CHAPTER 8. THE THUE EQUATION.

Acknowledgements. The research for this chapter has been done in cooperation with N. Tzanakis from Iraklion. The results have been published in Tzanakis and de Weger [1987]. See also Tzanakis [1987] and de Weger [1987^b].

8.1. Introduction.

Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be a binary form with integral coefficients, of degree at least three, and irreducible. Let m be a nonzero integer. The diophantine equation

$$F(X,Y) = m$$

in $X, Y \in \mathbb{Z}$ is called a Thue equation. It plays a central role in the theory of diophantine equations. In 1909 Thue proved that it has only finitely many solutions (cf. Thue [1909]). His proof was ineffective. An effective proof was given by Baker [1968]. See Chapter 5 of Shorey and Tijdeman [1986] for a survey of results on Thue equations. By using Lemma 2.4 in Baker's argument, we derive a fully explicit upper bound for the solutions of the Thue equation. Then we show how the methods developed in Chapter 3 can be used to actually find all the solutions of a Thue equation. Our method works in principle for any Thue equation, and in practice for any Thue equation of not too large degree, provided that some algebraic data on the form F are available.

Variants of the method we use here have been used in practice to solve Thue equations by Ellison, Ellison, Pesek, Stahl and Stall [1975], Steiner [1986], Pethö and Schulenberg [1987], and Blass, Glass, Meronk and Steiner [1987^a], [1987^b]. When determining all cubes in the Fibonacci sequence, Pethö [1983] solved a Thue equation by the Gelfond-Baker method, but with a completely different way to find all the solutions below the upper bound.

8.2. From the Thue equation to 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 $n \geq 3$ and let m be a nonzero integer. Consider the Thue equation

$$F(X,Y) = m \ , \tag{8.1}$$

in the unknowns $X, Y \in \mathbb{Z}$. If F is reducible over $\mathbb{Q}$, then (8.1) can be reduced to a system of finitely many equations of type (8.1) 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 ≥ 3 . Then we may assume from now on that F is irreducible over $\mathbb{Q}$. 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 (8.1). Therefore, throughout this chapter we suppose that the algebraic equation $g(x) = 0$ has at least one real root. We number its roots as follows: $\xi^{(1)}, \ldots, \xi^{(s)}$ ($s \geq 1$) are the real roots and $\xi^{(s+1)} = \overline{\xi^{(s+t+1)}}, \ldots, \xi^{(s+t)} = \overline{\xi^{(s+2t)}}$ are the non-real roots, so that we have t (≥ 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 (8.1) into four classes:

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

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

III) the 'large' solutions, with $Y_2 < |Y| \leq Y_3$. 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 8.1 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 8.1. *Let* $X, Y \in \mathbb{Z}$ *satisfy* (8.1). *Put* $\beta = X - \xi \cdot Y$,

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

$$C_1 = \frac{2^{n-1} \cdot |m|}{\min\limits_{1 \le i \le s} |g'(\xi^{(i)})|}, \quad C_2 = \tfrac{1}{2} \cdot \min_{1 \le i < j \le n} |\xi^{(i)} - \xi^{(j)}|,$$

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

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

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

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

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

Proof. Let $i_0 \in \{ 1, \ldots, n \}$ be such that $|\beta^{(i_0)}| = \min\limits_{1 \le i \le n} |\beta^{(i)}|$. We have from (8.1)

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

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

$$|Y| \cdot |\xi^{(i)} - \xi^{(i_0)}| = |\beta^{(i)} - \beta^{(i_0)}| \le |\beta^{(i)}| + |\beta^{(i_0)}| \le 2 \cdot |\beta^{(i)}|.$$

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

$$|\beta^{(i_0)}| = \frac{|m|}{|f_0|} \cdot \prod_{i \ne i_0} |\beta^{(i)}|^{-1} \le \frac{|m|}{|f_0|} \cdot \prod_{i \ne i_0} \left(\tfrac{1}{2} \cdot |Y| \cdot |\xi^{(i)} - \xi^{(i_0)}| \right)^{-1}$$

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

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

$$\left|\frac{X}{Y} - \xi^{(i_0)}\right| = \frac{|\beta^{(i_0)}|}{|Y|} \leq \frac{2^{n-1}\cdot|m|}{\left|g'(\xi^{(i_0)})\right|}\cdot|Y|^{-n}$$

$$\leq \left(\frac{Y_0}{|Y|}\right)^n \cdot \min_{s+1\leq i\leq s+t}|\mathrm{Im}\,\xi^{(i)}| ,$$

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

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

and thus $\left|\frac{X}{Y} - \xi^{(i_0)}\right| < \frac{1}{2}\cdot|Y|^{-2}$, since $\xi^{(i_0)}$ is irrational. Now (ii) follows from a well known result on continued fractions, cf. (3.6). □

Now let $|Y| > Y_1$ and $i_0 \in \{1, \ldots, s\}$ as in Lemma 8.1. Choose $j, k \in \{1, \ldots, n\}$ such that i_0, j, k are pairwise distinct and either $j, k \in \{1, \ldots, s\}$ 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,

$$\frac{\xi^{(i_0)}-\xi^{(j)}}{\xi^{(i_0)}-\xi^{(k)}}\cdot\frac{\beta^{(k)}}{\beta^{(j)}} - 1 = -\frac{\xi^{(k)}-\xi^{(j)}}{\xi^{(k)}-\xi^{(i_0)}}\cdot\frac{\beta^{(i_0)}}{\beta^{(j)}} . \tag{8.2}$$

By Lemma 8.1, the right hand side of (8.2) 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')

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

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

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

LEMMA 8.2. *Put*

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

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

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

$$|\Lambda| < \frac{1.39\cdot C_1 \cdot C_3}{C_2} \cdot |Y|^{-n} .$$

Proof. Consider first the real case. From $|Y| > Y_2^*$ and Lemma 8.1 it follows that the right hand side of (8.2) 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 (8.2) is equal to $e^{\Lambda}-1$, and now (8.2) implies, in view of Lemma 8.1 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 (cf. Lemma 2.2)

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

which proves our claim in the real case.
In the complex case the left hand side of (8.2) is equal to $e^{i\Lambda}-1$, and, as in the real case, we derive

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

Since $|e^{i\Lambda}-1| = 2 \cdot |\sin \Lambda/2|$, it follows that $|\sin \Lambda/2| < \frac{1}{4}$, and therefore by Lemma 2.3

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

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 $\epsilon_1, \ldots, \epsilon_r$, 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 [1932], Billevič [1956], [1964], Pohst and Zassenhaus [1982], Buchmann [1986], [1987]). We simply assume that a system of fundamental units is known. On the other hand, there exist only finitely many non-associates $\mu_1, \ldots, \mu_\nu$ in K such that $f_0 \cdot N(\mu_i) = m$ for $i = 1, \ldots, \nu$. (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 $\zeta \cdot \mu_i$, where ζ is a root of unity in K. (In the important case $|f_0| = |m| = 1$, it is clear that $M = \{ -1, 1 \}$). Then, for any integral solution (X,Y) of (8.1) there exist some $\mu \in M$ and $a_1, \ldots, a_r \in \mathbb{Z}$, such that

$$\beta = \mu \cdot \epsilon_1^{a_1} \cdot \ldots \cdot \epsilon_r^{a_r} .$$

Thus, the initial problem of solving (8.1) is reduced to that of finding all integral r-tuples $(a_1, \ldots, a_r)$ such that $\mu \cdot \epsilon_1^{a_1} \cdot \ldots \cdot \epsilon_r^{a_r}$ 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 (8.2). Thus the problem reduces to solving finitely many equations of the type

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

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

$$\Lambda = \log \left| \frac{\xi^{(i_0)} - \xi^{(j)}}{\xi^{(i_0)} - \xi^{(k)}} \cdot \frac{\mu^{(k)}}{\mu^{(j)}} \right| + \sum_{i=1}^{r} a_i \cdot \log \left| \frac{\epsilon_i^{(k)}}{\epsilon_i^{(j)}} \right| , \tag{8.3}$$

and in the complex case

$$\Lambda = \mathrm{Arg} \left(\frac{\xi^{(i_0)} - \xi^{(j)}}{\xi^{(i_0)} - \xi^{(k)}} \cdot \frac{\mu^{(k)}}{\mu^{(j)}} \right) + \sum_{i=1}^{r} a_i \cdot \mathrm{Arg} \left(\frac{\epsilon_i^{(k)}}{\epsilon_i^{(j)}} \right) + a_0 \cdot 2\pi , \tag{8.4}$$

with $a_0 \in \mathbb{Z}$, and $-\pi < \mathrm{Arg}(z) \leq \pi$ 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 8.2 we can find an explicit upper bound for $\max|a_i|$. This is what we do in the next section.

8.3. Upper bounds.

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

LEMMA 8.3. *Let* $I = \{ h_1, \ldots, h_r \} \subset \{ 1, \ldots, n \}$. *Put*

$$U_I = \left(\log|\epsilon_\ell^{(h_i)}| \right)_{1 \leq i \leq r, 1 \leq \ell \leq r} ,$$

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

$$U_I^{-1} = (u_{i\ell}) , \quad N[U_I^{-1}] = \max_{1 \leq i \leq r} \sum_{\ell=1}^{r} |u_{i\ell}| .$$

Put also

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

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

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

Then, for

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

we have

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

Proof. By $\beta = \mu\cdot\epsilon_1^{a_1}\cdot\ldots\cdot\epsilon_r^{a_r}$ we have

$$\begin{pmatrix} \log|\beta^{(h_1)}/\mu^{(h_1)}| \\ \vdots \\ \log|\beta^{(h_r)}/\mu^{(h_r)}| \end{pmatrix} = U_I\cdot\begin{pmatrix} a_1 \\ \vdots \\ a_r \end{pmatrix} . \tag{8.5}$$

On the other hand, for every $h \in \{ 1, \ldots, n \}$, using the end of the proof of Lemma 8.1,

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

and therefore

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

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

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

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

$$C_4\cdot|Y| \ge \left(\tfrac{1}{2} + \max_{1\le i_1<i_2\le n}|\xi^{(i_1)}-\xi^{(i_2)}| \right)\cdot|Y|\cdot|m|^{-1/n} > \frac{|Y|}{2|m|^{1/n}} > 1 .$$

Then,

$$\log\left|\frac{\beta^{(h)}}{\mu^{(h)}}\right| < \log\left(C_4\cdot|Y|\right) \text{ for } h = 1,\ldots,n , \quad \log\left(C_4\cdot|Y|\right) > 0 . \tag{8.6}$$

Next we show that

$$\left|\log\left|\frac{\beta^{(i)}}{\mu^{(i)}}\right|\right| < (n-1)\cdot\log(C_4\cdot|Y|) \quad \text{for} \quad i = 1, \ldots, n . \tag{8.7}$$

Indeed, in view of (8.6), a stronger inequality is true if $|\beta^{(i)}/\mu^{(i)}| \geq 1$. Suppose now that $|\beta^{(i)}/\mu^{(i)}| < 1$. By

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

it follows that

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

in view of (8.6). Now the inequality

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

follows from (8.5), (8.7), 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 8.1, for every $h \in I$, $|\beta^{(h)}/\mu^{(h)}| > C_2\cdot|Y|/\mu_+ > 1$ and now, in view of (8.6),

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

which implies our assertion. □

Lemmas 8.2 and 8.3 immediately yield

LEMMA 8.4. *Put*

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

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

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

Next we apply Lemma 2.4 (Waldschmidt). It yields in the real case (assuming that $\Lambda \neq 0$)

$$|\Lambda| > \exp\{-C_7 \cdot (\log A + C_8)\} , \tag{8.8}$$

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

The precise values for C_7 and C_8 are given in Section 2.3. It should be noted that in the complex case a_0 makes now its appearance, while it was not present in Lemmas 8.3 and 8.4. 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

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

we find from (8.4) and the proof of lemma 8.2 that $|a_0| < \frac{1}{2} + \frac{1}{2} \cdot r \cdot A + |\Lambda|/2\pi < 1 + r \cdot A \le r \cdot A$ if $A \ge 2$. Thus we may apply (8.8) 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 8.5. *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 8.2, $|e^{\Lambda} - 1| < \frac{1}{2}$ in the real case, and $|e^{i\Lambda} - 1| < \frac{1}{2}$ in the complex case. Note that $\beta^{(i_0)} \neq 0$. Hence (8.2) implies $\Lambda \neq 0$. Therefore Lemma 8.4 and (8.8) yield

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

The result now follows from Lemma 2.1. □

Remark. From this upper bound for A an upper bound for $|Y|$ can be derived, thus a value for Y_3 (cf. Section 8.2). We shall not do this. Note that Y_2' (cf. Lemma 8.4) is not necessarily equal to Y_2 (cf. Section 8.2).

8.4. Reducing the upper bound.

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

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

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

$$|\Lambda| < K_1 \cdot \exp\left(-K_2 \cdot A\right) , \quad A < K_3 . \tag{8.9}$$

In our case, it follows from (8.3) or (8.4) how to define q, δ and the μ_i's , and from Lemmas 8.4 and 8.5 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

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

so that $\Lambda = \delta + \Lambda_0$. We apply the methods of Chapter 3 to problem (8.9).

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

(i). Let $\delta = 0$, so that $\Lambda = \Lambda_0$. Then the linear form is homogeneous, and we apply the method of Section 3.7.

(ii) Let $\delta \neq 0$. Then the linear form is inhomogeneous, and we apply the method of Section 3.8.

(iii). Suppose now that the μ_i's are $\mathbb{Q}$-dependent. Let Γ be the approximation lattice for the linear form Λ , as defined in Section 3.7. Then we expect the lower bound for $|\underline{x}|$ ($\underline{x} \in \Gamma$, $\underline{x} \neq \underline{0}$) in general to be 'very small', since the vector having as coordinates the coefficients of the dependence relation will give rise to a very short vector in the lattice. So the reduction process, as applied in the two previous cases, will not work. In such a case we work as follows. Let M be a maximal subset of $\{\mu_1, \ldots, \mu_q\}$ consisting of $\mathbb{Q}$-independent numbers. With an appropriate choice of subscripts we may assume that $M = \{ \mu_1, \ldots, \mu_p \}$, $p < q$. Then we can find integers $d > 0$ and d_{ij} for $1 \leq i \leq p$, $p+1 \leq j \leq q$ such that

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

(These numbers d, d_{ij} can be found as coordinates of extremely short vectors in reduced bases). On the other hand, (8.9) is equivalent to

$$|\Lambda'| < K_1' \cdot \exp\left(-K_2 \cdot A\right) \quad , \quad A < K_3 \quad , \tag{8.10}$$

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

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

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

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

Therefore, if we put $A' = \max\limits_{1 \le i \le p} |a_i'|$, then $A' \le (q-p+1) \cdot D \cdot A$, and (8.10) implies

$$|\Lambda'| < K_1' \cdot \exp\left(-K_2' \cdot A'\right) \quad , \quad A' < K_3' \quad , \tag{8.11}$$

where

$$\Lambda' = \delta' + a_1' \cdot \mu_1' + \ldots + a_p' \cdot \mu_p' \quad , \quad K_1' = d \cdot K_1 \quad ,$$

$$K_2' = K_2 / (q-1+p) \cdot D \quad , \quad K_3' = (q-p+1) \cdot K_3 \quad .$$

Now, to solve (8.11) 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 necessary, until we find a very small upper bound for A' . After having found all solutions $(a_1', \ldots, a_p')$ of (8.11), we have a lower bound $L > 0$ for $|\Lambda'|$. It is reasonable to expect that L is not 'extremely small', because the integers $a_1', \ldots, a_p'$ being 'small' in absolute value cannot make $|\Lambda'|$ 'extremely small'. Now combine $|\Lambda'| \ge L$ with the first inequality of (8.10) to get

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

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 (found after some reduction steps as described above) 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 the technique of Fincke and Pohst, cf. Section 3.6. However, when solving a Thue equation, and not only an inequality for a linear form in logarithms, it may be better to avoid this method, and to use continued fractions of the roots $\xi^{(i)}$. In practice we can search for the solutions (X,Y) of (8.1) satisfying $Y_1 < |Y| \le C$ as follows, referring to Lemma 8.1. Here e.g. $C = Y_2$, and we can imagine C here as being a 'large' constant compared to Y_1 , but not 'very large' (cf. the introduction of Y_1, Y_2 in Section 8.2).

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

$$|\tilde{\xi}-\xi^{(i_0)}| < \frac{1}{6\cdot C^2} . \tag{8.12}$$

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 $a_{k+1} \ge 3$. Indeed, we have by (3.5)

$$\frac{1}{(a_{k+1}+2)\cdot|Y|^2} \le \frac{1}{(a_{k+1}+2)\cdot q_k^2} < |\xi^{(i_0)} - \frac{p_k}{q_k}| = |\xi^{(i_0)} - \frac{X}{Y}| \le \frac{C_1}{|Y|^n} .$$

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, $a_{k+1} \ge 3$, and by (3.5) we have

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

Therefore,

$$|\tilde{\xi} - \frac{p_k}{q_k}| \le |\tilde{\xi}-\xi^{(i_0)}| + |\xi^{(i_0)} - \frac{p_k}{q_k}| < \frac{1}{6\cdot C^2} + \frac{1}{3\cdot q_k^2} \le \frac{1}{2\cdot q_k^2}$$

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

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

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

$$a_{k+1} > \frac{|q_k|^{n-2}}{C_1} - 2 \ . \tag{8.13}$$

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

To sum up, we propose the following process for every real root $\xi^{(i_0)}$ for $i_0 = 1, \ldots, s$ (note that i_0 is a priori not known). (1) Compute a rational approximation $\tilde{\xi}$ of $\xi^{(i_0)}$ satisfying (8.12) (a truncation of its decimal expansion will do). (2) Expand $\tilde{\xi}$ into its continued fraction with partial quotients $a_0, a_1, a_2, \ldots, a_{k+1}$ and convergents p_i/q_i for all $i = 1, \ldots, k$ with $q_k \le C < q_{k+1}$. (3) Test all these convergents for the conditions (8.13) 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 $Z^n \mid m$. This simple observation excludes in general most of the reducible quotients X/Y , and all of them if m is an n-th-powerfree integer.

Having tested for all solutions in the range $|Y| \le C$ 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, Glass, Meronk and Steiner [1987[b]]). For every $(i,j) \in \{1,\ldots,r\} \times \{1,\ldots,n\}$ let φ_{ij} be the number $+1$ or -1 for which $|\epsilon_i^{(j)}|^{\nu_{ij}} \ge 1$, and put $E_j = \prod_{i=1}^{r} |\epsilon_i^{(j)}|^{\varphi_{ij}}$. Then

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

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

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

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

8.5. An application: integral points on the elliptic curve $y^2 = x^3 - 4\cdot x + 1$.

In this section we prove, as an application of the general theory described in the previous sections, the following result.

THEOREM 8.6. *The elliptic curve*

$$y^2 = x^3 - 4\cdot x + 1 \tag{8.14}$$

has only the following 22 integral points:

$$(x,\pm y) = (-2,1),\ (-1,2),\ (0,1),\ (2,1),\ (3,4),\ (4,7),\ (10,31),$$

$$(12,41),\ (20,89),\ (114,1217),\ (1274,45473)\ .$$

We prove this theorem 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 the previous sections.

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

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

Let the conjugates of ψ be $\psi^{(1)} = 0.254\ldots$, $\psi^{(2)} = -2.114\ldots$, $\psi^{(3)} = 1.860\ldots$. From a table of Delone and Faddeev ([1964], 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 independent units is ψ, $2 - \psi$. From Table I of Buchmann [1986] 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 (8.14) of the elliptic curve can be written as

$$y^2 = (x - \psi)\cdot(x^2 + x\cdot\psi + (\psi^2-4)) \tag{8.15}$$

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

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

which is prime, since its norm is -229. Therefore we would have that π is a unit times this prime, and then by (8.15), $x - \psi = \text{unit} \times (3 \cdot \psi^2 - 4) \times \text{square}$. Take norms, then we get $y^2 = \pm 229 \times \text{square}$, which is clearly impossible.

Now (8.15) implies

$$x - \psi = \pm \psi^i \cdot (2-\psi)^j \cdot \alpha^2 \ , \quad \alpha \in \mathbb{Z}[\psi] \ , \quad i, j \in \{ 0, 1 \} \ . \tag{8.16}$$

Since (8.14) is trivial to solve for $x \leq 0$ (the only solutions with $x \leq 0$ are the first three pairs stated in the theorem), we may assume that $x \geq 1$. Since $\psi^{(1)} = 0.254\ldots$, we see that the minus sign in (8.16) is impossible. Then, by $\psi^{(2)} = -2.114\ldots$, $i \neq 1$. We conclude therefore that

$$x - \psi = (2-\psi)^j \cdot (u+v \cdot \psi+w \cdot \psi^2)^2 \ , \quad u, v, w \in \mathbb{Z} \ , \ j \in \{ 0, 1 \} \ . \tag{8.17}$$

First case: $j = 0$. Then (8.17) implies, on equating corresponding coefficients in both sides,

$$x = u^2 - 2 \cdot v \cdot w, \ w^2 - 2 \cdot u \cdot v - 8 \cdot v \cdot w = 1, \ v^2 + 4 \cdot w^2 + 2 \cdot u \cdot w = 0 \ . \tag{8.18}$$

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 (8.18) 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

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

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

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 $u_1 = 2 \cdot u_2$, $z = 2 \, z_1$. Then $u_2^2 - v_1^2 = z_1^2$, where u_2 and v_1 are odd. By $u_2 \geq 0$ there exist $m, n \in \mathbb{Z}$ such that

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

It follows that

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

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 (8.18) 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 (8.18) we find $x = 20, 12$ respectively, which furnish the solutions $(x,\pm y) = (20,89), (12,41)$ for (8.14).

Secondly, we suppose that u_1 and z are odd. Then v_1 is even, so by $u_1 \geq 0$ 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

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

We may assume that $w = -m^2$. Substituting this in the second equation of (8.18) 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 (8.14).

Second case: $j = 1$. Then, equating the coefficients in (8.17) we get

$$x = 2\cdot u^2 + v^2 + 4\cdot w^2 + 2\cdot u\cdot w - 4\cdot v\cdot w \ , \tag{8.19}$$

$$\begin{cases} u^2 + 4\cdot v^2 + 18\cdot w^2 - 4\cdot u\cdot v + 8\cdot u\cdot w - 18\cdot v\cdot w = 1 \ , \\ \quad\ 2\cdot v^2 + \ 9\cdot w^2 - 2\cdot u\cdot v + 4\cdot u\cdot w - \ 8\cdot v\cdot w = 0 \ . \end{cases} \tag{8.20}$$

The first relation of (8.20) can be replaced by

$$u^2 - 2\cdot v\cdot w = 1 \ . \tag{8.21}$$

Note that u is odd. Put $z = v - 2\cdot w$. Then the second equation of (8.20) 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

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

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

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

Substituting this in (8.21) we obtain the Thue equation

$$m^4 - 4\cdot m^3\cdot n - 12\cdot m^2\cdot n^2 + 4\cdot n^4 = 1 \ . \tag{8.22}$$

In Theorem 8.7 below 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 (8.14) to $(x,\pm y) = (3,4)$.

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

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

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 (8.21), we obtain the Thue equation

$$n^4 - 12\cdot n^2\cdot m^2 - 8\cdot n\cdot m^3 + 4\cdot m^4 = 1 \ . \tag{8.23}$$

In Theorem 8.7 below 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 (8.14) to the solutions $(x,\pm y) = (2,1)$, $(10,31)$, $(1274,45473)$, $(114,1217)$. Thus this result completes the proof of theorem 8.6, provided the Thue equations (8.22), (8.23) have as their only solutions the pairs (m,n) mentioned above. We now proceed to prove this.

THEOREM 8.7. (i). *The Thue equation*

$$X^4 - 4\cdot X^3\cdot Y - 12\cdot X^2\cdot Y^2 + 4\cdot Y^4 = 1 \tag{8.24}$$

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

(ii). *The Thue equation*

$$X^4 - 12\cdot X^2\cdot Y^2 - 8\cdot X\cdot Y^3 + 4\cdot Y^4 = 1 \tag{8.25}$$

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

Proof. We use the notation and results of Sections 8.2 and 8.3.

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 Section 8.2 we have $n = 4$, $s = 4$, $t = 0$, and $\xi = \vartheta$ or $\xi = \varphi$. Simple computations show that for $\xi = \vartheta,\ \varphi$ we can take

$$Y_0 = 1\ ,\quad C_1 = 0.843\ ,\quad C_2 = 0.589\ ,\quad Y_1 = 2,\quad C_3 = 6.645\ ,$$

$$Y_2^* = 3\ ,\quad \mu_- = \mu_+ = 1\ ,\quad C_4 = 8.3374\ .$$

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 :

$$\vartheta^{(1)} \cong -1.080\ 286\ 352\ ,\quad \varphi^{(1)} \cong -1.851\ 360\ 980\ ,$$

$$\vartheta^{(2)} \cong\ \ 3.722\ 935\ 260\ ,\quad \varphi^{(2)} \cong\ \ 0.537\ 210\ 524\ ,$$

$$\vartheta^{(3)} \cong\ \ 0.334\ 111\ 716\ ,\quad \varphi^{(3)} \cong\ \ 5.986\ 021\ 747\ ,$$

$$\vartheta^{(4)} \cong -2.976\ 760\ 624\ ,\quad \varphi^{(4)} \cong -0.671\ 871\ 290\ .$$

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

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

On the other hand, (8.24) and (8.25) 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 (8.24) or (8.25), then $X - Y\cdot\vartheta$ or $X - Y\cdot\varphi$, respectively, is a unit of the order R . A system of fundamental units of R is given by

$$\epsilon_1 = 1 + \vartheta \ , \quad \epsilon_2 = 3 + \vartheta \ , \quad \epsilon_3 = \tfrac{1}{2}\cdot\vartheta^2 \ .$$

We do not prove this fact here. For a proof, see Tzanakis and de Weger [1987] Section III.2 and Appendix I.

Thus the solution of (8.24) and (8.25) is reduced to finding all $(a_1,a_2,a_3) \in \mathbb{Z}^3$ such that the unit $\pm\epsilon_1^{a_1}\cdot\epsilon_2^{a_2}\cdot\epsilon_3^{a_3}$ has the special shape $X - Y\cdot\vartheta$ or $X - Y\cdot\varphi$, respectively. In the notation of Lemma 8.3 we have, after some numerical computations, that we leave to the reader to check, that

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

(here, of course, $I = \{ 1, 2, 3, 4 \}$). Therefore we can take in Lemma 8.4

$$C_5 = 1.211 \ .$$

Also,

$$C_6 = 6.38771\times10^4 \ , \quad Y_2' = 3 \ .$$

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

Now, relation (8.3) becomes in our case

$$\Lambda = \log\left|\frac{\xi^{(i_0)}-\xi^{(j)}}{\xi^{(i_0)}-\xi^{(k)}}\right| + \sum_{i=1}^{3} a_i\cdot\log\left|\frac{\epsilon_i^{(k)}}{\epsilon_i^{(j)}}\right| \ , \tag{8.26}$$

where $\xi = \vartheta$ or φ. As mentioned in Section 2, once i_0 is fixed, we can choose j, k arbitrarily. Thus we can choose

$$\begin{cases} j = 3, \ k = 4 & \text{if } i_0 = 1 \text{ or } 2 \ , \\ j = 1, \ k = 2 & \text{if } i_0 = 3 \text{ or } 4 \ . \end{cases} \tag{8.27}$$

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 below,

$$C_7 = 5.71\times10^{38} \ , \quad C_8 = 6.17 \ ,$$

and therefore, by Lemma 8.5, if $|Y| > 3$, then for $A = \max_{1\le i\le 3} |a_i|$ we have the upper bound $C_9 = 3.26\times10^{40}$. As is easily checked, the only solutions of either (8.24) or (8.25) with $|Y| \le 3$ are those listed in the statement of

the theorem. Therefore we may assume that $|Y| > 3$, so that

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

Before we apply the reduction method of Section 3.8 we show that the application of Lemma 2.4 yields the above constants C_7, C_8. We apply this result in the case of Λ given by (8.26). In this case, we compute the V_i's for the various α_i's appearing in Λ, as follows. If $\alpha_i = |\epsilon_i^{(k)}/\epsilon_i^{(j)}|$ for $i = 1, 2, 3$, 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

$$H_i = \frac{\max\limits_{1\le h\le 4} |\epsilon_i^{(h)}|}{\min\limits_{1\le h\le 4} |\epsilon_i^{(h)}|} ,$$

and $H_i \ge 1$. Therefore, $h(\alpha_i) \le H_i$, and we can take

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

Since the latter term equals the logarithm of either $|\epsilon_i^{(k)}/\epsilon_i^{(j)}|$ 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, $h(\alpha_i) \le (\log a_0)/d + \log C_3$, where a_0 and d are 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 $i, h \in \{ 1, \ldots, 4 \}$ with $i \ne h$. It can be checked that the numbers χ_{ih} are algebraic integers for $\xi = \vartheta$ or φ. Now, for each permutation $\sigma = (\sigma_1\sigma_2\sigma_3\sigma_4) \in S_4$ 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

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

Note that

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

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 the elementary symmetric functions of $\chi(\sigma)$ for $\sigma \in S_4$, 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 $h(\alpha_i) \le 2 \cdot (\log 229)/d + \log C_3$ and it is clear that $|\log \alpha_i|/d \le \log C_3$. Since $\alpha_i \notin \mathbb{Q}$ we have $d \ge 2$, so we can take

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

Simple computations now show that

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

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

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

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

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

Therefore we apply Lemma 2.4 (Waldschmidt) with $n = 4$, $D \le 24$, $e(n) = 73$,

$$\alpha_1 = \left| \frac{\epsilon_1^{(k)}}{\epsilon_1^{(j)}} \right| \ , \quad \alpha_2 = \left| \frac{\epsilon_3^{(k)}}{\epsilon_3^{(j)}} \right| \ , \quad \alpha_3 = \left| \frac{\epsilon_2^{(k)}}{\epsilon_2^{(j)}} \right| \ , \quad \alpha_4 = \left| \frac{\xi^{(i_0)} - \xi^{(j)}}{\xi^{(i_0)} - \xi^{(k)}} \right| \ ,$$

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

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

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

We have now to apply the reduction process described in Section 3.7. In our

situation we have to solve (8.9) with

$$K_1 = C_6 = 6.38771\times10^4 \ , \quad K_2 = \frac{n}{C_5} = \frac{4}{1.211} > 3.303 \ , \quad K_3 = 3.26\times10^{40}$$

(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 (8.26) and (8.27):

$$\begin{cases} \delta = \delta_1 := \log\left|\dfrac{\xi^{(1)}-\xi^{(3)}}{\xi^{(1)}-\xi^{(4)}}\right| \quad \text{or} \quad \delta = \delta_2 := \log\left|\dfrac{\xi^{(2)}-\xi^{(3)}}{\xi^{(2)}-\xi^{(4)}}\right| \ , \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{where} \quad \xi = \vartheta \ \text{ or } \ \varphi \ , \\ \mu_i = \log\left|\dfrac{\epsilon_i^{(4)}}{\epsilon_i^{(3)}}\right| \ , \quad \text{for} \quad i = 1, 2, 3 \ , \end{cases} \tag{8.28}$$

or

$$\begin{cases} \delta = \delta_3 := \log\left|\dfrac{\xi^{(3)}-\xi^{(1)}}{\xi^{(3)}-\xi^{(2)}}\right| \quad \text{or} \quad \delta = \delta_4 := \log\left|\dfrac{\xi^{(4)}-\xi^{(1)}}{\xi^{(4)}-\xi^{(2)}}\right| \ , \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{where} \quad \xi = \vartheta \ \text{ or } \ \varphi \ , \\ \mu_i = \log\left|\dfrac{\epsilon_i^{(2)}}{\epsilon_i^{(1)}}\right| \ , \quad \text{for} \quad i = 1, 2, 3 \ . \end{cases} \tag{8.29}$$

Numerical details are given in Tzanakis and de Weger [1987]. We take $c_0 = 10^{140}$, and we work with the lattice with associated matrix

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

Note that in each of the four cases of (8.28) (resp. (8.29)) we have the same lattice, Γ_1 (resp. Γ_2), say. In each 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 Section 8.4.

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 $\mathcal{B}$, $\mathcal{U}$, $\mathcal{U}^{-1}$, as defined in Section 3.7. In our cases, the coordinates of the vectors of the reduced bases (i.e. the elements of $\mathcal{B}$) 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

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

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

$$|\underline{b}_1| > 3.247\times10^{46} \quad \text{in the case of lattice } \Gamma_1 ,$$

$$|\underline{b}_1| > 4.846\times10^{46} \quad \text{in the case of lattice } \Gamma_2 ,$$

$$\|s_3\| > 0.029 \quad \text{in all 8 cases.}$$

This means that in view of Lemma 3.5, in all cases $i_0 = 3$, and

$$\ell(\Gamma_i, \underline{x}) > 0.029 \cdot \tfrac{1}{2} \cdot 3.247\times10^{46} > 4.708\times10^{44} .$$

Then the assumptions of Lemma 3.10 are fulfilled with $n = 3$, $\gamma = 1$, $C = c_0$, $c = K_1$, $\delta = K_2$, $X_0 = X_1 = K_3$, since $\sqrt{27}\cdot K_3 < 1.112\times10^{40}$, which implies

$$A < \frac{1}{3.303} \cdot \log\left(10^{140} \cdot 6.38771\times10^{4} / 3.26\times10^{40}\right) < 72.8 .$$

It follows that $A \le 72$. We repeat the procedure with $K_3 = 72$ and $c_0 = 10^{12}$. We found from our computations

$$|\underline{b}_1| > 1.293\times10^{4} \quad \text{in the case of lattice } \Gamma_1 ,$$

$$|\underline{b}_1| > 1.092\times10^{4} \quad \text{in the case of lattice } \Gamma_2 ,$$

$$\|s_3\| > 0.143 \quad \text{in all 8 cases.}$$

This means that in view of Lemma 3.5, in all cases $i_0 = 3$, and

$$\ell(\Gamma_i, \underline{x}) > 0.143 \cdot \tfrac{1}{2} \cdot 1.092\times10^{4} > 7.807\times10^{2} .$$

Then the assumptions of Lemma 3.10 are fulfilled, since $\sqrt{27}\cdot K_3 < 3.742\times10^{2}$, which implies

$$A < \frac{1}{3.303} \cdot \log\left(10^{12} \cdot 6.38771\times10^{4} / 72\right) < 10.5 .$$

It follows that $A \le 10$. We enumerated all remaining possibilities, and found no other solutions of (8.24) and (8.25) than mentioned in the theorem. This completes the proof of Theorem 8.7, hence also that of Theorem 8.6. □

The computations for the proof of Theorem 8.7 took 35 sec.

8.6. The Thue-Mahler equation, an outline.

Let $F(X,Y)$ be as in Section 8.1. Let $p_1, \ldots, p_s$ be fixed distinct prime numbers. The diophantine equation

$$F(X,Y) = \pm \prod_{i=1}^{s} p_i^{n_i}$$

in the variables $X, Y \in \mathbb{Z}$, $n_1, \ldots, n_s \in \mathbb{N}_0$, with $(X,Y) = 1$, is known as a Thue-Mahler equation. It was proved by Mahler [1933] that this equation has only finitely many solutions, and by Coates [1970] that they can, at least in principle, be determined effectively, since an effectively computable upper bound for the variables can be derived from the p-adic theory of linear forms in logarithms. For the history of this equation we refer to Shorey and Tijdeman [1986], Chapter 7.

We believe that it is possible to solve Thue-Mahler equations, not only in principle, but in practice. This can be done by reducing the above mentioned upper bounds, using a combination of real and p-adic computational diophantine approximation techniques, based on the L^3-algorithm for reducing bases of lattices (cf. Sections 3.7, 3.8, 3.11 and 3.12). The method can be considered as a p-adic analogue of the method for solving Thue equations, on which we report in the preceding sections.

Such an idea (but without using the L^3-algorithm) was used by Agrawal, Coates, Hunt and van der Poorten [1980], who determined all solutions of the equation

$$X^3 - X^2 \cdot Y + X \cdot Y^2 + Y^3 = \pm 11^n .$$

This is one of the only two examples in the literature where a Thue-Mahler equation has been solved completely, the other one being

$$X^3 + 3 \cdot Y^3 = 2^n ,$$

which was solved by Tzanakis [1984] by a different method. Both examples are of the simplest kind, in view of the fact that the cubic field $\mathbb{Q}(\vartheta)$, where ϑ is a root of $F(x,1) = 0$, has only one fundamental unit, and there occurs only one prime. Therefore it is sufficient to use two-dimensional real continued fractions and one-dimensional p-adic continued fractions, instead of the more complicated L^3-algorithm (which was not yet available in 1980). With the use of the L^3-algorithm the method can in principle be extended to

the general situation, where there are more than one fundamental units, and more than one primes. In a forthcoming publication, Tzanakis and the present author plan to give details and worked-out examples.