JOURNAL OF NUMBER THEORY 31, 99–132 (1989)

On the Practical Solution of the Thue Equation

N. TZANAKIS

Department of Mathematics, University of Crete, Iraklion, Crete, Greece

B. M. M. DE WEGER

Faculty of Applied Mathematics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

Communicated by M. Waldschmidt

Received November 2, 1987

This paper gives in detail a practical general method for the explicit determination of all solutions of any Thue equation. It uses a combination of Baker's theory of linear forms in logarithms and recent computational diophantine approximation techniques. An elaborated example is presented. © 1989 Academic Press, Inc.

I. INTRODUCTION

In 1909 A. Thue [25] proved his famous result: If $F(X, Y) \in \mathbb{Z}[X, Y]$ is an irreducible binary form of degree at least 3, and m is a given nonzero rational integer, then the equation

$$F(X, Y) = m \tag{1}$$

has only finitely many integral solutions (X, Y). Thue's proof was ineffective. His method is described also in Sprindžuk's book [23, Chap. I, Sect. 2]. Other noneffective proofs of Thue's theorem based on Siegel's ideas can be found, for example, in Dickson [11, Chap. X] and Mordell [18, Chap. 23], and a (noneffective) proof based on the p-adic method of Skolem (with the restriction that the algebraic equation $F(x, 1) = 0$ have at least one pair of complex conjugate roots) can be found in Borevich and Shafarevich [7, Chap. 4, Sect. 6]. Prof. J. W. S. Cassels drew our attention to a paper of Cl. Chabauty (Démonstration nouvelle d'un théorème de Thue et Mahler sur les formes binaires, *Bull. Sci. Math.* (2) **65** (1941), 112–130, MR 7, 147b, and "Reviews in Number Theory," Vol. 2, D60-5), which also deals with the totally real case, based on the p-adic method of Skolem. We note here, as a side-remark, that for a class of Thue equations

0022-314X/89 $3.00

(especially for those related to a cubic or a half-real quartic field), Skolem's p-adic method provides a practical method for finding explicitly all solutions. However, in this paper we do not discuss at all such a method. We restrict ourselves to referring to Tzanakis [26], which includes useful references.

The first effective proof of Thue's theorem was given in 1968 by A. Baker [2], as a conseuquence of his deep study of linear forms in the logarithms of algebraic numbers. Baker's result yields an explicit upper bound for $\max(|X|, |Y|)$ for the solutions (X, Y) of (1). We refer to Shorey and Tijdeman [22, Chap. 5] for a survey of related results.

Baker's theory alone cannot provide us with a practical method for solving explicitly a particular equation (1), since the upper bound it yields is far too large. A combination with computational techniques from diophantine approximation theory makes this task possible. Thus, since Baker's work, a few papers have appeared in which all solutions of particular equations (1), mainly of degree 3, are explicitly found, making use of Baker's results. A typical example is the paper by Ellison *et al.* [13]. Other papers of the same flavour, which use an improvement of a Baker-type theorem due to Waldschmidt [30], are those by Steiner [24], Pethö and Schulenberg [19], and Blass *et al.* [6], the last two including also a discussion of the general equation (1) with $m = \pm 1$, and its solution in practice.

The present paper gives in detail a practical general method for the explicit determination of all solutions of any particular Thue equation (1). It uses a combination of Baker's theory and recent computational diophantine approximation techniques, that are applied for the first time to the solution of (1). A brief outline of our method can be found also in Tzanakis and de Weger [27 (see also de Weger [32])].

As one would expect, we work in the field $\mathbb{Q}(\xi)$, where $F(\xi, 1) = 0$. Then the computation of fundamental units in a convenient order of $\mathbb{Q}(\xi)$, as well as the factorization of m into prime ideals of this order, is needed. Such problems constitute in their general setting a whole area of current research. Therefore they are not discussed here. We merely suppose that we possess a set of fundamental units, and that we know the prime ideal factorization of m. Nevertheless, in an appendix we give in detail a method for computing a triple of fundamental units in any order of a totally real quartic field, due to Billevič, because in the examples that we give such a field is involved.

In Section I we discuss the general Thue equation (1). This section consists of three subsections. In Subsection 1 we reduce the solution of (1) to a finite number of inequalities of the form

$$|\Lambda| < \text{constant} \cdot |Y|^{-n}, \tag{2}$$

where n is the degree of $F(X, Y)$, and

$$\Lambda = \operatorname{Log} \delta + a_1 \cdot \operatorname{Log} \delta_1 + \cdots + a_q \cdot \operatorname{Log} \delta_q.$$

Here, $\delta, \delta_1, ..., \delta_q$ are explicitly given, in general complex, algebraic numbers, Log denotes the principal logarithm, and $a_1, ..., a_q$ are variables in $\mathbb{Z}$, in such a way that the determination of all integral solutions $a_1, ..., a_q$ of (2) implies the determination of all solutions (X, Y) of (1). The number q is equal to r or $r+1$, where r is the number of fundamental units in the field $\mathbb{Q}(\xi)$.

In Subsection 2 we pass from inequality (2) to an inequality

$$|\Lambda| < K_1 \cdot \exp(-K_2 \cdot A), \tag{3}$$

where $A = \max |a_i|$, and K_1, K_2 are explicitly known positive constants. Since, as we show, $\Lambda \neq 0$, we can apply Waldschmidt's theorem [30] to compute positive constants C_7, C_8 such that

$$|\Lambda| > \exp(-C_7 \cdot (\log A + C_8)). \tag{4}$$

Then (3) and (4) are combined to give a "very large" upper bound K_3 of A.

In Subsection 3 we discuss in full generality the problem of solving inequality (3) under the restriction $A < K_3$. We discuss in detail a process, based on Lovász' Lattice Basis Reduction Algorithm (cf. Lenstra *et al.* [16]), which reduces the upper bound of A to a new upper bound, which is of the size of the logarithm of the previous one. In the same section we compare our reduction process to others already used in the analogous problems.

In Section III we apply our general method to find all integral points on the elliptic curve

$$y^2 = x^3 - 4 \cdot x + 1. \tag{5}$$

Equation (5) arises naturally from the following problem of S. P. Mohanty: to find all triangular numbers $T_n = n \cdot (n+1)/2$ which equal a product of three consecutive integers. By solving (5) (see Theorem A of Section III) we find all such numbers T_n. There are six of them. The elliptic curve (5) is also interesting from the fact that it has rather many integral points, namely 22. The largest ones are $(x, \pm y) = (1274, 45473)$.

In Subsection 1 of Section III we reduce the solution of (5) to a number of quartic Thue equations. Only two of those have irreducible forms $F(X, Y)$, whereas the other ones are trivially solved in this section.

In Subsection 2 we solve the pair of irreducible quartic Thue equations (see Theorem B of Section III), according to the general method of Sec-

tion II. These equations are related to a totally real quartic field (the same field for both equations), so that three fundamental units are involved.

Section IV includes two appendixes. In Appendix I we state a theorem of Billevič [4] about the computation of a set of fundamental units in a totally real quartic field, and we apply it in the particular case of the quartic field appearing in Subsection 2 of Section III.

In Appendix II we state the previously mentioned theorem of Waldschmidt, in the form that we use it, and we discuss its application in practice. As a corollary we compute the constants C_7, C_8 which appear in inequality (4).

All computer calculations related to the reduction process applied in Subsection 2 of Section III were performed on an IBM 3083 computer at the University of Leiden. Most of them have been duplicated on an IBM 4361 computer at the University of Crete (by an independent package of programs). The computation of the fundamental units has been performed on the latter computer. Computational details cannot be included here. They can be found in the preprint version [28] of this paper.

II. The General Thue Equation

1. *From the Thue Equation to an Inequality Involving a Linear Form in Logarithms*

In this section we show how the solution of the general Thue equation implies an inequality involving a linear form in the logarithms of algebraic numbers with rational integral coefficients (unknowns). Let

$$F(X, Y) = \sum_{i=0}^{n} f_i \cdot X^{n-i} \cdot Y^i \in \mathbb{Z}[X, Y]$$

be a binary form of degree $n \geqslant 3$ and let m be a nonzero integer. Consider the Thue equation

$$F(X, Y) = m, \tag{1.1}$$

in the unknowns X, $Y \in \mathbb{Z}$. If F is reducible over $\mathbb{Q}$, then (1.1) can be reduced to a system of finitely many equations of type (1.1) with irreducible binary forms. For such equations of degree 1 or 2 it is well known how to determine the solutions. Therefore we may assume from now on that F is irreducible over $\mathbb{Q}$ and of degree $\geqslant 3$. Let $g(x) = F(x, 1)$. If $g(x) = 0$ has no real roots then one can trivially find small upper bounds for $\max(|X|, |Y|)$ for the solutions (X, Y) of (1.1) (see, e.g., [6]). Therefore, throughout this paper we suppose that the algebraic equation $g(x) = 0$ has

at least one real root. We number its roots as follows: $\xi^{(1)}, \ldots, \xi^{(s)}$ $(s \geqslant 1)$ are the real roots and $\xi^{(s+1)} = \overline{\xi^{(s+t+1)}}, \ldots, \xi^{(s+t)} = \overline{\xi^{(s+2t)}}$ are the nonreal roots, so that we have t $(\geqslant 0)$ pairs of complex conjugate roots, and $s + 2 \cdot t = n$.

Consider the field $K = \mathbb{Q}(\xi)$, where $g(\xi) = 0$. We will define three positive real numbers $Y_1 < Y_2 < Y_3$ that will divide the set of possible solutions (X, Y) of (1.1) into four classes:

(I) the "very small" solutions, with $|Y| \leqslant Y_1$. They will be found by enumeration of all possibilities.

(II) the "small" solutions, with $Y_1 < |Y| \leqslant Y_2$. They will be found by evaluating the continued fraction expansions of the real $\xi^{(i)}$'s.

(III) the "large" solutions, with $Y_2 < |Y| \leqslant Y_3$. They will be proved not to exist by a computational diophantine approximation technique.

(IV) the "very large" solutions, with $|Y| > Y_3$. They will be proved not to exist by the theory of linear forms in logarithms.

The value of Y_3 follows from the Gelfond–Baker theory of linear forms in logarithms. The value of Y_2 follows from the restrictions that we use as we try to prove that no "large" solutions exist. The value of Y_1 follows from Lemma 1.1 below. This lemma shows that if $|Y|$ is large enough then X/Y is "extremely close" to one of the real roots $\xi^{(i)}$. In a typical example Y_3 may be as large as $10^{10^{50}}$, Y_2 as large as 10^{10}, and Y_1 as small as 10.

LEMMA 1.1. *Let* $X, Y \in \mathbb{Z}$ *satisfy* (1.1). *Put* $\beta = X - \xi \cdot Y \in K$,

$$Y_0 = \begin{cases} \left\lceil \left(\dfrac{2^{n-1} \cdot |m|}{\min_{1 \leqslant i \leqslant t} |g'(\xi^{(s+i)})| \cdot \min_{1 \leqslant i \leqslant t} |\operatorname{Im} \xi^{(s+i)}|} \right)^{1/n} \right\rceil & \text{if } t \geqslant 1 \\ 1 & \text{if } t = 0, \end{cases}$$

$$C_1 = \frac{2^{n-1} \cdot |m|}{\min_{1 \leqslant i \leqslant s} |g'(\xi^{(i)})|}, \qquad C_2 = \frac{1}{2} \cdot \min_{1 \leqslant i < j \leqslant n} |\xi^{(i)} - \xi^{(j)}|,$$

$$Y_1 = \max(Y_0, \lceil (4 \cdot C_1)^{1/(n-2)} \rceil).$$

(i) *If* $|Y| > Y_0$ *then there exists an* $i_0 \in \{1, \ldots, s\}$ *such that*

$$|\beta^{(i_0)}| \leqslant C_1 \cdot |Y|^{-(n-1)},$$

$$|\beta^{(i)}| \geqslant C_2 \cdot |Y| \qquad \text{for} \quad i \in \{1, \ldots, n\}, i \neq i_0.$$

(ii) *If* $|Y| > Y_1$ *then* X/Y *is a convergent from the continued fraction expansion of* $\xi^{(i_0)}$.

Proof. Let $i_0 \in \{1, ..., n\}$ be such that $|\beta^{(i_0)}| = \min_{1 \leqslant i \leqslant n} |\beta^{(i)}|$. We have from (1.1)

$$|f_0| \cdot \prod_{i=1}^{n} |\beta^{(i)}| = |m|.$$

By the minimality of $|\beta^{(i_0)}|$ we have for all i

$$\begin{aligned} |Y| \cdot |\xi^{(i)} - \xi^{(i_0)}| &= |\beta^{(i)} - \beta^{(i_0)}| \\ &\leqslant |\beta^{(i)}| + |\beta^{(i_0)}| \leqslant 2 \cdot |\beta^{(i)}|. \end{aligned}$$

Hence $|\beta^{(i)}| \geqslant C_2 \cdot |Y|$. Further,

$$\begin{aligned} |\beta^{(i_0)}| &= \frac{|m|}{|f_0|} \cdot \prod_{i \neq i_0} |\beta^{(i)}|^{-1} \\ &\leqslant \frac{|m|}{|f_0|} \cdot \prod_{i \neq i_0} \left(\frac{1}{2} \cdot |Y| \cdot |\xi^{(i)} - \xi^{(i_0)}| \right)^{-1} \\ &= \frac{2^{n-1} \cdot |m|}{|f_0 \cdot \prod_{i \neq i_0} (\xi^{(i)} - \xi^{(i_0)})| \cdot |Y|^{n-1}} \\ &= \frac{2^{n-1} \cdot |m|}{|g'(\xi^{(i_0)})| \cdot |Y|^{n-1}}. \end{aligned}$$

Now, if $i_0 > s$ (and hence $t \geqslant 1$), then, by the definition of Y_0,

$$\begin{aligned} \left| \frac{X}{Y} - \xi^{(i_0)} \right| = \frac{|\beta^{(i_0)}|}{|Y|} &\leqslant \frac{2^{n-1} \cdot |m|}{|g'(\xi^{(i_0)})|} \cdot |Y|^{-n} \\ &\leqslant \left(\frac{Y_0}{|Y|} \right)^n \cdot \min_{s+1 \leqslant i \leqslant s+t} |\operatorname{Im} \xi^{(i)}|, \end{aligned}$$

which is impossible if $|Y| > Y_0$. Hence $i_0 \leqslant s$, and now (i) follows at once. Moreover, if $|Y| > Y_1$, then

$$\left| \frac{X}{Y} - \xi^{(i_0)} \right| = |\beta^{(i_0)}| \cdot |Y|^{-1} \leqslant C_1 \cdot |Y|^{-n} \leqslant \frac{1}{4} \cdot Y_1^{n-2} \cdot |Y|^{-n} \leqslant \frac{1}{2} \cdot |Y|^{-2},$$

and thus $|X/Y - \xi^{(i_0)}| < \frac{1}{2} \cdot |Y|^{-2}$, since $\xi^{(i_0)}$ is irrational. Now (ii) follows from a well-known theorem on continued fractions (see, e.g., Hardy and Wright [15, Theorem 184]). ∎

Now let $|Y| > Y_1$ and $i_0 \in \{1, ..., s\}$ as in Lemma 1.1. Choose $j, k \in \{1, ..., n\}$ such that i_0, j, k are pairwise distinct and either $j, k \in \{1, ..., s\}$

or $j+t=k$ (so that $\xi^{(k)}=\overline{\xi^{(j)}}$), but further the choice of j, k is free. By $\beta^{(i)}=X-Y\cdot\xi^{(i)}$ for $i=i_0, j, k$ we get, on eliminating the X and Y,

$$\beta^{(i_0)}\cdot(\xi^{(j)}-\xi^{(k)})+\beta^{(j)}\cdot(\xi^{(k)}-\xi^{(i_0)})+\beta^{(k)}\cdot(\xi^{(i_0)}-\xi^{(j)})=0,$$

or, equivalently,

$$\frac{\xi^{(i_0)}-\xi^{(j)}}{\xi^{(i_0)}-\xi^{(k)}}\cdot\frac{\beta^{(k)}}{\beta^{(j)}}-1=-\frac{\xi^{(k)}-\xi^{(j)}}{\xi^{(k)}-\xi^{(i_0)}}\cdot\frac{\beta^{(i_0)}}{\beta^{(j)}}. \tag{1.2}$$

By Lemma 1.1, the right-hand side of (1.2) is "extremely small." Put, if j, $k\in\{1,\ldots,s\}$ (let us call it "the real case")

$$\Lambda=\log\left|\frac{\xi^{(i_0)}-\xi^{(j)}}{\xi^{(i_0)}-\xi^{(k)}}\cdot\frac{\beta^{(k)}}{\beta^{(j)}}\right|$$

and if j, $k\in\{s+1,\ldots,s+2\cdot t\}$ (let us call it "the complex case")

$$\Lambda=\frac{1}{i}\cdot\operatorname{Log}\left(\frac{\xi^{(i_0)}-\xi^{(j)}}{\xi^{(i_0)}-\xi^{(k)}}\cdot\frac{\beta^{(k)}}{\beta^{(j)}}\right),$$

where, in general, for $z\in\mathbb{C}$, $\operatorname{Log}(z)$ denotes the principal value of the logarithm of z (hence $-\pi<\operatorname{Im}\operatorname{Log}(z)\leqslant\pi$). By $\xi^{(k)}=\overline{\xi^{(j)}}$ we have $\Lambda\in\mathbb{R}$ and $|\Lambda|\leqslant\pi$.

The following lemma shows how small $|\Lambda|$ is.

LEMMA 1.2. *Put*

$$C_3=\max_{i_1\neq i_2\neq i_3\neq i_1}\left|\frac{\xi^{(i_1)}-\xi^{(i_2)}}{\xi^{(i_1)}-\xi^{(i_3)}}\right|,$$

$$Y_2^*=\max(Y_1,\lceil(2\cdot C_1\cdot C_3/C_2)^{1/n}\rceil).$$

If $|Y|>Y_2^*$ *then*

$$|\Lambda|<\frac{1\cdot 39\cdot C_1\cdot C_3}{C_2}\cdot|Y|^{-n}.$$

Proof. Consider first the real case. From $|Y|>Y_2^*$ and Lemma 1.1 it follows that the right-hand side of (1.2) is absolutely less than $\frac{1}{2}$ and, consequently,

$$\frac{\xi^{(i_0)}-\xi^{(j)}}{\xi^{(i_0)}-\xi^{(k)}}\cdot\frac{\beta^{(k)}}{\beta^{(j)}}>0.$$

It follows that the left-hand side of (1.2) is equal to $e^{\Lambda}-1$, and now (1.2) implies, in view of Lemma 1.1 and the definition of C_3,

$$|e^{\Lambda}-1|<C_3\cdot\frac{C_1\cdot|Y|^{-(n-1)}}{C_2\cdot|Y|}=\frac{C_1\cdot C_3}{C_2}\cdot|Y|^{-n}.$$

On the other hand, $|e^{\Lambda}-1|<\frac{1}{2}$ implies

$$|\Lambda|\leqslant 2\cdot\log 2\cdot|e^{\Lambda}-1|\leqslant 1.39\cdot|e^{\Lambda}-1|,$$

which proves our claim in the real case.

In the complex case the left-hand side of (1.2) is equal to $e^{i\Lambda}-1$, and, as in the real case, we derive

$$|e^{i\Lambda}-1|<\frac{C_1\cdot C_3}{C_2}\cdot|Y|^{-n}<\frac{1}{2}.$$

Since $|e^{i\Lambda}-1|=2\cdot|\sin\Lambda/2|$, it follows that $|\sin\Lambda/2|<\frac{1}{4}$, and therefore

$$|\Lambda|\leqslant 2\cdot\frac{\frac{1}{4}}{\sin\frac{1}{4}}\cdot|\sin\Lambda/2|=\frac{\frac{1}{4}}{\sin\frac{1}{4}}\cdot|e^{i\Lambda}-1|\leqslant 1.02\cdot|e^{i\Lambda}-1|,$$

which proves the lemma in the complex case. ∎

In the ring of integers of the field K (as well as in any other order R of K) there exists a system of fundamental units $\varepsilon_1, \ldots, \varepsilon_r$, where $r=s+t-1$ (Dirichlet's Unit Theorem). Note that since F is irreducible and we have supposed $s>0$, the only roots of unity belonging to K are ± 1. We shall not discuss here the problem of finding such a system (for efficient methods see, e.g., Berwick [3], Billevič [4, 5], Pohst and Zassenhaus [21], Buchmann [8, 9]). We simply assume that a system of fundamental units is known. On the other hand, there exist only finitely many nonassociates $\mu_1, \ldots, \mu_v$ in K such that $f_0\cdot N(\mu_i)=m$ for $i=1, \ldots, v$. (We use $N(\cdot)$ to denote the norm of the extension $K/\mathbb{Q}$.) We also assume that a complete set of such μ_i's is known. Let M be the set of all $\pm\mu_i$'s. (In the important case $|f_0|=|m|=1$, it is clear that $M=\{+1,-1\}$.) Then, for any integral solution (X, Y) of (1.1) there exist some $\mu\in M$ and $a_1, \ldots, a_r\in\mathbb{Z}$, such that

$$\beta=\mu\cdot\varepsilon_1^{a_1}\cdots\varepsilon_r^{a_r}.$$

Thus, the initial problem of solving (1.1) is reduced to that of finding all integral r-tuples $(a_1, \ldots, a_r)$ such that $\mu\cdot\varepsilon_1^{a_1}\cdots\varepsilon_r^{a_r}$ for some $\mu\in M$ be of the special shape $X-Y\cdot\xi$, with $X, Y\in\mathbb{Z}$. As we have seen, X and Y can be eliminated, so that we obtain (1.2). Thus the problem reduces to solving finitely many equations of the type

$$\frac{\xi^{(i_0)}-\xi^{(j)}}{\xi^{(i_0)}-\xi^{(k)}}\cdot\frac{\mu^{(k)}}{\mu^{(j)}}\cdot\prod_{i=1}^{r}\left(\frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(j)}}\right)^{a_i}-1$$
$$=-\frac{\xi^{(k)}-\xi^{(j)}}{\xi^{(k)}-\xi^{(i_0)}}\cdot\frac{\mu^{(i_0)}}{\mu^{(j)}}\cdot\prod_{i=1}^{r}\left(\frac{\varepsilon_i^{(i_0)}}{\varepsilon_i^{(j)}}\right)^{a_i}$$

(the so-called "unit equation"). In the real case we have

$$\Lambda=\log\left|\frac{\xi^{(i_0)}-\xi^{(j)}}{\xi^{(i_0)}-\xi^{(k)}}\cdot\frac{\mu^{(k)}}{\mu^{(j)}}\right|+\sum_{i=1}^{r}a_i\cdot\log\left|\frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(j)}}\right|, \tag{1.3}$$

and in the complex case

$$\Lambda=\operatorname{Arg}\left(\frac{\xi^{(i_0)}-\xi^{(j)}}{\xi^{(i_0)}-\xi^{(k)}}\cdot\frac{\mu^{(k)}}{\mu^{(j)}}\right)+\sum_{i=1}^{r}a_i\cdot\operatorname{Arg}\left(\frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(j)}}\right)+a_0\cdot 2\pi, \tag{1.4}$$

with $a_0\in\mathbb{Z}$, and $-\pi<\operatorname{Arg}(z)\leqslant\pi$ for every $z\in\mathbb{C}$. Note that Λ in the real case, and $i\cdot\Lambda$ in the complex case, is a linear form in (principal) logarithms of algebraic numbers, where the coefficients a_i are integers. The Gelfond–Baker theory provides an explicit lower bound for $|\Lambda|$ in terms of $\max|a_i|$. Using this in combination with Lemma 1.2 we can find an explicit upper bound for $\max|a_i|$. This is what we do in the next section.

2. *Upper Bounds for the Absolute Values of the Unknowns*

Let $A=\max_{1\leqslant i\leqslant r}|a_i|$. First we find an upper bound for A in terms of $|Y|$.

LEMMA 2.1. *Put* $I=\{h_1,\ldots,h_r\}\subset\{1,\ldots,n\}$, *and*

$$U_I=(\log|\varepsilon_l^{(h_i)}|)_{1\leqslant i\leqslant r,\,1\leqslant l\leqslant r}$$

(*where* i *indicates a row and* l *a column of the matrix*),

$$U_I^{-1}=(u_{il}),\qquad N[U_I^{-1}]=\max_{1\leqslant i\leqslant r}\sum_{l=1}^{r}|u_{il}|.$$

Put also

$$\mu_-=\min_{\substack{1\leqslant i\leqslant n\\ \mu\in M}}|\mu^{(i)}|,\qquad \mu_+=\max_{\substack{1\leqslant i\leqslant n\\ \mu\in M}}|\mu^{(i)}|,$$

$$C_4=\frac{\frac{1}{2}+\max_{1\leqslant i_1<i_2\leqslant n}|\xi^{(i_1)}-\xi^{(i_2)}|}{\mu_-},$$

$$C_5=\min((n-1)\cdot\min_I N[U_I^{-1}],\max_I N[U_I^{-1}]).$$

Then, for

$$|Y| > \max(Y_1, 2 \cdot |m|^{1/n}, \mu_+ / C_2),$$

we have

$$A < C_5 \cdot \log(C_4 \cdot |Y|).$$

Proof. By $\beta = \mu \cdot \varepsilon_1^{a_1} \cdot \cdots \cdot \varepsilon_r^{a_r}$ we have

$$\begin{pmatrix} \log |\beta^{(h_1)}/\mu^{(h_1)}| \\ \vdots \\ \log |\beta^{(h_r)}/\mu^{(h_r)}| \end{pmatrix} = U_I \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_r \end{pmatrix}. \tag{2.1}$$

On the other hand, for every $h \in \{1, ..., n\}$, using the end of the proof of Lemma 1.1,

$$\begin{aligned} |\beta^{(h)}| = |X - Y \cdot \xi^{(h)}| &\leqslant |X - Y \cdot \xi^{(i_0)}| + |Y| \cdot |\xi^{(i_0)} - \xi^{(h)}| \\ &\leqslant \frac{1}{2 \cdot |Y|} + |Y| \cdot |\xi^{(i_0)} - \xi^{(h)}| \\ &< \left(\frac{1}{2} + \max_{1 \leqslant i_1 < i_2 \leqslant n} |\xi^{(i_1)} - \xi^{(i_2)}| \right) \cdot |Y|, \end{aligned}$$

and therefore

$$\left| \frac{\beta^{(h)}}{\mu^{(h)}} \right| < C_4 \cdot |Y| \qquad \text{for} \quad h = 1, ..., n.$$

Note that $C_4 \cdot |Y| > 1$. Indeed, by

$$\prod_{i=1}^{n} |\mu^{(i)}| = \frac{|m|}{|f_0|} \leqslant |m|$$

it follows that $\min_{1 \leqslant i \leqslant n} |\mu^{(i)}| \leqslant |m|^{1/n}$, hence $\mu_- \leqslant |m|^{1/n}$. Therefore

$$\begin{aligned} C_4 \cdot |Y| &\geqslant \left(\frac{1}{2} + \max_{1 \leqslant i_1 < i_2 \leqslant n} |\xi^{(i_1)} - \xi^{(i_2)}| \right) \cdot |Y| \cdot |m|^{-1/n} \\ &> \frac{|Y|}{2\,|m|^{1/n}} > 1. \end{aligned}$$

Then,

$$\log \left| \frac{\beta^{(h)}}{\mu^{(h)}} \right| < \log(C_4 \cdot |Y|) \quad (h = 1, ..., n), \quad \log(C_4 \cdot |Y|) > 0. \tag{2.2}$$

Next we show that

$$\left|\log\left|\frac{\beta^{(i)}}{\mu^{(i)}}\right|\right| < (n-1)\cdot\log(C_4\cdot|Y|) \qquad \text{for} \quad i=1,\ldots,n. \tag{2.3}$$

Indeed, in view of (2.2), a stronger inequality is true if $|\beta^{(i)}/\mu^{(i)}| \geqslant 1$. Suppose now that $|\beta^{(i)}/\mu^{(i)}| < 1$. By

$$\prod_{h=1}^{n}\left|\frac{\beta^{(h)}}{\mu^{(h)}}\right| = 1$$

it follows that

$$\begin{aligned}\left|\log\left|\frac{\beta^{(i)}}{\mu^{(i)}}\right|\right| &= -\log\left|\frac{\beta^{(i)}}{\mu^{(i)}}\right| \\ &= \sum_{\substack{h=1\\ h\neq i}}^{n}\log\left|\frac{\beta^{(h)}}{\mu^{(h)}}\right| < (n-1)\cdot\log(C_4\cdot|Y|),\end{aligned}$$

in view of (2.2). Now the inequality

$$A < (n-1)\cdot\min_{I} N[U_I^{-1}]\cdot\log(C_4\cdot|Y|)$$

follows from (2.1), (2.3), the definition of $N[U_I^{-1}]$, and the fact that, as we have not put so far any restriction on I, this could be chosen so that $N[U_I^{-1}]$ be minimal. It remains to show that

$$A < \max_{I} N[U_I^{-1}]\cdot\log(C_4\cdot|Y|).$$

Choose I such that $i_0 \notin I$. Then, by Lemma 1.1, for every $h\in I$, $|\beta^{(h)}/\mu^{(h)}| > C_2\cdot|Y|/\mu_+ > 1$ and now, in view of (2.2),

$$\left|\log\left|\frac{\beta^{(h)}}{\mu^{(h)}}\right|\right| < \log(C_4\cdot|Y|),$$

which implies our assertion. ∎

Lemmas 1.2 and 2.1 immediately yield

LEMMA 2.2. *Put*

$$C_6 = \frac{1.39\cdot C_1\cdot C_3\cdot C_4^n}{C_7}, \quad Y_2' = \max(Y_2^*, 2\cdot|m|^{1/n}, \mu_+/C_2).$$

If $|Y| > Y_2'$ *then*

$$|\Lambda| < C_6 \cdot \exp\left(\frac{-n}{C_5} \cdot A\right).$$

Next we apply the following result of Waldschmidt [30] from the theory of linear forms in logarithms.

LEMMA 2.3. *If* $\Lambda \neq 0$ *then in the real case*

$$|\Lambda| > \exp(-C_7 \cdot (\log A + C_8)), \tag{2.4}$$

and in the complex case this holds when A *is replaced by* $A' = \max_{0 \leq i \leq r} |a_i|$.

The precise values for C_7 and C_8 are given in Appendix II. It should be noted that in the complex case a_0 now appears while in Lemmas 2.1 and 2.2 it was not presented. In order to obtain an upper bound for A we must find an upper bound for A' in terms of A. Indeed, using the relation

$$\operatorname{Arg}(z_1 \cdot z_2) = \operatorname{Arg}(z_1) + \operatorname{Arg}(z_2) + k \cdot 2\pi, \quad k \in \{-1, 0, 1\},$$

it is not difficult to see from (1.4) that $|a_0| < \frac{1}{2} + \frac{1}{2} \cdot r \cdot A + 0.51/2\pi < r \cdot A$ if $A \geqslant 2$. Thus we may apply (2.4) in both cases with A if we replace C_8 by C_8', where

$$C_8' = C_8 \qquad \text{in the real case,}$$

$$C_8' = C_8 + \log r \qquad \text{in the complex case.}$$

We can now give an upper bound for A.

LEMMA 2.4. *Put*

$$C_9 = \frac{2 \cdot C_5}{n} \cdot \left(\log C_6 + C_7 \cdot C_8' + C_7 \cdot \log \frac{C_5 \cdot C_7}{n}\right).$$

If $|Y| > Y_2'$, *then* $A < C_9$.

Proof. As we have seen in the proof of Lemma 1.2, $|e^{\Lambda} - 1| < \frac{1}{2}$ in the real case, and $|e^{i\Lambda} - 1 < \frac{1}{2}$ in the complex case. Note that $\beta^{(i_0)} \neq 0$. Hence (1.2) implies $\Lambda \neq 0$. Therefore Lemmas 2.2 and 2.3 yield

$$A < \frac{C_5}{n} \cdot (\log C_6 + C_7 \cdot C_8' + C_7 \cdot \log A).$$

The result now follows from Pethö and de Weger [20] (Lemma 2.3). ∎

Remark. From this upper bound for A one can derive easily an upper bound for $|Y|$, thus a value for Y_3 (we shall not do this explicitly). This means that, theoretically, the problem of solving the Thue equation (1.1) can be treated completely in a finite number of steps. In practice, however, this is not satisfactory, because the upper bound for A (and hence certainly that for $|Y|$, which is of size $\exp(A)$) is so large that it is completely unrealistic to speak about checking all possibilities, even with the most powerful computers of today. To give an idea, in the quartic Thue equations that we solve in Section III, the corresponding A has an upper bound of the size of 10^{40}. Nevertheless, such a large bound is still very useful, because from it, using a computational reduction technique based on diophantine approximation theory, we can obtain a considerably smaller upper bound, which usually is of the size of the logarithm of the initial upper bound. (This reduction process can be applied successively more than once.) This is the object of the next section.

3. *Reducing the Upper Bound*

We are now left with a problem of the following type. Let be given real numbers $\delta, \mu_1, \ldots, \mu_q$ ($q \geqslant 2$, the case $q = 1$ is trivial). Write

$$\Lambda = \delta + a_1 \cdot \mu_1 + \cdots + a_q \cdot \mu_q,$$

where the a_i's belong to $\mathbb{Z}$, and put $A = \max_{1 \leqslant i \leqslant q} |a_i|$. If K_1, K_2, K_3 are given positive numbers, then find all q-tuples $(a_1, \ldots, a_q) \in \mathbb{Z}^q$ satisfying

$$|\Lambda| < K_1 \cdot \exp(-K_2 \cdot A), \qquad A < K_3. \tag{3.1}$$

In our case, it follows from (1.3) or (1.4) how to define q, δ, and the μ_i's, and from Lemmas 2.2 and 2.4 how to define K_1, K_2, K_3. In general, K_1 and K_2 are "small" constants, whereas K_3 is "very large." Put

$$\Lambda_0 = a_1 \cdot \mu_1 + \cdots + a_q \cdot \mu_q,$$

so that $\Lambda = \delta + \Lambda_0$. We call Λ an inhomogeneous form (if $\delta \neq 0$) and Λ_0 a homogeneous form. In general, we have $\delta \neq 0$ and for our reduction process we can use the "generalized lemma of Davenport," as Ellison [12] gives it. This has been used in practice for solving Thue equations by Agrawal *et al.* [1], Steiner [24], Blass *et al.* [6], and Pethö and Schulenberg [19]. Roughly speaking, to apply this lemma, one has to find good simultaneous rational approximations P_i/Q of the numbers μ_i/μ_q ($1 \leqslant i < q$) (i.e., rational approximations with the same denominator Q), and then to test for $Q \cdot \delta/\mu_q$ not to be very near to an integer. To find such approximations P_i/Q, the algorithm of Lenstra, Lenstra, and Lovász [16] for lattice basis reduction (which we will refer to as the L^3-algorithm) can be used (see [16, (1.38), (1.39)]).

We prefer to use the L^3-algorithm for solving (3.1) in a different way. We propose this alternative method for the following reasons. First, it is a generalization of a similar method for the homogeneous case (cf. de Weger [31, Sect. 4]). Second, it actually produces $a_1, ..., a_q \in \mathbb{Z}$ for which $|A|$ is almost as small as possible under the condition $A < K_3$; i.e., it actually finds a solution that almost satisfies the Kronecker theorem on inhomogeneous diophantine approximation. Third, it can be generalized to the p-adic case (see [31, Sect. 5]), for the homogeneous case; in a forthcoming paper (see Tzanakis and de Weger [29]) we apply the p-adic analogue of this reduction process to solve a Thue-Mahler equation).

We shall apply the "integral version" of the L^3-algorithm, as given in [31, Sect. 3]. The advantage of this version is that in it only integers are involved, and every division is exact (as proved theoretically), thus avoiding at this stage rounding off errors.

In what follows in this section, we use the letter σ (with subscript) to denote positive constants which are, in general, very small compared to K_3. Also, we consider lattices in $\mathbb{Z}^q$. For such a lattice Γ, by a "matrix associated with Γ" we mean a matrix whose column vectors form a basis of Γ (the points or vectors of $\mathbb{R}^q$ resp. $\mathbb{Z}^q$ will be considered as $q \times 1$ matrices).

Below we distinguish three cases. In the first two we suppose that the μ_i's are $\mathbb{Q}$-independent.

(i) (Cf. also [31, Sect. 4]). Choose c_0 somewhat larger than K_3^q, so that $c_0 = \sigma_1 \cdot K_3^q$ ($\sigma_1 > 1$), and consider the lattice Γ associated with the matrix

$$\mathscr{A} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \\ \lceil c_0 \cdot \mu_1 \rceil & \cdots & \lceil c_0 \cdot \mu_{q-1} \rceil \lceil c_0 \cdot \mu_q \rceil \end{pmatrix}.$$

Find a reduced basis $\mathbf{b}_1, ..., \mathbf{b}_q$ of Γ. Then, by Proposition (1.11) of [16], we have for every $\mathbf{x} \in \Gamma$, $\mathbf{x} \neq \mathbf{0}$,

$$|\mathbf{x}| \geqslant 2^{-(q-1)/2} \cdot |\mathbf{b}_1| =: l_0. \tag{3.2}$$

Since the μ_i's are $\mathbb{Q}$-independent, we expect that l_0 is not "very small," and in practice it is of the size of K_3. A heuristic argument for this is the following. By the properties of the reduced basis, its vectors are "almost orthogonal" and of "almost the same length." On the other hand, the volume of the fundamental parallelepiped of this basis equals $\lceil c_0 \cdot \mu_q \rceil$, which is of the size of K_3^q. Therefore we expect that each $|\mathbf{b}_i|$ be of the size of K_3; i.e., we expect that

$$l_0 = \sigma_2 \cdot K_3,$$

where σ_2 depends on σ_1. Suppose now that $(a_1, ..., a_q) \in \mathbb{Z}^q$ satisfies (3.1), and consider the lattice point

$$\mathscr{A} \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_q \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_{q-1} \\ \lambda_0 \end{pmatrix} \in \Gamma,$$

where

$$\lambda_0 = a_1 \cdot [c_0 \cdot \mu_1] + \cdots + a_q \cdot [c_0 \cdot \mu_q].$$

Hence, $|\lambda_0 - c_0 \cdot \Lambda| \leqslant q \cdot A < q \cdot K_3$, from which

$$|\lambda_0| < c_0 \cdot |\Lambda| + q \cdot K_3. \tag{3.3}$$

On the other hand, by (3.2), $a_1^2 + \cdots + a_{q-1}^2 + \lambda_0^2 \geqslant l_0^2$, from which

$$(q-1) \cdot K_3^2 + \lambda_0^2 \geqslant \sigma_2^2 \cdot K_3^2,$$

and this combined with (3.3) gives

$$c_0 \cdot |\Lambda| + q \cdot K_3 \geqslant K_3 \cdot \sqrt{(\sigma_2^2 - (q-1))}.$$

With an appropriate choice of c_0 we can have σ_2 large enough, e.g., such that the square root in the expression above be "somewhat larger" than q. This means that we have

$$c_0 \cdot |\Lambda| > \sigma_3 \cdot K_3$$

(for $\sigma_3 = \sqrt{(\sigma_2^2 - (q-1))} - q$). Combined with the first inequality of (3.1) this yields

$$A < \frac{1}{K_2} \cdot \left(\log \left(\frac{\sigma_1 \cdot K_1}{\sigma_3} \right) + (q-1) \cdot \log K_3 \right),$$

which means that an upper bound for A has been found which is of the size of the logarithm of the previous upper bound. We can formulate a precise result. Indeed, if we substitute l_0, σ_2, and σ_3 in the previous arguments by their precise values, we see that we have already proved the following result.

PROPOSITION 3.1. *If* $|\mathbf{b}_1| > \sqrt{((q^2+q-1)\cdot 2^{q-1})}\cdot K_3$, *then every solution of* (3.1), *in the case* $\delta = 0$, *satisfies*

$$A < \frac{1}{K_2}\cdot(\log(c_0\cdot K_1) - \log(\sqrt{(2^{-(q-1)}\cdot|\mathbf{b}_1|^2 - (q-1)\cdot K_3^2)} - q\cdot K_3)).$$

If necessary we can repeat this process with the new upper bound in place of K_3, to obtain an even smaller bound.

(ii) Let $\delta \neq 0$. In this case we consider the same lattice Γ as in the case $\delta = 0$, and we compute the reduced basis $\mathbf{b}_1, ..., \mathbf{b}_q$. Let $\mathscr{B}$ be the matrix associated with this basis. By the version of the L^3-algorithm that we use we can compute at the same time the matrix $\mathscr{U}$ such that $\mathscr{B} = \mathscr{A}\cdot\mathscr{U}$, and its inverse $\mathscr{U}^{-1}$. Note that, because of the simple form of $\mathscr{A}$, $\mathscr{A}^{-1}$ can be computed very easily, and therefore we can with little extra effort compute $\mathscr{B}^{-1}$. Now consider the point

$$\mathbf{x} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -[c_0\cdot\delta] \end{pmatrix} \in \mathbb{Z}^q,$$

and let $s_1, ..., s_q \in \mathbb{R}$ be its coordinates with respect to the basis $\mathbf{b}_1, ..., \mathbf{b}_q$. The s_i's can be easily found from the relation

$$\begin{pmatrix} s_1 \\ \vdots \\ s_q \end{pmatrix} = \mathscr{B}^{-1}\cdot\mathbf{x}.$$

PROPOSITION 3.2. *Let* $i^* = \max(i\colon 1 \leqslant i \leqslant q$ *and* $s_i \notin \mathbb{Z})$. *If*

$$2^{-(q-1)/2}\cdot\|s_{i^*}\|\cdot|\mathbf{b}_1| \geqslant \sqrt{\left(4\cdot q^2 + 3\cdot q - \frac{3}{4}\right)}\cdot K_3,$$

then every solution of (3.1), *in the case* $\delta \neq 0$, *satisfies*

$$A < \frac{1}{K_2}\cdot\log\left(\frac{c_0\cdot K_1}{q\cdot K_3}\right).$$

Proof. Denote by $l(\mathbf{x}, \Gamma)$ the minimal distance of $\mathbf{x}$ from the points of Γ. By Lemma 3.5 of de Weger [32] we have

$$l(\mathbf{x}, \Gamma) \geqslant 2^{-(q-1)/2}\cdot\|s_{i^*}\|\cdot|\mathbf{b}_1|.$$

Therefore, in view of our hypothesis, it follows that $l(\mathbf{x}, \Gamma) \geqslant \sqrt{(4 \cdot q^2 + 3 \cdot q - \frac{3}{4})} \cdot K_3$. We now have, similarly as in Lemma 3.7 of [32], that if $l(\mathbf{x}, \Gamma) \geqslant \sqrt{(4 \cdot q^2 + 3 \cdot q - \frac{3}{4})} \cdot X_0$ (for some constant $X_0 \geqslant 2$), then the inequality $|\Lambda| > K_1 \cdot \exp(-K_2 \cdot A)$ has no solution in the range $(1/K_2) \cdot \log(c_0 \cdot K_1/q \cdot K_3) \leqslant A \leqslant X_0$, with K_3 in place of X_0, to get immediately the desired upper bound for A. ∎

The new upper bound is of the size of $\log K_3$, as in case (i). Proposition 3.2 can be applied provided that $\|s_{i^*}\|$ is not extremely small. If, however, it is so small that the inequality of the hypothesis is not satisfied, then we have to try with another value of c_0, or to apply Lemma 3.6 from [32] instead of its Lemma 3.5. As in case (i) we can apply the above proposition successively more than once (see Section III, the last part of Subsection 2).

(iii) Suppose now that the μ_i's are $\mathbb{Q}$-dependent. Then we expect the lower bound for $|\mathbf{x}|$ ($\mathbf{x} \in \Gamma$, $\mathbf{x} \neq \mathbf{0}$) in general to be "very small," since the vector having as coordinates the coefficients of the dependence relation will, multiplied by $\mathscr{A}$, give rise to a very short vector in the lattice. So the reduction process described in the two previous cases will not work. In such a case we work as follows. Let M be a maximal subset of $\{\mu_1, ..., \mu_q\}$ consisting of $\mathbb{Q}$-independent numbers. With an appropriate choice of subscripts we may assume that $M = \{\mu_1, ..., \mu_p\}$, $p < q$. Then we can find integers $d > 0$ and d_{ij} ($1 \leqslant i \leqslant p$, $p+1 \leqslant j \leqslant q$) such that

$$d \cdot \mu_j = \sum_{i=1}^{p} d_{ij} \cdot \mu_i, \qquad j = p+1, ..., q.$$

(Note that these numbers d, d_{ij} can be found as coordinates of extremely short vectors in reduced bases.) On the other hand, (3.1) is equivalent to

$$|\Lambda'| < K_1' \cdot \exp(-K_2 \cdot A), \qquad A < K_3, \tag{3.4}$$

where $\Lambda' = d \cdot \Lambda$ and $K_1' = d \cdot K_1$. Now, with $\delta' = d \cdot \delta$ and

$$a_i' = d \cdot a_i + \sum_{j=p+1}^{q} d_{ij} \cdot a_j$$

we obtain

$$\Lambda' = \delta' + \sum_{i=1}^{p} a_i' \cdot \mu_i.$$

Put $D = \max(|d|, |d_{ij}|: 1 \leqslant i \leqslant p, p+1 \leqslant j \leqslant q)$. Then

$$|a_i'| \leqslant (q - p + 1) \cdot D \cdot A \qquad \text{for} \quad i = 1, ..., p.$$

Therefore, if we put $A' = \max_{1 \leqslant i \leqslant p} |a_i'|$, then $A' \leqslant (q-p+1) \cdot D \cdot A$, and (3.4) implies

$$|\Lambda'| < K_1' \cdot \exp(-K_2' \cdot A'), \qquad A' < K_3', \tag{3.5}$$

where

$$\Lambda' = \delta' + a_1' \cdot \mu_1' + \cdots + a_p' \cdot \mu_p', \qquad K_1' = d \cdot K_1,$$
$$K_2' = K_2/(q-1+p) \cdot D, \qquad K_3' = (q-p+1) \cdot K_3.$$

Now, to solve (3.5) we apply the reduction process described in (i) or (ii), depending on whether $\delta' = 0$ or $\delta' \neq 0$, and maybe more than once, if needed, until we find a very small upper bound for A'. Having found all solutions $(a_1', ..., a_p')$ of (3.5), we have at the same time a lower bound $L > 0$ for $|\Lambda'|$. It is reasonable to expect that L is not "extremely small," because the integers $a_1', ..., a_p'$ being "small" in absolute value cannot make Λ' "extremely small." Now combine $|\Lambda'| \geqslant L$ with the first inequality of (3.4) to get

$$A < \frac{1}{K_2} \cdot \log\left(\frac{K_1}{L}\right).$$

Since L is not "very small," as argued heuristically, the above upper bound for A is "small."

Returning now to the general case, we point out that if the reduced upper bound for A is not small enough to admit enumeration of the remaining possibilities in a reasonable time, then it might be necessary, or at least advisable, to use some technique for finding all vectors of a given lattice, whose length is less than some givn "small" bound. In de Weger [31] it is described how the algorithm of Fincke and Pohst [14] can be used to find all such vectors, and how this can be used to reduce the bound for A even further, in the homogeneous case. In the inhomogeneous case, we might analogously want to find all lattice points

$$\mathbf{y} = \begin{pmatrix} a_1 \\ \vdots \\ a_{q-1} \\ \lambda_0 \end{pmatrix}, \tag{3.6}$$

with, say, $|\mathbf{y} - \mathbf{x}| \leqslant K_0$, where K_0 is a given constant somewhat larger than $l(\mathbf{x}, \Gamma)$, and λ_0 and $\mathbf{x}$ are as in case (ii) above.

Let $\mathbf{x} = \sum_{i=1}^{q} s_i \cdot \mathbf{b}_i$, as in case (ii), and let $r_i \in \mathbb{Z}$ satisfy $|r_i - s_i| \leqslant \frac{1}{2}$, $i = 1, ..., q$. If we put $\mathbf{z} = \sum_{i=1}^{q} r_i \cdot \mathbf{b}_i$, then (3.6) implies $|\mathbf{y} - \mathbf{z}| < K_0'$, where $K_0' = K_0 + |\mathbf{x} - \mathbf{z}|$. Write $\mathbf{y}$ as $\mathbf{u} + \mathbf{z}$, where $\mathbf{u}$ now belongs to the lattice.

Then $|\mathbf{u}| < K_0'$, and by the Fincke and Pohst method we can find all possible $\mathbf{u}$, which gives all possible $\mathbf{y}$.

However, when solving a Thue equation, and not only an inequality for a linear form in logarithms, it may be advisable to avoid this Fincke and Pohst method, and to use continued fractions of the roots $\xi^{(i)}$. In practice we can search for the solutions (X, Y) of (1.1) satisfying $Y_1 < |Y| \leqslant C$ as follows, referring to Lemma 1.1. Here, e.g., $C = Y_2$, and we can imagine C here as being a "large" constant compared to Y_1, but not a "very large" one (cf. the introduction of Y_1, Y_2 in Subsection 1).

Let $\tilde{\xi}$ be a rational approximation of $\xi^{(i_0)}$, such that

$$|\tilde{\xi} - \xi^{(i_0)}| < \frac{1}{6 \cdot C^2}. \tag{3.7}$$

Since $|Y| > Y_1$, X/Y must be a convergent, p_k/q_k say, from the continued fraction expansion of $\xi^{(i_0)}$. Denote by $a_0, a_1, a_2, \ldots$ the partial quotients in this expansion. First we claim that $a_{k+1} \geqslant 3$. Indeed, we have

$$\begin{aligned} \frac{1}{(a_{k+1}+2) \cdot |Y|^2} \leqslant \frac{1}{(a_{k+1}+2) \cdot q_k^2} &< \left| \xi^{(i_0)} - \frac{p_k}{q_k} \right| \\ &= \left| \xi^{(i_0)} - \frac{X}{Y} \right| \leqslant \frac{C_1}{|Y|^n}, \end{aligned}$$

where the second inequality from the left is a well-known result on continued fractions. If $a_{k+1} = 1$ or 2, then we would have $|Y|^{n-2} < 4 \cdot C_1$, which is absurd, since $|Y| > Y_1 > (4 \cdot C_1)^{1/(n-2)}$. Thus, $a_{k+1} \geqslant 3$, and by a well-known result on continued fractions we have

$$\left| \xi^{(i_0)} - \frac{p_k}{q_k} \right| < \frac{1}{a_{k+1} \cdot q_k^2} \leqslant \frac{1}{3 \cdot q_k^2}.$$

(The above-mentioned well-known inequalities follow from Theorems 163 and 171 from Hardy and Wright [15].) Therefore,

$$\begin{aligned} \left| \tilde{\xi} - \frac{p_k}{q_k} \right| &\leqslant |\tilde{\xi} - \xi^{(i_0)}| + \left| \xi^{(i_0)} - \frac{p_k}{q_k} \right| \\ &< \frac{1}{6 \cdot C^2} + \frac{1}{3 \cdot q_k^2} \leqslant \frac{1}{2 \cdot q_k^2} \end{aligned}$$

and this means that p_k/q_k is in fact a convergent from the continued fraction expansion of $\tilde{\xi}$ also. Moreover, in view of the inequalities

$$\frac{1}{(a_{k+1}+2) \cdot q_k^2} < \left| \xi^{(i_0)} - \frac{p_k}{q_k} \right| \leqslant \frac{C_1}{|Y|^n} \leqslant \frac{C_1}{|q_k|^n},$$

a_{k+1} must be sufficiently large compared to q_k, namely

$$a_{k+1} > \frac{|q_k|^{n-2}}{C_1} - 2. \tag{3.8}$$

This inequality can be checked easily for all k such that $q_k \leqslant C$.

To sum up, we propose the following process for every real root $\xi^{(i_0)}$ for $i_0 = 1, ..., s$ (note that i_0 is a priori not known). (1) Compute a rational approximation $\tilde{\xi}$ of $\xi^{(i_0)}$ (a truncation of its decimal expansion will do) satisfying (3.7). (2) Expand $\tilde{\xi}$ into its continued fraction with partial quotients $b_0, b_1, b_2, ..., b_{k+1}$ and convergents p_i/q_i for all $i = 1, ..., k$ with $q_k \leqslant C < q_{k+1}$. (3) Test all these convergents for the conditions (3.8) and $F(p_i, q_i) = m$. Concerning this last test, note that if $X/Y = p_i/q_i$, then $X = Z \cdot p_i$, $Y = Z \cdot q_i$ for some $Z \in \mathbb{Z}$ with $Z^n | m$. This simple observation excludes in general most of the reducible quotients X/Y, and all of them if m is an nth-powerfree integer.

Having tested for all solutions in the range $|Y| \leqslant C$ we may suppose that $|Y| > C$. For such solutions (X, Y) we can obtain a lower bound for the corresponding A as follows (the idea is due to A. Pethö; cf. also Section 1 of Blass *et al.* [6]). For every $(i, j) \in \{1, ..., r\} \times \{1, ..., n\}$ let v_{ij} be the number $+1$ or -1 for which $|\varepsilon_i^{(j)}|^{v_{ij}} \geqslant 1$, and put $E_j = \prod_{i=1}^{r} |\varepsilon_i^{(j)}|^{v_{ij}}$. Then

$$|\beta^{(j)}| = |\mu^{(j)}| \cdot \prod_{i=1}^{r} |\varepsilon_i^{(j)}|^{a_i} \leqslant \mu_+ \cdot E_j^A$$

and hence, for any pair j_1, j_2 with $j_1 \neq j_2$, we have

$$|Y| = \frac{|\beta^{(j_1)} - \beta^{(j_2)}|}{|\xi^{(j_1)} - \xi^{(j_2)}|} \leqslant \mu_+ \cdot \frac{E_{j_1}^A + E_{j_2}^A}{|\xi^{(j_1)} - \xi^{(j_2)}|},$$

and from this we can find a lower bound for A, if we know that $|Y| > C$. Of course, for another pair j_1, j_2 we may find a different lower bound, and therefore we can take the larger one.

III. An Application: Integral Points on the Elliptic Curve $y^2 = x^3 - 4 \cdot x + 1$

In this section we will prove, as an application of the general theory described in Section II, the following results.

Theorem A. *The elliptic curve*

$$y^2 = x^3 - 4 \cdot x + 1 \tag{E}$$

has only the following 22 *integral points*:

$$(x, \pm y) = (-2, 1), (-1, 2), (0, 1), (2, 1), (3, 4), (4, 7), (10, 31),$$
$$(12, 41), (20, 89), (114, 1217), (1274, 45473).$$

COROLLARY. *The only triangular numbers* (*i.e., numbers* $T_n = \frac{1}{2} \cdot n \cdot (n+1)$ *with* n *a positive integer*) *which are a product of three consecutive integers are* T_3, T_{15}, T_{20}, T_{44}, T_{608}, *and* T_{22736}.

Proof of the corollary. Consider the equation $T_n = m \cdot (m+1) \cdot (m+2)$. As noted by Mohanty [17], on putting $m = (x-2)/2$, $n = (y-1)/2$ with $x > 2$ even and $y > 1$ odd, the above equation is transformed into (E). Then, by Theorem A, the only possible values for y are 7, 31, 41, 89, 1217, and 45473, which proves our claim. ∎

We prove Theorem A in two main steps. First, we reduce the problem to the solution of two quartic Thue equations. Then we solve these equations using the general theory developed in Section II.[1]

1. *From the Elliptic Curve to a Couple of Totally Real Quartic Thue Equations*

Let L be the totally real field $\mathbb{Q}(\psi)$, where

$$\psi^3 - 4 \cdot \psi + 1 = 0.$$

Let the conjugates of ψ be $\psi^{(1)} = 0.254...$, $\psi^{(2)} = -2.114$, $\psi^{(3)} = 1.860...$. From a table of Delone and Faddeev [10, p. 141] we see that the class number of L is 1, its ring of integers is $\mathbb{Z}[\psi]$, its discriminant is 229, and a pair of units is ψ, $2-\psi$. From Table I of Buchmann [8] we see that $-7+2 \cdot \psi^2$, $2 \cdot \psi + \psi^2$ is a pair of fundamental units in $\mathbb{Z}[\psi]$. Since $-7+2 \cdot \psi^2 = -\psi^{-1} \cdot (2-\psi)$ and $2 \cdot \psi + \psi^2 = (2-\psi)^{-1}$ we see that ψ, $2-\psi$ is also a pair of fundamental units in $\mathbb{Z}[\psi]$.

The equation (E) of the elliptic curve can be written as

$$y^2 = (x-\psi) \cdot (x^2 + x \cdot \psi + (\psi^2 - 4)) \tag{1.1}$$

and the factors on the right-hand side are relatively prime. Indeed, if π were a common prime divisor of them, then π would divide

$$(x^2 + x \cdot \psi + (\psi^2 - 4)) - (x + 2 \cdot \psi) \cdot (x - \psi) = 3 \cdot \psi^2 - 4,$$

[1] *Note added in proof.* The elementary proof of our Theorem A presented in [17] is not correct. As noted by A. Bremner, at two crucial points ([17], p. 92, line 4, 5 and −10, −9) implications of the type $(d \mid a \cdot b \Rightarrow d \mid a$ or $d \mid b)$ are used.

which is prime, since its norm is -229. Therefore we would have that π is a unit times this prime, and then by (1.1), $x-\psi = \text{unit} \times (3\cdot\psi^2-4)\times$ square. Taking norms we get $y^2 = \pm 229 \times \text{square}$, which is clearly impossible.

Now (1.1) implies

$$x-\psi = \pm\psi^i\cdot(2-\psi)^j\cdot\alpha^2, \qquad \alpha\in\mathbb{Z}[\psi],\ i,\ j\in\{0,1\}. \tag{1.2}$$

Since (E) is trivial to solve for $x\leqslant 0$ (the only solutions with $x\leqslant 0$ are the first three pairs stated in the theorem), we may assume that $x\geqslant 1$. Since $\psi^{(1)} = 0.254...$, we see that the minus sign in (1.2) is impossible. Then, by $\psi^{(2)} = -2.114...$, $i\neq 1$. We conclude therefore that

$$x-\psi = (2-\psi)^j\cdot(u+v\cdot\psi+w\cdot\psi^2)^2, \quad u,v,w\in\mathbb{Z},\ j\in\{0,1\}. \tag{1.3}$$

First case: $j=0$. Then (1.3) implies, on equating corresponding coefficients in both sides,

$$x = u^2-2\cdot v\cdot w, \quad w^2-2\cdot u\cdot v-8\cdot v\cdot w = 1, \quad v^2+4\cdot w^2+2\cdot u\cdot w = 0. \tag{1.4}$$

Note that w is odd and v is even, hence $4\mid 2\cdot u\cdot w$, so u is even. Put $u = 2\cdot u_1$, $v = 2\cdot v_1$. The last equation of (1.4) now reads

$$w^2+u_1\cdot w+v_1^2 = 0.$$

Consider this as a quadratic equation in w. Its discriminant must be a square, z^2 say. Then

$$u_1^2-4\cdot v_1^2 = z^2, \qquad w = \tfrac{1}{2}(-u_1\pm z).$$

Note that u_1 and z have the same parity. We may assume $u\geqslant 0$.

First, suppose that u_1 and z are even. Since $w^2+u_1\cdot w+v_1^2=0$ and w is odd, we find $u_1\equiv 2 \pmod 4$, and v_1 is odd. Put $u_1 = 2\cdot u_2$, $z = 2z_1$. Then $u_2^2-v_1^2 = z_1^2$, where u_2 and v_1 are odd. By $u_2\geqslant 0$ there exist m, $n\in\mathbb{Z}$ such that

$$u_2 = m^2+n^2, \qquad v_1 = m^2-n^2, \qquad z_1 = 2\cdot m\cdot n.$$

It follows that

$$u = 4\cdot(m^2+n^2), \qquad v = 2\cdot(m^2-n^2), \qquad w = -(m\pm n)^2.$$

Since the sign of z, and thus that of n, is of no importance, we may assume $w = -(m+n)^2$. After substitution in the second equation of (1.4) we obtain the Thue equation

$$m^4+36\cdot m^3\cdot n+6\cdot m^2\cdot n^2-28\cdot m\cdot n^3+n^4 = 1.$$

The left-hand side can be factored as

$$(m+n)\cdot(m^3+35\cdot m^2\cdot n-29\cdot m\cdot n^2+n^3),$$

and therefore it can be solved very easily. Its only solutions are $\pm(m, n)=(1, 0), (0, 1)$. They lead to $\pm(u, v, w)=(4, 2, -1), (4, -2, -1)$, and then by (1.4) we find $x=20, 12$, respectively, which furnish the solutions $(x, \pm y)=(20, 89), (12, 41)$ for (E).

Second, we suppose that u_1 and z are odd. Then v_1 is even, so by $u_1 \geqslant 0$ there exist $m, n \in \mathbb{Z}$ with

$$u_1=m^2+n^2, \quad 2\cdot v_1=2\cdot m\cdot n, \quad z=m^2-n^2.$$

It follows that

$$u=2\cdot(m^2+n^2), \quad v=2\cdot m\cdot n, \quad w=-m^2 \quad \text{or} \quad w=-n^2.$$

We may assume that $w=-m^2$. Substituting this in the second equation of (1.4) we find the Thue equation

$$m^4+8\cdot m^3\cdot n-8\cdot m\cdot n^3=1.$$

The left-hand side is again reducible. The only solutions, as is easily seen, are $\pm(m, n)=(1, 0), (1, 1), (1, -1)$. Since m and n cannot have the same parity, only the first pair is accepted. It leads to $(u, v, w)=(2, 0, -1)$, and hence to $(x, \pm y)=(4, 7)$ for (E).

Second Case: $j=1$. Then, equating the coefficients in (1.3) we get

$$x=2\cdot u^2+v^2+4\cdot w^2+2\cdot u\cdot w-4\cdot v\cdot w, \tag{1.5}$$

$$\begin{aligned} u^2+4\cdot v^2+18\cdot w^2-4\cdot u\cdot v+8\cdot u\cdot w-18\cdot v\cdot w&=1,\\ 2\cdot v^2+9\cdot w^2-2\cdot u\cdot v+4\cdot u\cdot w-8\cdot v\cdot w&=0. \end{aligned} \tag{1.6}$$

The first relation of (1.6) can be replaced by

$$u^2-2\cdot v\cdot w=1. \tag{1.7}$$

Note that u is odd. Put $z=v-2\cdot w$. Then the second equation of (1.6) yields

$$w^2=2\cdot z\cdot(u-z).$$

First, we suppose that z is odd. Then there exist $m, n \in \mathbb{Z}$ such that

$$z=m^2, \qquad u-z=2\cdot n^2,$$

where we use that $u \geqslant 0$ and $(u, w) = 1$. Thus, choosing signs properly,

$$u = m^2 + 2 \cdot n^2, \quad v = m^2 + 4 \cdot m \cdot n, \quad w = 2 \cdot m \cdot n.$$

Substituting this in (1.7) we obtain the Thue equation

$$m^4 - 4 \cdot m^3 \cdot n - 12 \cdot m^2 \cdot n^2 + 4 \cdot n^4 = 1. \tag{1.8}$$

In the next section, in Theorem B, we prove that this equation has only the solutions $\pm(m, n) = (1, 0)$, leading to $(u, v, w) = (1, 1, 0)$, and finally for (E) to $(x, \pm y) = (3, 4)$.

Second, we suppose that z is even. Then there exist $m, n \in \mathbb{Z}$ with

$$z = 2 \cdot m^2, \qquad u - z = n^2.$$

Thus, choosing signs properly, we find

$$u = 2 \cdot m^2 + n^2, \quad v = 2 \cdot m^2 + 4 \cdot m \cdot n, \quad w = 2 \cdot m \cdot n.$$

Now, substituting into (1.7), we obtain the Thue equation

$$n^4 - 12 \cdot n^2 \cdot m^2 - 8 \cdot n \cdot m^3 + 4 \cdot m^4 = 1. \tag{1.9}$$

In the next section, in Theorem B, we prove that this equation has only the solutions $\pm(m, n) = (0, 1), (1, -1), (3, 1), (-1, 3)$. They lead respectively to $(u, v, w) = (1, 0, 0), (3, -2, -2), (19, 30, 6), (11, -10, -6)$, which lead for (E) to the solutions $(x, \pm y) = (2, 1), (10, 31), (1274, 45473), (114, 1217)$. Thus, this result completes the proof of Theorem A, provided the Thue equations (1.8), (1.9) have as their only solutions the pairs (m, n) mentioned above. The proof of this fact will be the object of the next subsection.

2. *Solving the Thue Equations*

In this section we will prove the following result.

THEOREM B. (i) *The Thue equation*

$$X^4 - 4 \cdot X^3 \cdot Y - 12 \cdot X^2 \cdot Y^2 + 4 \cdot Y^4 = 1 \tag{2.1}$$

has only the solutions $\pm(X, Y) = (1, 0)$.

(ii) *The Thue equation*

$$X^4 - 12 \cdot X^2 \cdot Y^2 - 8 \cdot X \cdot Y^3 + 4 \cdot Y^4 = 1 \tag{2.2}$$

has only the solutions $\pm(X, Y) = (1, 0), (1, -1), (1, 3), (3, -1)$.

Proof. We will use the natation and results of Section II. Let the algebraic numbers ϑ and φ be defined by

$$\vartheta^4 - 12 \cdot \vartheta^2 - 8 \cdot \vartheta + 4 = 0, \quad \varphi^4 - 4 \cdot \varphi^3 - 12 \cdot \varphi^2 + 4 = 0.$$

Since $\varphi = 2/\vartheta$, it follows that ϑ and φ generate the same field K over $\mathbb{Q}$. In the notation of Subsection II.1 we have $n = 4$, $s = 4$, $t = 0$, and $\xi = \vartheta$ or $\xi = \varphi$. Simple numerical computations show that we can take

$$Y_0 = 1,\ C_1 = 0.843,\ C = 0.589,\ Y_1 = 2,\ C_3 = 6.645,$$
$$Y_2^* = 3,\ \mu_- = \mu_+ = 1,\ C_4 = 8.3374.$$

In these computations we estimate C_1, C_3, C_4 from above and C_2 from below, making use of the following approximations for the conjugates of ϑ and φ:

$$\vartheta^{(1)} \cong -1.080\,286\,352, \qquad \varphi^{(1)} \cong -1.851\,360\,980,$$
$$\vartheta^{(2)} \cong \ \ 3.722\,935\,260, \qquad \varphi^{(2)} \cong \ \ 0.537\,210\,524,$$
$$\vartheta^{(3)} \cong \ \ 0.334\,111\,716, \qquad \varphi^{(3)} \cong \ \ 5.986\,021\,747,$$
$$\vartheta^{(4)} \cong -2.976\,760\,624, \qquad \varphi^{(4)} \cong -0.671\,871\,290.$$

Now we work in the order R of K with $\mathbb{Z}$-basis $\{1, \vartheta, \frac{1}{2} \cdot \vartheta^2, \frac{1}{2} \cdot \vartheta^3\}$ (note that $\frac{1}{2} \cdot \vartheta^2$ is an algebraic integer). Note that

$$\varphi = \frac{2}{\vartheta} = 4 + 6 \cdot \vartheta - \frac{1}{2} \cdot \vartheta^3 \in R.$$

On the other hand, (2.1) and (2.2) are respectively equivalent to $\mathrm{Norm}_{K/\mathbb{Q}}(X - Y \cdot \vartheta) = 1$ and $\mathrm{Norm}_{K/\mathbb{Q}}(X - Y \cdot \varphi) = 1$, which means that if (X, Y) is a solution of (2.1) or (2.2), then $X - Y \cdot \vartheta$ or $X - Y \cdot \varphi$, respectively, is a unit of the order R. In Appendix II we prove that a system of fundamental units of R is

$$\eta_1 = 3 - 11 \cdot \vartheta - \vartheta^2 + \vartheta^3, \quad \eta_2 = 324 - 948 \cdot \vartheta - \tfrac{189}{2} \cdot \vartheta^2 + \tfrac{175}{2} \cdot \vartheta^3,$$
$$\eta_3 = 622 - 1820 \cdot \vartheta - \tfrac{363}{2} \cdot \vartheta^2 + 168 \cdot \vartheta^3.$$

But, as is obvious from the known solutions of (2.2),

$$\varepsilon_1 = 1 + \vartheta, \qquad \varepsilon_2 = 3 + \vartheta$$

are units of R, and so is

$$\varepsilon_3 = \tfrac{1}{2} \cdot \vartheta^2.$$

Note that $\varepsilon_1 = -\eta_1^{-1}$, $\varepsilon_2 = \eta_2^{-1} \cdot \eta_3$, $\varepsilon_3 = \eta_1^{-3} \cdot \eta_3$. Since the matrix of the exponents is unimodular, it follows that ε_1, ε_2, ε_3 is also a system of fundamental units of R.

Thus the solution of (2.1) and (2.2) is reduced to finding all $(a_1, a_2, a_3) \in \mathbb{Z}^3$ such that the unit $\pm \varepsilon_1^{a_1} \cdot \varepsilon_2^{a_2} \cdot \varepsilon_3^{a_3}$ has the special shape $X - Y \cdot \vartheta$ or $X - Y \cdot \varphi$, respectively. In the notation of Lemma II.2.1 we have, after some numerical computations, that we leave to the reader to check, that

$$\min_I N[U_I^{-1}] = 0.634950..., \quad \max_I N[U_I^{-1}] = 1.210070...$$

(here, of course, $I = \{1, 2, 3, 4\}$). Therefore we can take

$$C_5 = 1.211.$$

Also,

$$C_6 = 6.38771 \times 10^4, \qquad Y_2' = 3.$$

(The values of C_5 and C_6 are estimated from above.)

Now, the relation (1.3) from Section II in our case becomes

$$\begin{aligned} \Lambda = \log \left| \frac{\xi^{(i_0)} - \xi^{(j)}}{\xi^{(i_0)} - \xi^{(k)}} \right| + a_1 \cdot \log \left| \frac{\varepsilon_1^{(k)}}{\varepsilon_1^{(j)}} \right| \\ + a_2 \cdot \log \left| \frac{\varepsilon_2^{(k)}}{\varepsilon_2^{(j)}} \right| + a_3 \cdot \log \left| \frac{\varepsilon_3^{(k)}}{\varepsilon_3^{(j)}} \right|, \end{aligned} \tag{2.3}$$

where $\xi = \vartheta$ or φ. As mentioned in Subsection II.1, once i_0 is fixed, we can choose j, k arbitrarily. Thus we can choose

$$\begin{aligned} &j = 3, k = 4 \qquad \text{if} \quad i_0 = 1 \text{ or } 2, \\ &j = 1, k = 2 \qquad \text{if} \quad i_0 = 3 \text{ or } 4. \end{aligned} \tag{2.4}$$

Therefore, for each $\xi \in \{\vartheta, \varphi\}$ we have four possibilities for Λ. For each of these eight cases we have, as will be shown in Appendix II,

$$C_7 = 5.71 \times 10^{38}, \qquad C_8 = 6.17,$$

and therefore, by Lemma II.2.4, if $|Y| > 3$, then for $A = \max_{1 \leqslant i \leqslant 3} |a_i|$ we have the upper bound 3.26×10^{40}. As is easily checked, the only solutions of either (2.1) or (2.2) with $|Y| \leqslant 3$ are those listed in the statement of the theorem. Therefore we may assume that $|Y| > 3$, so that

$$A < 3.26 \times 10^{40}.$$

We now have to apply the reduction process described in Subsection II.3. In our situation we have to solve II (3.1) with

$$K_1 = C_6 = 6.38771 \times 10^4, \quad K_2 = \frac{n}{C_5} = \frac{4}{1.211} > 3.303, \quad K_3 = 3.26 \times 10^{40}$$

(K_2 is estimated from below), and

$$\Lambda = \delta + a_1 \cdot \mu_1 + a_2 \cdot \mu_2 + a_3 \cdot \mu_3,$$

where for δ and the μ_i's we have the following possibilities, in view of (2.3) and (2.4):

$$\delta = \delta_1 := \log \left| \frac{\xi^{(1)} - \xi^{(3)}}{\xi^{(1)} - \xi^{(4)}} \right| \quad \text{or}$$

$$\delta = \delta_2 := \log \left| \frac{\xi^{(2)} - \xi^{(3)}}{\xi^{(2)} - \xi^{(4)}} \right|, \text{ where } \xi = \vartheta \text{ or } \varphi, \tag{2.5}$$

$$\mu_i = \log \left| \frac{\varepsilon_i^{(4)}}{\varepsilon_i^{(3)}} \right|, \qquad i = 1, 2, 3,$$

or

$$\delta = \delta_3 := \log \left| \frac{\xi^{(3)} - \xi^{(1)}}{\xi^{(3)} - \xi^{(2)}} \right| \quad \text{or}$$

$$\delta = \delta_4 := \log \left| \frac{\xi^{(4)} - \xi^{(1)}}{\xi^{(4)} - \xi^{(2)}} \right|, \text{ where } \xi = \vartheta \text{ or } \varphi, \tag{2.6}$$

$$\mu_i = \log \left| \frac{\varepsilon_i^{(2)}}{\varepsilon_i^{(1)}} \right|, \qquad i = 1, 2, 3.$$

We now take $c_0 = 10^{140}$, and we work with the lattice with associated matrix

$$\mathscr{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ [c_0 \cdot \mu_1] & [c_0 \cdot \mu_2] & [c_0 \cdot \mu_3] \end{pmatrix}.$$

Note that in each of the four cases of (2.5) (resp. (2.6)) we have the same lattice, Γ_1 (resp. Γ_2), say. In any case $\delta \neq 0$, and we had no numerical evidence that the μ_i's are $\mathbb{Q}$-dependent. Therefore we worked as in case (ii) of Subsection II.3.

For each Γ_i we have applied the integral version of the L^3-algorithm, and each time we have computed the integral 3×3-matrices $\mathscr{B}$, $\mathscr{U}$, $\mathscr{U}^{-1}$, as defined in Subsection II.3. In our cases, the coordinates of the vectors of

the reduced bases (i.e., the elements of $\mathscr{B}$) turned out to have 46 to 48 digits; i.e., the lengths of the reduced basis vectors are of the size of $c_0^{1/3}$, as expected. In each of the eight cases we computed the coordinates s_1, s_2, s_3 of the vector

$$\mathbf{x} = \begin{pmatrix} 0 \\ 0 \\ -[c_0 \cdot \delta] \end{pmatrix}$$

with respect to the reduced basis $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ of the lattice. From our computations we found

$$|\mathbf{b}_1| > 3.247 \times 10^{46} \qquad \text{in the case of lattice } \Gamma_1,$$
$$|\mathbf{b}_1| > 4.846 \times 10^{46} \qquad \text{in the case of lattice } \Gamma_2,$$
$$\|s_3\| > 0.029 \qquad \text{in all 8 cases.}$$

In the notation of Proposition II.3.2 we have $i^* = 3$, and in view of the above estimation of $|\mathbf{b}_1|$ it is easily checked that the hypothesis of this proposition is satisfied. Therefore

$$A < \frac{1}{3.303} \cdot \log(10^{140} \cdot 6.38771 \times 10^4/3 \cdot 3.26 \times 10^{40}) < 72.4.$$

It follows that $A \leqslant 72$.

We repeat the procedure with $K_3 = 72$ and $c_0 = 10^{12}$. We found from our computations

$$|\mathbf{b}_1| > 1.293 \times 10^4 \qquad \text{in the case of lattice } \Gamma_1,$$
$$|\mathbf{b}_1| > 1.092 \times 10^4 \qquad \text{in the case of lattice } \Gamma_2,$$
$$\|s_3\| > 0.143 \qquad \text{in all 8 cases.}$$

As before, the hypothesis of Proposition II.3.2 is satisfied, and consequently

$$A < \frac{1}{3.303} \cdot \log(10^{12} \cdot 6.38771 \times 10^4/3 \times 72) < 10.1.$$

It follows that $A \leqslant 10$. We enumerated all remaining possibilities and found no other solutions of (2.1) and (2.2) than mentioned in the theorem. This completes the proof of Theorem B, hence also that of Theorem A. ∎

The total computation time on the IBM 3083 computer at Leiden used for proving Theorem B was about 35 sec. In [28] some numerical details are given.

IV. Appendixes

APPENDIX I: *Fundamental Units in a Totally Real Quartic Field*

To find the fundamental units of an order of a totally real quartic field one can apply the following result of Billevič [4, Theorem 5], conveniently formulated here.

THEOREM. *Let R be an order of a totally real quartic field, generated by $1, \omega_2, \omega_3, \omega_4$. Consider the lattice of $\mathbb{R}^4$ spanned by*

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} \omega_i^{(1)} \\ \omega_i^{(2)} \\ \omega_i^{(3)} \\ \omega_i^{(4)} \end{pmatrix}, \qquad i = 2, 3, 4.$$

We identify the algebraic integer $\alpha = x_1 + x_2 \cdot \omega_2 + x_3 \cdot \omega_3 + x_4 \cdot \omega_4 \in R$ with the lattice point

$$\alpha = \begin{pmatrix} \alpha^{(1)} \\ \alpha^{(2)} \\ \alpha^{(3)} \\ \alpha^{(4)} \end{pmatrix} = x_1 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \sum_{i=2}^{4} x_i \cdot \begin{pmatrix} \omega_i^{(1)} \\ \omega_i^{(2)} \\ \omega_i^{(3)} \\ \omega_i^{(4)} \end{pmatrix}.$$

Consider the "1-sequence" of lattice points α such that

$$|\alpha^{(i)}| < 1, \qquad i = 2, 3, 4, \qquad \alpha^{(1)} > 1.$$

This 1-sequence is a partially ordered set:

$$\alpha \ll \beta \qquad \text{iff} \quad \alpha^{(1)} < \beta^{(1)}.$$

Let ε_1 be the first unit in the 1-sequence. Let ε_2 be the next unit in the 1-sequence such that $|\varepsilon_2^{(i)}| > |\varepsilon_1^{(i)}|$ for at least one $i \in \{2, 3, 4\}$. We now proceed to define ε_3. For $v \in R$ we put

$$\Delta(v) = \begin{vmatrix} \log |v^{(2)}| & \log |v^{(3)}| & \log |v^{(4)}| \\ \log |\varepsilon_1^{(2)}| & \log |\varepsilon_1^{(3)}| & \log |\varepsilon_1^{(4)}| \\ \log |\varepsilon_2^{(2)}| & \log |\varepsilon_2^{(3)}| & \log |\varepsilon_2^{(4)}| \end{vmatrix}.$$

Let μ be the first unit in the 1-sequence after ε_2 that satisfies $\Delta(\mu) \neq 0$.

Case 1. If $\mu^{(1)} > \varepsilon_1^{(1)} \cdot \varepsilon_2^{(1)}$, put $\varepsilon_3 = \mu$.

Case 2. If $\mu^{(1)} < \varepsilon_1^{(1)} \cdot \varepsilon_2^{(1)}$, then let ξ be the first element (not necessarily a unit) in the 1-sequence with $\xi^{(1)} > \sqrt{(\varepsilon_1^{(1)} \cdot \varepsilon_2^{(1)} \cdot \mu^{(1)})}$.

Case 2.1. If there is no unit ε in the 1-sequence with $\mu \ll \varepsilon \leqq \xi$, put $\varepsilon_3 = \mu$.

Case 2.2. If there exist units ε in the 1-sequence such that $\mu \ll \varepsilon \leqq \xi$ then:

Case 2.2.1. If for every unit ε of Case 2.2 at least one of

$$\Delta(\varepsilon) \cdot \Delta(\mu) \leqslant 0, \qquad |\Delta(\varepsilon)| \geqslant |\Delta(\mu)|$$

is true, then put $\varepsilon_3 = \mu$.

Case 2.2.2. If there exist units ε of Case 2.2 such that both

$$\Delta(\varepsilon) \cdot \Delta(\mu) > 0, \qquad |\Delta(\varepsilon)| < |\Delta(\mu)|$$

are true, then we denote by $\{\mu_1, ..., \mu_k\}$ the set of all such units ε, and we define

$$\Delta_0 = \min_{1 \leqslant i \leqslant k} \{|\Delta(\mu_i)|, |\Delta(\mu) - \Delta(\mu_i)|\}.$$

Then:

Case 2.2.2.1. If $\Delta_0 = |\Delta(\mu_j)|$ for some $j \in \{1, ..., k\}$, put $\varepsilon_3 = \mu_j$.

Case 2.2.2.2. If $\Delta_0 = |\Delta(\mu) - \Delta(\mu_j)|$ for some $j \in \{1, ..., k\}$, put $\varepsilon_3 = \mu \cdot \mu_j^{-1}$.

Then ε_1, ε_2, ε_3 form a system of fundamental units of R.

EXAMPLE. Let ϑ and the order R of $\mathbb{Q}(\vartheta)$ be defined as in Subsection III.2. A $\mathbb{Z}$-basis for R is $\{1, \vartheta, \frac{1}{2} \cdot \vartheta^2, \frac{1}{2} \cdot \vartheta^3\}$. By solving

$$1 < x_1 + x_2 \cdot \vartheta^{(1)} + x_3 \cdot \tfrac{1}{2} \cdot \vartheta^{(1)^2} + x_4 \cdot \tfrac{1}{2} \cdot \vartheta^{(1)^3} < c_1,$$
$$-1 < x_1 + x_2 \cdot \vartheta^{(i)} + x_3 \cdot \tfrac{1}{2} \cdot \vartheta^{(i)^2} + x_4 \cdot \tfrac{1}{2} \cdot \vartheta^{(i)^3} < 1, \quad i = 2, 3, 4$$

in $x_1, ..., x_4 \in \mathbb{Z}$, with $c_1 = 15000$ we found that the first six units in the 1-sequence, according to the partial ordering "$\ll$," are

$$E_1 = 3 - 11 \cdot \vartheta - 2 \cdot \tfrac{1}{2} \cdot \vartheta^2 + 2 \cdot \tfrac{1}{2} \cdot \vartheta^3, \quad E_1^{(1)} = 12.45...,$$
$$E_2 = 45 - 130 \cdot \vartheta - 26 \cdot \tfrac{1}{2} \cdot \vartheta^2 + 24 \cdot \tfrac{1}{2} \cdot \vartheta^3, \quad E_2^{(1)} = 155.13...,$$
$$E_3 = 324 - 948 \cdot \vartheta - 189 \cdot \tfrac{1}{2} \cdot \vartheta^2 + 175 \cdot \tfrac{1}{2} \cdot \vartheta^3, \quad E_3^{(1)} = 1127.51...,$$
$$E_4 = 555 - 1625 \cdot \vartheta - 324 \cdot \tfrac{1}{2} \cdot \vartheta^2 + 300 \cdot \tfrac{1}{2} \cdot \vartheta^3, \quad E_4^{(1)} = 1932.31...,$$
$$E_5 = 622 - 1820 \cdot \vartheta - 363 \cdot \tfrac{1}{2} \cdot \vartheta^2 + 336 \cdot \tfrac{1}{2} \cdot \vartheta^3, \quad E_5^{(1)} = 2164.50...,$$
$$E_6 = 4036 - 11808 \cdot \vartheta - 2355 \cdot \tfrac{1}{2} \cdot \vartheta^2 + 2180 \cdot \tfrac{1}{2} \cdot \vartheta^3, \quad E_6^{(1)} = 14043.67....$$

Now we observe the following, having in mind the above theorem. E_1 can be taken as the first fundamental unit. Since $|E_2^{(i)}| < |E_1^{(i)}|$ for $i = 2, 3, 4$ but $|E_3^{(4)}| > |E_1^{(4)}|$, we can take E_3 as a second fundamental unit (in fact, $E_2 = E_1^2$). Since $\Delta(E_4) = 0$ (in fact, $E_4 = E_1^3$) and $\Delta(E_5) \neq 0$, we have $\mu = E_5$. Then $\mu^{(1)} < E_1^{(1)} \cdot E_3^{(1)}$, so we are in Case 2. Since $E_6^{(1)} > \sqrt{(E_1^{(1)} \cdot E_3^{(1)} \cdot E_5^{(1)})}$, it follows that $\xi \leqq E_6$ (remember that ξ may be a nonunit; in fact, we do not need to know ξ explicitly in our case). Therefore, either we are in Case 2.1 (if $\xi \ll E_6$) and we take E_5 as the third fundamental unit, or the only unit of Case 2.2 is E_6 (if $\xi = E_6$). Since $\Delta(E_6) = 0$ (in fact, $E_6 = E_1 \cdot E_3$), we are in Case 2.2.1, hence again we take E_5 as the third fundamental unit. Thus we have proved that E_1, E_3, E_5 form a system of fundamental units for the order R.

APPENDIX II: *Application of Waldschmidt's Theorem*

In Lemma II.2.3 we have made use of the following result, which is a special case of a theorem due to Waldschmidt [30].

THEOREM. *Let K be a number field with $[K : \mathbb{Q}] = D$. Let $\alpha_1, ..., \alpha_n \in K$, and $b_1, ..., b_n \in \mathbb{Z}$ $(n \geqslant 2)$. Let $V_1, ..., V_n$ be positive real numbers satisfying $1/D \leqslant V_1 \leqslant \cdots \leqslant V_n$ and $V_j \geqslant \max(h(\alpha_j), |\log \alpha_j|/D)$ for $j = 1, ..., n$. Here, $h(\alpha)$ is the absolute logarithmic height, defined by*

$$h(\alpha) = \frac{1}{d} \cdot \log\left(a_0 \cdot \prod_{i=1}^{d} \max(1, |\alpha^{(i)}|)\right),$$

where $d = [\mathbb{Q}(\alpha) : \mathbb{Q}]$, a_0 is the positive leading coefficient of the minimal polynomial of α over $\mathbb{Z}$, and $\alpha^{(i)}$, $i = 1, ..., d$ are the conjugates of α. Let $V_j^+ = \max(V_j, 1)$ for $j = n$, $n-1$, and put

$$\Lambda = b_1 \cdot \log \alpha_1 + \cdots + b_n \cdot \log \alpha_n,$$

where for every $j \in \{1, ..., n\}$, $\log \alpha_j$ is an arbitrary but fixed determination of the logarithm of α_j. Put $B = \max_{1 \leqslant i \leqslant n} |b_i|$. If $\Lambda \neq 0$ then

$$|\Lambda| > \exp(-2^{e(n)} \cdot n^{2n} \cdot D^{n+2} \cdot V_1 \cdot \cdots \cdot V_n \cdot \log(e \cdot D \cdot V_{n-1}^+) \cdot (\log B + \log(e \cdot D \cdot V_n^+))),$$

where $e(n) = \min(8 \cdot n + 51,\ 10 \cdot n + 33,\ 9 \cdot n + 39)$.

We apply this theorem in the case of Λ given by III (2.3). In this case, we compute the V_i's for the various α_i's appearing in Λ, as follows. If $\alpha_i = |\varepsilon_i^{(k)}/\varepsilon_i^{(j)}|$, $i = 1, 2, 3$, then α_i is a unit and hence a_0 (appearing in the

computation of $h(\alpha_i)$) is equal to 1. Clearly, every conjugate of α_i is in absolute value less than

$$H_i = \frac{\max_{1 \leqslant h \leqslant 4} |\varepsilon_i^{(h)}|}{\min_{1 \leqslant h \leqslant 4} |\varepsilon_i^{(h)}|},$$

and $H_i \geqslant 1$. Therefore, $h(\alpha_i) \leqslant H_i$, and we can take

$$V_i = \max(\log H_i, |\log |\varepsilon_i^{(k)}/\varepsilon_i^{(j)}||).$$

Since the latter term equals the logarithm of either $|\varepsilon_i^{(k)}/\varepsilon_i^{(j)}|$ or its inverse, it follows that

$$V_i = \log H_i.$$

If $\alpha_i = |\xi^{(i_0)} - \xi^{(j)}|/|\xi^{(i_0)} - \xi^{(k)}|$, then all conjugates of α_i are in absolute value less than C_3. Therefore, $h(\alpha_i) \leqslant (\log a_0)/d + \log C_3$, where a_0 and d ae as in the definition of $h(\alpha)$ for $\alpha = \alpha_i$. An upper bound for a_0 can be computed as follows. Consider the algebraic numbers $\chi_{ih} = \frac{1}{2} \cdot (\xi^{(i)} - \xi^{(h)})$ for i, $h \in \{1, ..., 4\}$, $i \neq h$. It can be checked that the numbers χ_{ih} are algebraic integers for $\xi = \vartheta$ or φ. Now, for each permutation $\sigma \in S_4$ (we write σ_i instead of $\sigma(i)$), we consider the number $\chi(\sigma) = \chi_{\sigma_1\sigma_2}/\chi_{\sigma_1\sigma_3}$ (independent of σ_4), and the polynomial

$$P(X) = \prod_{\sigma \in S_4} (X - \chi(\sigma)).$$

Consider also the number

$$\Delta = \prod_{1 \leqslant i < h \leqslant 4} \chi_{ih}.$$

Note that

$$\Delta^2 = \frac{1}{2^{12}} \cdot \prod_{1 \leqslant i < h \leqslant 4} (\xi_i - \xi_h)^2 = \frac{1}{2^{12}} \cdot D,$$

where D is the discriminant of the defining polynomial of ξ, and therefore $\Delta^2 = 229$. On the other hand, the coefficients of $P(X)$ are, up to the sign, equal to the elementary symmetric functions of $\chi(\sigma)$, $\sigma \in S_4$, and so they are symmetrical expressions of the $\xi^{(i)}$'s with rational coefficients. This means that $P(X) \in \mathbb{Q}[X]$. On the other hand, by the definition of Δ, any coefficient of $P(X)$ multiplied by Δ^4 is a polynomial of the χ_{ih}'s with coefficients in $\mathbb{Z}$ and therefore it is an algebraic integer. Combine this with the fact that $P(X) \in \mathbb{Q}[X]$ to see that $229^2 \cdot P(X) \in \mathbb{Z}[X]$. Hence, since α_i is a root of $P(X)$, its leading coefficient a_0 is at most 229^2. To conclude, we

have $h(\alpha_i) \leqslant 2 \cdot (\log 229)/d + \log C_3$ and it is clear that $|\log \alpha_i|/d \leqslant \log C_3$. Since $\alpha_i \notin \mathbb{Q}$, we have $d \geqslant 2$ and thus we can take

$$V_i = \log 229 + \log C_3.$$

Simple computations now show that

$$\log H_1 = 4.074586..., \qquad \log H_2 = 5.667432...,$$

$$\log H_3 = 4.821584...,$$

$$\log C_3 = 1.262065... \quad \text{if} \quad \xi = \vartheta,$$

$$\log C_3 = 1.893823... \quad \text{if} \quad \xi = \varphi,$$

$$\log 229 + \log C_3 \leqslant 7.327545....$$

Therefore we apply Wadschmidt's result with $n = 4$, $D \leqslant 24$, $e(n) = 73$,

$$\alpha_1 = \left|\frac{\varepsilon_1^{(k)}}{\varepsilon_1^{(j)}}\right|, \alpha_2 = \left|\frac{\varepsilon_3^{(k)}}{\varepsilon_3^{(j)}}\right|, \alpha_3 = \left|\frac{\varepsilon_2^{(k)}}{\varepsilon_2^{(j)}}\right|, \alpha_4 = \left|\frac{\xi^{(i_0)} - \xi^{(j)}}{\xi^{(i_0)} - \xi^{(k)}}\right|,$$

for $\xi = \vartheta$ or φ, and $b_1 = a_1$, $b_2 = a_3$, $b_3 = a_2$, $b_4 = 1$, $B = A$, $V_1 = \log H_1$, $V_2 = \log H_3$, $V_3 = V_3^+ = \log H_2$, $V_4 = V_4^+ = \log 229 + \log C_3$. Thus we find that

$$|\Lambda| > \exp(-C_7 \cdot (\log A + C_8)),$$

with $C_7 = 5.71 \times 10^{38}$ and $C_8 = 6.17$.

Acknowledgments

The second named author thanks the University of Crete for its hospitality during his visit to Crete in the fall of 1986. He was supported by the Netherlands Foundation for Mathematics (SMC) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).

References

1. M. K. Agrawal, J. H. Coates, D. C. Hunt, and A. J. Van Der Poorten, Elliptic curves of conductor 11, *Math. Comp.* **35** (1980), 991–1002.
2. 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.
3. W. E. H. Berwick, Algebraic number-fields with two independent units, *Proc. London Math. Soc.* **34** (1932), 360–378.
4. K. K. Billevič, On the units of algebraic fields of third and fourth degree, *Mat. Sb.* **40**, **82** (1956), 123–137. [Russian]
5. K. K. Billevič, A theorem on the units of algebraic fields of nth degree, *Mat. Sb.* **64**, **106** (1964), 145–152. [Russian]
6. J. Blass, A. M. W. Glass, D. B. Meronk, and R. P. Steiner, Practical solutions to Thue equations over the rational integers, preprint, Bowling Green State University, 1987.

7. Z. I. Borevich and I. R. Shafarevich, "Number Theory," Academic Press, New York/London, 1973.
8. J. Buchmann, A generalization of Voronoi's unit algorithm I&II, *J. Number Theory* **20** (1986), 177–191 and 192–209.
9. J. Buchmann, The generalized Voronoi algorithm in totally real algebraic number fields, to appear.
10. B. N. Delone and D. K. Faddeev, "The Theory of Irrationalities of the Third Degree," Vol. 10, Transl. of Mathematical Monographs, Amer. Math. Soc., 1964.
11. L. E. Dickson, "Introduction to the Theory of Numbers," Dover, New York, 1957.
12. W. J. Ellison, "Recipes for Solving Diophantine Problems by Baker's Method," Sém. Théorie des Nombres, Univ. de Bordeaux I, 1970–1971, Lab. Th. Nombr. CNRS, Exp. 11.
13. W. J. Ellison, F. Ellison, J. Pesek, C. E. Stahl, and D. S. Stall, The diophantine equation $y^2 + k = x^3$, *J. Number Theory* **4** (1972), 107–117.
14. U. Fincke and M. Pohst, Improved methods for calculating vectors of short length in a lattice, including a complexity analysis, *Math. Comp.* **44** (1985), 463–471.
15. G. H. Hardy and E. M. Wright, "An Introduction to the Theory of Numbers," 5th ed., Oxford Univ. Press, London/New York, 1979.
16. A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász, Factoring polynomials with rational coefficients, *Math. Ann.* **261** (1982), 515–534.
17. S. P. Mohanty, Integer points of $y^2 = x^3 - 4x + 1$, *J. Number Theory* **30** (1988), 86–93.
18. L. J. Mordell, "Diophantine Equations," Academic Press, New York/London, 1969.
19. A. Pethö and R. Schulenberg, Effektives Lösen von Thue Gleichungen, *Publ. Math. Debrecen*, in press.
20. A. Pethö and B. M. M. De Weger, Products of prime powers in binary recurrence sequences. I. The hyperbolic case, with an application to the generalized Ramanujan–Nagell equation, *Math. Comp.* **47** (1986), 713–727.
21. M. Pohst and H. Zassenhaus, On effective computation of fundamental units I & II, *Math. Comp.* **38** (1982), 275–291 and 293–329.
22. T. N. Shorey and R. Tijdeman, "Exponential Diophantine Equations," Cambridge Univ. Press, London/New York, 1986.
23. V. G. Sprindžuk, "Classical Diophantine Equations in Two Unknowns," Nauka, Moscow, 1982. [Russian]
24. R. P. Steiner, On Mordell's equation: A problem of Stolarsky, *Math. Comp.* **46** (1986), 703–714.
25. A. Thue, Über Annäherungswerte algebraischer Zahlen, *J. Reine Angew. Math.* **135** (1909), 284–305.
26. N. Tzanakis, On the diophantine equation $x^2 - Dy^4 = k$, *Acta Arith.* **46** (1986), 257–269.
27. N. Tzanakis, On the practical solution of the Thue equation, an outline, *in* "Proceedings of the Colloquium on Number Theory, Budapest, 1987," in press.
28. N. Tzanakis and B. M. M. De Weger, "On the Practical Solution of the Thue Equation," Memorandum No. 668, Faculty of Applied Mathematics, University of Twente, 1987.
29. N. Tzanakis and B. M. M. De Weger, On the practical solution of the Thue–Mahler equation, in preparation.
30. M. Wadschmidt, A lower bound for linear forms in logarithms, *Acta Arith.* **37** (1980), 257–283.
31. B. M. M. De Weger, Solving exponential diophantine equations using lattice basis reduction algorithms, *J. Number Theory* **26** (1987), 325–367.
32. B. M. M. De Weger, "Algorithms for Diophantine Equations," PhD thesis, University of Leiden, 1987. Also to appear as a CWI Tract, Centre for Mathematics and Computer Science, Amsterdam, 1988.