}, \end{aligned}$$

where $\lambda = \log \rho$. Then

$$\begin{aligned} \log|\Lambda| \ \ge\ & -\frac{a_1 a_2}{9\lambda} A^2 - \frac{2}{3}(a_1 + a_2) A - \frac{16}{3}\sqrt{2 a_1 a_2} B^{3/2} \\ & - \log\left(\frac{a_1 a_2}{\lambda} B^2 \right) - \frac{3}{2}\lambda - 2h - \frac{3}{20} \end{aligned}$$

where

$$A = \frac{4h}{\lambda} + 4 + \frac{\lambda}{h} \qquad \textit{and} \qquad B = 1 + \frac{h}{\lambda}.$$

Here, the height function $h(\alpha)$ is as defined in Section 3.

Suppose that a and n are positive integers ($a \ge 2$ and $n \ge 3$) for which equation (5) possesses a positive solution $(x, y) \ne (1, 1)$. By Lemmas 2.4 and 2.6 we may assume that

$$a < 0.3n \qquad \text{and} \qquad \min\{x, y\} > an. \tag{18}$$

Further, to prove Theorem 1.1 for $2 \le a \le 83$, it suffices to consider $n \ge 349$ while, if $a \ge 84$, careful application of Theorem 2.3 (and hence inequality (8)) along with Lemma 2.6 allows us to restrict our attention to $n \ge 331$. This implies, additionally, that we have

$$x > \max\{(a+1)^2, 662\}. \tag{19}$$

We apply Theorem 4.1 with $\alpha_1 = x/y$, $\alpha_2 = 1 + 1/a$, $b_1 = n$, $b_2 = 1$, and (as in [Mi])

$$\rho = \begin{cases} 5.8 & \text{if} \quad a = 2, \\ 1 + \frac{\log(a+1)}{\log(1+1/a)} & \text{if} \quad a \ge 3. \end{cases}$$

We therefore have

$$|\Lambda| = \left|\log\left(1+\frac{1}{a}\right) - n\log\left(\frac{x}{y}\right)\right| < \frac{1}{ax^n}, \tag{20}$$

so that

$$n\log\left(\frac{x}{y}\right) < \log\left(1+\frac{1}{a}\right) + \frac{1}{ax^n}.$$

Taking $\lambda = \log\rho$, we have, by calculus,

$$\log(a+1) < \lambda < 1.365\log(a+1), \tag{21}$$

which, in conjunction with (18) and the fact that $n \geq 331$, implies

$$(\lambda-1)\log(x/y) + 2\log x < 2.002\log x.$$

It follows that we may take

$$\begin{aligned} a_1 &= 2.002\log x, \\ a_2 &= \begin{cases} 4.15 & \text{if} \quad a=2, \\ 3\log(a+1) & \text{if} \quad a\geq 3 \end{cases} \end{aligned}$$

and verify that $a_1 + a_2 \geq 4\max\{1,\lambda\}$ and $a_1^{-1} + a_2^{-1} \leq \min\{1,\lambda^{-1}\}$ (using (19) and (21)). We further let

$$h = \max\{5\lambda, 1.2\log n\}$$

and note that this is justified, since the inequality

$$\log\left(\frac{n}{a_2}+\frac{1}{a_1}\right) + \log\lambda + 1.56 < 1.2\log n$$

is readily checked for $a=2$ (using (19)) and follows from (19) and (21) for larger a.

Suppose first that $h = 1.2\log n > 5\lambda$ so that we may apply (18), $n \geq 331$ and (21) to deduce the inequality

$$n > \max\left\{(a+1)^4, 1500\right\}. \tag{22}$$

If $a=2$, then (22) implies that we have (in the notation of Theorem 4.1) $A < 3.4\log n$ and $B < A/4$. Applying Theorem 4.1 yields

$$\begin{aligned} \log|\Lambda| \quad > \quad & -6.1\log^2 n\log x - 4.6\log n\log x - 18.6\log^{3/2} n\log^{1/2} x \\ & -11.9\log n - \log(\log^2 n\log x) - 4.2. \end{aligned}$$

It follows, then, from $x > 2n > 3000$ that

$$\log|\Lambda| > -\left(6.1\log^2 n + 23.2\log n + 13.3\right)\log x$$

which, since $n > 1500$, contradicts

$$\log|\Lambda| < -n\log x. \tag{23}$$

Similarly, if $a \geq 3$ and $h > 5\lambda$, we have $A < 5.81\log n/\lambda$, $B < 1.44\log n/\lambda$ and thus (21) implies that

$$\begin{aligned} \log|\Lambda| \quad > \quad & -22.6\lambda^{-2}\log^2 n\log x - 7.8\lambda^{-1}\log n\log x - 14.4\log n \\ & -32.0\lambda^{-1}\log^{3/2} n\log^{1/2} x - \log(\lambda^{-2}\log^2 n\log x) - 2.7. \end{aligned}$$

Since $\lambda > 1.76$, (18) and (22) therefore give

$$\log|\Lambda| > -\left(7.3\log^2 n + 22.7\log n + 15.4\right)\log x$$

which again contradicts $n > 1500$ and (23).

It follows that we may assume that $h = 5\lambda$ (so that $A = 24.2$ and $B = 6$) and apply Theorem 4.1 to deduce lower bounds upon Λ of the form

$$\log|\Lambda| > -c_1\log x - c_2\sqrt{\log x} - \log\log x - c_3 \tag{24}$$

for positive constants c_1, c_2 and c_3 which depend only on a (i.e. not upon n). Applying (21) allows us to take

$$c_1 = 423.2, \qquad c_2 = 271.7\sqrt{\log(a+1)} \qquad \text{and} \qquad c_3 = 64.1\log(a+1) + 5.6$$

which, together with (20), implies

$$\log|\Lambda| > -648.5\log x$$

so that from (23) we may conclude that $n < 649$. This, with (18), implies that $a \leq 194$ and, arguing more precisely with Theorem 2.3 and (8), we may in fact assume that $a \leq 166$.

We now show that (5) has no nontrivial solutions if $84 \leq a \leq 166$. Computing c_1, c_2 and c_3 in (24), we find that for each such a we have $c_1 \leq 328.68$, $c_2 \leq 614.59$ and $c_3 \leq 330.58$ (all obtained for $a = 166$). Combining (23) and (24), we find, for $n \geq 557$, that $x < 84n$, contradicting (18). Since there are no new solutions to (5) with $a \geq 84$ and $n < 331$, it follows that we need only consider prime values of n with $331 \leq n \leq 547$. Applying Theorem 2.3 and (8) for each such n, we may restrict our attention to a with $84 \leq a \leq a_0$ with a_0 given in the following table.

n	a_0	n	a_0	n	a_0	n	a_0	n	a_0
331	86	373	97	419	108	457	118	499	129
337	88	379	98	421	109	461	119	503	130
347	90	383	99	431	112	463	120	509	131
349	91	389	101	433	112	467	121	521	134
353	92	397	103	439	113	479	124	523	135
359	93	401	104	443	114	487	126	541	139
367	95	409	106	449	116	491	127	547	141

If $84 \leq a \leq 86$, we may take $c_1 = 325.2, c_2 = 574.2$ and $c_3 = 290.0$ in (24). Combining this with (18) and (23) allows us to deduce the following upper bounds upon a solution (x, y) to (5), for these values of a:

$n = 331$	$x < 10^{4302}$	$n = 353$	$x < 10^{195}$
$n = 337$	$x < 10^{1051}$	$n = 359$	$x < 10^{133}$
$n = 347$	$x < 10^{314}$	$367 \leq n \leq 523$	$x < 10^{88}$
$n = 349$	$x < 10^{264}$	$541 \leq n \leq 547$	$x < 84n$

Similar arguments produce like bounds for the remaining a with $84 \leq a \leq 141$. We remark that the values c_1, c_2 and c_3 increase monotonically with a in this range and satisfy $c_1 < 327.81$, $c_2 < 604.78$ and $c_3 < 320.49$. With this in mind, it is easy to see that the upper bounds obtained upon solutions to (5) for larger values of a are all rather smaller than the worst cases noted in the above table (i.e. of the form $x < 10^{1093}$ or better). To show that equation (5) has exactly the one positive solution $x = y = 1$ for $a \geq 84$, then, it remains to check beneath these upper bounds, through examination of the relevant continued fraction expansions. We once again apply Pari GP to compute the initial terms in the expansions to $\sqrt{1+\frac{1}{a}}$, for the pairs (a, n) under consideration. Inequalities (9) and (18) imply that a solution $(x, y) \neq (1, 1)$ to (5) yields a partial quotient a_j with

$$a_j \geq (an)^{n-1} > 10^{1466} \tag{25}$$

and we find, after checking at most 8467 partial quotients (corresponding to $(a, n) = (86, 331)$ where we have 8466 convergents with denominators smaller than 10^{4302}), that the largest one we encounter is $a_{34} = 581420$ (with $(a, n) = (84, 461)$).

The proof of Theorem 1.1 for $2 \leq a \leq 83$ proceeds along similar lines, only with more possibilities for the values of the exponent n (since we can no longer assume that n is prime). We again compute c_1, c_2 and c_3 in (24) for each choice of a and for $349 \leq n < 649$. Combining this with inequality (23) gives bounds upon solutions to (5) at least as strong as $x < 10^{553}$ (obtained for $(a, n) = (3, 349)$). Another convergent check yields no new solutions to (5) (since the largest partial quotient encountered is $a_{31} = 942288$ for $(a, n) = (60, 433)$, contradicting (9)).

5. Conclusions

As we have noted, the remaining cases of equation (2) with $b = a + 1, 2 \leq a \leq \min\{0.3n, 83\}$ and $17 \leq n \leq 347$ may be "effectively" treated by arguments analogous to those in Section 3. The main computational difficulty, at present, appears to be the problem of finding fundamental units in the corresponding number fields. Improvements in the hypergeometric method discussed in Section 2 or in the theory of linear forms in logarithms (Sections 3 and 4) might enable one to the complete the proof that (2) has at most one positive solution in all cases, though it is not unlikely that a fundamentally new idea will be required.

Acknowledgements

We are grateful to both Maurice Mignotte and the anonymous referee for their valuable remarks.

References

[Ba1] A. Baker, "Rational approximations to certain algebraic numbers", *Proc. London Math. Soc. (3)* **14** [1964], 385–398. MR **28:**5029

[Ba2] A. Baker, "Rational approximations to $\sqrt[3]{2}$ and other algebraic numbers", *Quart. J. Math. Oxford Ser. (2)* **15** [1964], 375–383. MR **30:**1977

[Ba3] A. Baker, "Contributions to the theory of diophantine equations I: On the representation of integers by binary forms", *Philos. Trans. Roy. Soc. London Ser. A* **263** [1968], 173–191. MR **37:**4005

[BD] A. Baker and H. Davenport, "The equations $3x^2 - 2 = y^2$ and $8x^2 - 7 = z^2$", *Quart. J. Math. (Oxford) (2)* **20** [1969], 129–137. MR **40:**1333

[BW] A. Baker and G. Wüstholz, "Logarithmic forms and group varieties", *J. Reine Angew. Math.* **442** [1993], 19–62. MR **94i:**11050

[Be1] M. Bennett, "Effective measures of irrationality for certain algebraic numbers", *J. Austral. Math. Soc.*, to appear.

[Be2] M. Bennett, "Explicit lower bounds for rational approximation to algebraic numbers", *Proc. London Math. Soc.*, to appear.

[BH] Yu. Bilu and G. Hanrot, "Solving Thue equations of high degree", J. Number Theory **60** [1996], 373–392. CMP 97:02

[Ch] G.V. Chudnovsky, "On the method of Thue-Siegel", *Ann. Math. II Ser.* **117** [1983], 325–382. MR **85g:**11058

[Co] H. Cohen, *Algorithmic Algebraic Number Theory*, Springer Verlag, Berlin, 1993.

[DM] H. Darmon and L. Merel, "Winding quotients and some variants of Fermat's Last Theorem", *J. Reine Angew. Math.*, to appear.

[De] B.N. Delone, "Solution of the indeterminate equation $X^3q + Y^3 = 1$", *Izv. Akad. Nauk SSR (6)* **16** [1922], 253–272.

[DF] B.N. Delone and D.K. Faddeev, *The Theory of Irrationalities of the Third Degree*, Translation of Math. Monographs, Vol. 10, Am. Math. Soc., 1964. MR **28:**3955

[Do] Y. Domar, "On the Diophantine equation $|Ax^n - By^n| = 1, n \geq 5$", *Math. Scand.* **2** [1954], 29–32. MR **16:**13c

[Ev1] J.-H. Evertse, *Upper Bounds for the Numbers of Solutions of Diophantine Equations*, PhD Thesis, Leiden, 1983.

[Ev2] J.-H. Everste, "On the equation $ax^n - by^n = c$", *Comp. Math.* **47** [1982], 289–315. MR **84d:**10022

[Ev3] J.-H. Evertse, "On the representation of integers by binary cubic forms of positive discriminant", *Invent. Math.* **73** [1983], 117–138. MR **85e:**11026a

[H] S. Hyrrö, "Über die Gleichung $ax^n - by^n = z$ und das Catalansche Problem", *Ann. Acad. Sci. Fenn. Ser. A I*, No. 355, 1964, 50 pp.

[LMN] M. Laurent, M. Mignotte and Y. Nesterenko, "Formes linéaires en deux logarithmes et déterminants d'interpolation", *J. Number Theory* **55** [1995], 285–321. MR **96h:**11073

[LLL] A.K. Lenstra, H.W. Lenstra jr. and L. Lovász, "Factoring polynomials with rational coefficients", *Math. Ann.* **261** [1982], 515–534. MR **84a:**12002

[Le] W.J. LeVeque, *Topics in Number Theory*, Addison-Wesley, Reading, Mass. 1956. MR **18:**283d

[Lj1] W. Ljunggren, "Einige Eigenschaften der Einheitenreeller quadratischer und rein biquadratischer Zahlkörper mit Anwendung auf die Lösung einer Klasse von bestimmter Gleichungen vierten Grades", *Det Norske Vidensk. Akad. Oslo Skrifter* **I**, No. 12 [1936], 1–73.

[Lj2] W. Ljunggren, "On an improvement of a theorem of T. Nagell concerning the Diophantine equation $Ax^3 + By^3 = C$", *Math. Scan.* **1** [1953], 297–309. MR **15:**400d

[Mi] M. Mignotte, "A note on the equation $ax^n - by^n = c$", Acta Arith. **75** [1996], 287–295. MR **97c:**11042

[MW] M. Mignotte and B.M.M. de Weger, "On the diophantine equations $x^2 + 74 = y^5$ and $x^2 + 86 = y^5$", *Glasgow Math. J.* **38** [1996], 77–85. MR **97b:**11044

[Mo] L.J. Mordell, *Diophantine Equations*, Academic Press, London, 1969. MR **40:**2600

[Mu] J. Mueller, "Counting solutions of $|ax^r - by^r| \leq h$", *Quart. J. Math. Oxford (2)* **38** [1987], 503–513. MR **89b:**11026

[MS] J. Mueller and W.M. Schmidt, "Thue's equation and a conjecture of Siegel", *Acta Math.* **160** [1988], 207–247. MR **89g:**11029

[N] T. Nagell, "Solution complète de quelques équations cubiques à deux indéterminées", *J. de Math. (9)* **4** [1925], 209–270.

[R] P. Ribenboim, *Catalan's Conjecture*, Academic Press, London, 1994. MR **95a:**11029

[Sh] T.N. Shorey, "Perfect powers in values of certain polynomials at integer points", *Math. Proc. Cambridge Philos. Soc.* **99** [1986], 195–207. MR **87c:**11032

[ST] T.N. Shorey and R. Tijdeman, "New applications of Diophantine approximations to Diophantine equations", *Math. Scand.* **39** [1976], 5–18. MR **56:**5425

[Si] C.L. Siegel, "Die Gleichung $ax^n - by^n = c$", *Math. Ann.* **144** [1937], 57–68.

[Ta] V. Tartakovskiĭ, "Auflösung der Gleichung $x^4 - \rho y^4 = 1$", *Izv. Akad. Nauk SSR (6)* **20** [1926], 301–324.

[Th] A. Thue, "Über Annäherungenswerte algebraischen Zahlen", *J. Reine Angew. Math.* **135** [1909], 284-305.

[Ti] R. Tijdeman, "Applications of the Gel'fond-Baker method to rational number theory", *Colloq. Math. Soc. János Bolyai* **13** [1976], 399–416. MR **58:**27753

[TW] N. Tzanakis and B.M.M. de Weger, "On the practical solution of the Thue equation", *J. Number Theory* **31** [1989], 99–132. MR **90c:**11018

[dW] B.M.M. de Weger, *Algorithms for diophantine equations*, CWI-tract no. 65, Centrum voor Wiskunde en Informatica, Amsterdam, 1989. MR **90m:**11205

Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48109
E-mail address: `mabennet@math.lsa.umich.edu`

Mathematical Institute, University of Leiden, Leiden, The Netherlands, and Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
E-mail address: `deweger@few.eur.nl`