ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 40, Number 2, 2010

ON THUE EQUATIONS OF SPLITTING TYPE OVER FUNCTION FIELDS

VOLKER ZIEGLER

ABSTRACT. In this paper we consider Thue equations of splitting type over the ring $k[T]$, i.e., they have the form

$$X(X-p_1Y)\cdots(X-p_{d-1}Y)-Y^d=\xi,$$

with $p_1,\dots,p_{d-1}\in k[T]$ and $\xi\in k$. In particular, we show that such Thue equations have only trivial solutions provided the degree of p_{d-1} is large, with respect to the degree of the other parameters $p_1,\dots,p_{d-2}$.

1. Introduction. Let $F\in\mathbf{Z}[X,Y]$ be a homogeneous, irreducible polynomial of degree $d\geq 3$. Then the Diophantine equation

$$F(X,Y)=m,\quad m\in\mathbf{Z}\setminus\{0\}$$

is called a Thue equation in honor of Axel Thue [**23**] who proved the finiteness of the number of solutions. Since then several Thue equations and also families of Thue equations were solved. In particular, families of Thue equations of the form

$$X(X-a_1Y)\cdots(X-a_{d-1}Y)+Y^d=\pm 1, \tag{1}$$

with $a_1,\dots,a_{d-1}$ were studied by several authors, e.g., Heuberger [**9**], Lee [**12**], Mignotte and Tzanakis [**16**], Pethő [**18**], Pethő and Tichy [**19**], Thomas [**22**] and Wakabayshi [**24**]. This type of Thue equation is called splitting type. Obviously these Thue equations have solutions $\pm(1,0),\pm(0,1),\pm(a_1,1),\dots,\pm(a_{d-1},1)$, which are called trivial. Thomas [**22**] investigated Thue equations of splitting type of degree $d=3$ with $a_1=p_1(n)$, $a_2=p_2(n)$, where p_1,p_2 are monic polynomials

2010 AMS *Mathematics subject classification.* Primary 11D59, 11D25, 11Y50.
Keywords and phrases. Thue equations, Thomas' conjecture, Function fields.
The author is supported by the Austrian Science Fund FWF (project number: S9603).
Received by the editors on August 21, 2007, and in revised form on October 15, 2007.
DOI:10.1216/RMJ-2010-40-2-723

with $\deg(p_1) < \deg(p_2)$ and $n \in \mathbf{Z}$ large. Under some complicated degree conditions for p_1 and p_2, Thomas proved that the trivial solutions are the only solutions provided $n \in \mathbf{Z}$ is large. These investigations led him to the conjecture that for $a_1 = p_1(n), \dots, a_{d-1} = p_{d-1}(n)$, and n sufficiently large, $p_1, \dots, p_{d-1}$ monic polynomials with $\deg(p_1) < \cdots < \deg(p_{d-1})$ Diophantine equation (1) has only trivial solutions. This conjecture was finally settled by Heuberger [**10**] under some complicated degree conditions. However, counterexamples exist to Thomas's conjecture for $d = 3$. Ziegler [**25**] observed that the Thue equations

$$X(X - nY)(X - (n^4 + 3n)Y) + Y^3 = \pm 1$$

and

$$X(X - nY)(X - (n^4 - 2n)Y) + Y^3 = \pm 1$$

have the nontrivial solutions $\pm(n^9 + 3n^6 + 4n^3 + 1, n^8 + 3n^5 + 3n^2)$ and $\pm(n^3 - 1, -n^8 + 3n^5 - 3n^2)$, respectively. These counterexamples were found by solving Thue equation (1) for $d = 3$ over the function field $\mathbf{C}(T)$, i.e., assume $X, Y, a_1, a_2 \in \mathbf{C}[T]$ (see [**25**]).

Thue equations over function fields were investigated by Gill [**7**], Osgood [**17**], Schmidt [**21**], Mason [**14**] (see also [**15**]), Lettl [**13**] and many others. Also the case of global function fields has been considered by Gaál and Pohst [**5, 6**] recently. Families of Thue equations were investigated by Fuchs and Ziegler [**3, 4**] and Fuchs and Jadrijević [**2**]. A first attempt to prove a function field analogon of Thomas's conjecture was made by Ziegler [**25**], who considered equation (1) in the case of $d = 3$. The purpose of this paper is to prove an analogon of Thomas's conjecture for general d. Therefore, we consider the equation

$$X(X - p_1Y) \cdots (X - p_{d-1}Y) + Y^d = \xi, \tag{2}$$

over $k[T]$, where $p_1, \dots, p_{d-1} \in k[T]$, $\xi \in k$ and k an algebraic closed field of characteristic 0. In particular we prove the following theorem:

Theorem 1. *Let* $0 < \deg(p_1) < \cdots < \deg(p_{d-2}) < \deg(p_{d-1})$ *and assume* $c_d \deg(p_{d-2}) < \deg(p_{d-1})$ *with* $c_d = 1.031d(d-1)(d-1)!(2d-3)4^{d-1}$. *Then Thue equation* (2) *has only trivial solutions* $\zeta(1, 0)$, $\zeta(0, 1)$, $\zeta(p_1, 1), \dots, \zeta(p_{d-1}, 1)$, *with* $\zeta^d = \xi$.

This is an analogon of a result of Halter-Koch, et al. [**8**] for function fields. The plan of the paper is as follows. In the next section we give a short overview of the tools we need for a proof of the theorems. Then we investigate the unit structure of the relevant function fields (see Section 3). With the knowledge of Section 3 we are able to adopt a method described by Heuberger et al. [**11**] and find a lower bound for $\deg Y$. An upper bound for $\deg Y$ is found by Mason's fundamental lemma [**15**, Lemma 2, Chapter 1]. Comparing upper and lower bounds yields Theorem 1. Note that for the rest of the paper we will assume that k is an algebraic closed field of characteristic 0.

2. Preliminaries. First, we state Mason's fundamental lemma [**15**, Lemma 2, Chapter 1], which is a special case of the ABC-theorem for function fields (see, e.g., [**20**, Theorem 7.17]).

Lemma 1. *Let $K/k(T)$ be a function field of genus g, let us denote by M_K the set of all valuations in K, let $H_K(\alpha) := -\sum_{\omega\in M_K}\min(0,\omega(\alpha))$ denote the height of $\alpha \in K$ and let $\gamma_1,\gamma_2,\gamma_3 \in K$ with $\gamma_1+\gamma_2+\gamma_3 = 0$. Let $\mathcal{V}$ be a finite set of valuations such that for all $\omega \notin \mathcal{V}$ we have $\omega(\gamma_1) = \omega(\gamma_2) = \omega(\gamma_3)$; then*

$$H_K(\gamma_1/\gamma_2) \le \max(0, 2g-2+|\mathcal{V}|).$$

Let $K/k(T)$ be an extension of function fields; then we are interested in the integral closure of $k[T]$ in K, which is denoted by $\mathfrak{O}_K$. Obviously $\mathfrak{O}_K$ is a Dedekind ring and all primes of $k[T]$ are tamely ramified. Assume K is Galois over $k(T)$; then we know

$$\mathrm{N}_{K/k(T)}(\mathcal{D}_{\mathfrak{O}_K/k[T]}) = \prod_{a\in k}(T-a)^{(e_a-1)g_a} = \delta_{\mathfrak{O}_K/k[T]}, \tag{3}$$

where $\mathcal{D}_{B/A}$ and $\delta_{B/A}$ denote the different and the discriminant of B over A respectively. Moreover, $(T-a)\mathfrak{O}_L = (\mathfrak{p}_1\cdots\mathfrak{p}_{g_a})^{e_a}$, i.e., e_a is the ramification index of $(T-a)$ in $\mathfrak{O}_L$. The equation above holds since the residue class degree is 1 in this case (see also [**25**]). Besides the valuations which are obtained from the primes $(T-a)$ with $a \in k$, infinite valuations also exist which are obtained by the unique maximal ideal of the discrete valuation ring

$$\mathfrak{O}_\infty := \{f(T)/g(T) : f, g \in \mathbf{C}[T], \deg(f) \le \deg(g)\} \subset k(T).$$

The following result is useful to determine ramifications and valuations:

Proposition 1 (Puiseux). *Let the function field $K/k(T)$ be defined by the polynomial*

$$P(X,T) = X^d + P_{d-1}(T)X^{d-1} + \cdots + P_0(T)$$

with coefficients $P_0, \ldots, P_{d-1} \in k(T)$. Then for each $a \in k$ there exists a formal Puiseux series

$$y_{i,j} = \sum_{h=m_i}^{\infty} c_{h,i}\zeta_i^{hj}(T-a)^{h/e_{a,i}}, \quad (1 \leq j \leq e_{a,i}, 1 \leq i \leq r_a),$$

where $c_{h,i} \in k$ and $\zeta_i \in k$ is an $e_{a,i}$th root of unity such that

$$P(X,T) = \prod_{i=1}^{r_a}\prod_{j=1}^{e_{a,i}}(X - y_{i,j}).$$

Moreover, let $\mathfrak{P}_1, \ldots, \mathfrak{P}_{r_a}$ be the primes of K lying above the prime $(T-a)$. Then $e_{a,i} = e(\mathfrak{P}_i|(T-a))$ for $i = 1, \ldots, r_a$ for some appropriate order of the indices.

Note that a similar statement holds also for infinite valuations. Furthermore, the m_i are the valuations of α with respect to the primes above $(T-a)$, where α is a root of $P(X,T)$.

An essential tool in our proof is Mason's fundamental lemma (see Lemma 1). But, for an application of this lemma, we need a tool to compute the genus of function fields. The Riemann-Hurwitz formula (see, e.g., [**20**, Theorem 7.16]) yields such a tool.

Proposition 2 (Riemann-Hurwitz). *Let L/K be a geometric extension of function fields with constant field k. Let g_K and g_L be the genera of K and L, respectively. Then*

$$2g_L - 2 = [L:K](2g_K - 2) + \sum_{w \in M_L}(e_w - 1), \tag{4}$$

where M_L is the set of valuations of L and e_w denotes the ramification index of $w \in M_L$ in the extension L/K.

A geometric extension L/K is a finite algebraic extension of function fields such that $L \cap \bar{k} = k$, where k is the constant field of K and L. Note, if k is algebraic closed, then every finite algebraic extension is geometric.

3. Unit structure. Let $F(X) = X(X-p_1)\cdots(X-p_{d-1})+1 \in k[T,X]$, α a root of $F(X)$, $K = k(T,\alpha)$ and $L \supset K$ the splitting field of F over $k(T)$. Moreover, let us assume K and L are imbedded by a fixed morphism into $\overline{K} := \cup_{r=1}^{\infty} k((T^{-1/r}))$, which is the algebraic closure of $k(T)$. We fix this embedding for the rest of the paper. Furthermore, let ν be the valuation such that $\nu(f) = -\deg(f)$ for any $f \in \overline{K}^*$ and $\nu(0) = -\infty$. In the sequel we will use both the deg- and the ν-notation. In order to distinguish between polynomials $\in k[T]$ and algebraic functions $\in \overline{K}$ we will use the deg-notation only for polynomials.

Let $\tilde{d}_i = \nu(\alpha_i)$, for $0 \le i \le d-1$, with $\tilde{d}_0 \ge \cdots \ge \tilde{d}_{d-1}$, where $\alpha_0, \ldots, \alpha_{d-1} \in L \subset \overline{K}$ are the conjugates of α. Then we have:

Lemma 2. *Let $d_i = \deg(p_i)$ and $d_0 = -\sum_{i=1}^{d-1} d_i$. Then we have $\nu(\alpha_i) = \tilde{d}_i = -d_i$ for $0 \le i \le d-1$.*

Proof. Suppose $x = \nu(\alpha_i)$ and $-d_k > x > -d_{k+1}$ for some $k = 0, 1, \ldots, d-2$ or $-d_{d-1} > -x$ and $k = d-1$. Then we have

$$0 = \nu(1) = \nu(\alpha_i(\alpha_i - p_1)\cdots(\alpha_i - p_{d-1})) = (k+1)x - \sum_{j=k+1}^{d-1} d_j < 0,$$

a contradiction. Similarly, we get a contradiction if $x > -d_0$. Hence, each $\tilde{d}_j$ is equal to some $-d_i$. Therefore, we have to prove that $\tilde{d}_j \neq \tilde{d}_i$ if $i \neq j$. Obviously, $\tilde{d}_0 = -d_0$, respectively $\tilde{d}_{d-1} = -d_{d-1}$, since otherwise $0 = \nu(1) = \nu(\alpha_0 \cdots \alpha_{d-1}) < 0$, respectively

$$\begin{aligned} -d_{d-1} &= -\deg(p_1 + \cdots + p_{d-1}) = \nu(\alpha_0 + \cdots + \alpha_{d-1}) \\ &\ge \nu(\alpha_{d-1}) = \tilde{d}_{d-1} > -d_{d-1}. \end{aligned}$$

Let j be the largest index such that $\tilde{d}_j = \tilde{d}_{j-1} = \cdots = \tilde{d}_{j-k+1}$ for some $k > 1$. Then we obtain the following equations:

$$\begin{aligned}
-d_{d-1} &= -\deg\left(\sum_j p_j\right) = \nu\left(\sum_j \alpha_j\right) = \nu(\alpha_{d-1}) = \tilde{d}_{d-1},\\
-d_{d-1} - d_{d-2} &= -\deg\left(\sum_{j_1<j_2} p_{j_1}p_{j_2}\right)\\
&= \nu\left(\sum_{j_1<j_2} \alpha_{j_1}\alpha_{j_2}\right) = \nu(\alpha_{d-1}\alpha_{d-2})\\
&= \tilde{d}_{d-1} + \tilde{d}_{d-2}\\
&\vdots\\
-d_{d-1} - \cdots - d_{j+1} &= -\deg\left(\sum_{j_1<\cdots<j_{d-j-1}} p_{j_1}\cdots p_{j_{d-j-1}}\right),\\
&= \nu\left(\sum_{j_1<\cdots<j_{d-j-1}} \alpha_{j_1}\cdots \alpha_{j_{d-j-1}}\right)\\
&= \nu(\alpha_{d-1}\cdots\alpha_{j+1})\\
&= \tilde{d}_{d-1} + \cdots + \tilde{d}_{j+1}.
\end{aligned}$$

This yields $-d_i = \tilde{d}_i$ for all $i > j$. But we have

$$\begin{aligned}
-d_{d-1} - \cdots - d_j &= -\deg\left(\sum_{j_1<\cdots<j_{d-j}} p_{j_1}\cdots p_{j_{d-j}}\right)\\
&= \nu\left(\sum_{j_1<\cdots<j_{d-j}} \alpha_{j_1}\cdots \alpha_{j_{d-j}}\right)\\
&\geq \nu(\alpha_{d-1}\cdots\alpha_j) = -d_{d-1} - \cdots - d_{j+1} + \tilde{d}_j\\
&\geq -d_{d-1} - \cdots - d_{j+1} - d_j
\end{aligned}$$

and therefore $\tilde{d}_j = \tilde{d}_{j-1} = \cdots = \tilde{d}_{j-k+1} = -d_j$. This implies

$$\begin{aligned}
-d_{d-1} - \cdots - d_{j-k+1} &= -\deg\left(\sum_{j_1<\cdots<j_{d-j+k-1}} p_{j_1}\cdots p_{j_{d-j+k-1}}\right)\\
&= \nu\left(\sum_{j_1<\cdots<j_{d-j+k-1}} \alpha_{j_1}\cdots \alpha_{j_{d-j+k-1}}\right)\\
&= \nu(\alpha_{d-1}\cdots\alpha_{j-k+1}) = -d_{d-1} - \cdots - d_{j+1} - kd_j
\end{aligned}$$

which is a contradiction to the assumption $d_0 < d_1 < \cdots < d_{d-1}$ unless $k = 1$. $\square$

The lemma above implies $F(X)$ is irreducible. Otherwise, $F(X) = P_1(X)P_2(X)$ such that $P_1(\alpha_0) \neq 0$ and $\{\alpha_i : i \in N \subset \{1, 2, \ldots, d-1\}\}$ is a complete set of conjugate units such that $\nu(\alpha_i) < 0$ for each $i \in N$, which is a contradiction. Furthermore, we conclude that K is unramified at infinity since all the infinite valuations of α are distinct by the lemma above. Similarly we deduce that L is not ramified at infinity, since for each automorphism $\sigma \in \mathrm{Gal}\,(L/k(T))$ not the identity we can find two indices j and k such that $\nu(\alpha_j - \alpha_k) \neq \nu(\sigma(\alpha_j - \alpha_k))$. Therefore, we have:

Lemma 3. *The polynomial $F(X)$ is irreducible and the fields K and L are unramified over $k(T)$ at infinity.*

Let us denote by ∞_i the valuation that is induced by the imbedding $\sigma_i : K \to \overline{K}$ such that $\alpha \mapsto \alpha_i$, and let $|\alpha|_{\infty_i} := \nu(\sigma_i \alpha)$. Then we consider the free Abelian group $\mathcal{D}_\infty$ generated by $\infty_0, \ldots, \infty_{d-1}$. Because of the previous lemma we know that all the valuations ∞_i are distinct. The group $\mathcal{D}_\infty$ is called the group of polar divisors. We also define the polar height $H_K^{(\infty)}(a_1\infty_1 + \cdots + a_{d-1}\infty_{d-1}) = -\sum(\min\{0, a_i\})$. Similarly, we can define for every $a \in k$ the group of local divisors $\mathcal{D}_a$, which is freely generated by the valuations above the finite valuation induced by the prime $(T - a)$. Let $\omega_1, \ldots, \omega_{g_a}$ be these valuations; then $H_K^{(a)}(a_1\omega_1 + \cdots + a_{g_a}\omega_{g_a}) = -\sum(\min\{0, a_i\})$ is called the local height. For every $\alpha \in K$, we can define the principal divisor $(\alpha) = \sum_{\omega \in M_K} \omega(\alpha)\omega$ of α and similarly we can define the (principal) polar and local divisors of α. Note that M_K is the set of all valuations of K.

Proposition 3. *Let $\varepsilon \in k[T, \alpha]^* \subset \mathfrak{O}_K^*$ with $\varepsilon \notin k$. Then we have*

$$H_K(\varepsilon) \geq -d_0 = \sum_{i=1}^{d-1} d_i.$$

Proof. Let us write

$$\varepsilon = h_0 + h_1\alpha + \cdots + h_{d-1}\alpha^{d-1},$$

with $h_0, \ldots, h_{d-1} \in k[T]$. Moreover, we define

$$\begin{aligned} m_0 &= \min\{-\deg(h_0), -\deg(h_1) - d_0, \ldots, -\deg(h_{d-1}) - (d-1)d_0\},\\ m_1 &= \min\{-\deg(h_0), -\deg(h_1) - d_1, \ldots, -\deg(h_{d-1}) - (d-1)d_1\},\\ &\vdots\\ m_{d-1} &= \min\{-\deg(h_0), -\deg(h_1) - d_{d-1}, \ldots, -\deg(h_{d-1})\\ &\qquad - (d-1)d_{d-1}\}. \end{aligned}$$

We see that the polar divisor of ε has ∞_j-coefficient m_j if the corresponding minimum occurs only once. Note that the minimal index j with $m_i = -\deg(h_j) - jd_i$ is decreasing with i.

Now, let us assume $m_l = -\deg h_j - jd_l$ is the singular minimum with l maximal. Then we have

$$\begin{gathered} -\deg(h_j) - jd_l \le -\deg(h_{j_1}) - j_1 d_l\\ -\deg(h_{j_1}) - j_1 d_{l+1} \le -\deg(h_{j_2}) - j_2 d_{l+1}\\ \vdots\\ -\deg(h_{j_{d-l-1}}) - j_{d-l-1}d_{d-1} = -\deg(h_{j_{d-l}}) - j_{d-l}d_{d-1} \end{gathered}$$

with $j = j_0 \le j_1 < \cdots < j_{d-l-1} < j_{d-l}$ and each j_i minimal with $1 \le i \le d-l-1$. Moreover, j_{d-l} is maximal such that $-\deg(h_{j_{d-l}}) - j_{d-l}d_{d-1} = m_{d-1}$. Since $j_{d-l} \le d-1$, we conclude $j \le l$ and by the inequalities above, we obtain

$$-\deg(h_j) \le -\sum_{i=l}^{d-1}(j_{d-i} - j_{d-i-1})d_i - \deg(h_{j_{d-l}}) \le -\sum_{i=d-1-(j_{d-l}-j)}^{d-1} d_i.$$

Therefore, $m_l = -\deg(h_j) - jd_l \le -\sum_{i=d-1-j_{d-l}}^{d-1} d_i$.

Furthermore, there are at least $d - j_{d-l}$ singular minima m_i and at least $d - j_{d-l} - 1$ such minima with $i > 0$. Indeed, if $m_i = \deg(h_{j_i}) + j_i d_i$ and $m_{i+1} = \deg(h_{j_{i+1}}) + j_{i+1}d_{i+1}$ with j_i and j_{i+1} minimal and m_i not a singular minimum, then $j_i < j_{i+1}$. Therefore, we have singular minima m_{i_k} other than m_l with $1 \leq i_1 < \cdots < i_{d-2-j_{d-l}}$. We claim $m_{i_k} \leq -d_{i_k}$, with $k = 1, \ldots, d-2-j_{d-l}$. Each minima m_i is of the form $-\deg(h_j) - jd_i$. If $j \neq 0$, we are done. Otherwise, $-\deg(h_0) < -\deg(h_j) - jd_i \leq -d_i$ for some j with $\deg(h_j) \neq -\infty$. Note that the case $h_j = 0$ for each $j > 0$ leads to $\varepsilon = h_0 \in k[T]$ and this leads to $\varepsilon \in k^*$. Altogether we deduce $m_{i_k} \leq -d_k$ and

$$H_K(\varepsilon) \geq -m_l - \sum_{k=1}^{d-2-j_{d-l}} m_{i_k} \geq \sum_{k=d-1-j_{d-l}}^{d-1} d_k - \sum_{k=1}^{d-2-j_{d-l}} m_{i_k} = \sum_{k=1}^{d-1} d_k. \square$$

By an analog of Dirichlet's unit theorem for function fields, we know that there are at most $d-1$ multiplicatively independent units that generate the unit group of $k[T,\alpha]$. This fundamental system of units spans a lattice, i.e., consider the map $\log : k[T,\alpha]^* \to \mathbf{Z}^{d-1}$ with $\log(\varepsilon) = (|\varepsilon|_{\infty_1}, \ldots, |\varepsilon|_{\infty_{d-1}})$, then the image of $k[T,\alpha]^*$ is a lattice $\Lambda \subset \mathbf{Z}^{d-1}$. Moreover, $\log(\varepsilon)$ is the vector of components of the polar divisor of ε except the component of ∞_0. It is obvious that a set of units is multiplicatively independent if and only if the corresponding vectors in the lattice are independent. Let us consider the set $\mathcal{E} = \{\varepsilon_1, \ldots, \varepsilon_{d-1}\}$ of $d-1$ independent units. Then the absolute value of the determinant of the matrix whose rows are $\log(\varepsilon_1), \ldots, \log(\varepsilon_{d-1})$ is usually called the regulator of $\mathcal{E}$. Note that the regulator is the same as the lattice constant of the lattice spanned by the vectors $\log(\varepsilon_1), \ldots, \log(\varepsilon_{d-1})$. Let us fix the following set of units $\mathcal{E} = \{\alpha, \alpha - p_1, \ldots, \alpha - p_{d-2}\}$. Then we have

Lemma 4. *The set $\mathcal{E}$ is a system of multiplicatively independent units and the corresponding regulator $R_{\mathcal{E}}$ is*

$$(2d_1 + d_2 + \cdots d_{d-1})(3d_2 + d_2 + \cdots + d_{d-1}) \cdots ((d-1)d_{d-2} + d_{d-1})d_{d-1}.$$

Proof. Note, if $l \neq k$, then

$$|\alpha - p_l|_{\infty_k} = \begin{cases} -d_l & \text{if } k < l, \\ -d_k & \text{if } k > l. \end{cases}$$

Since the sum of all valuations of an element is zero and $\alpha - p_l$ has nonzero valuations only at infinity we deduce $|\alpha - p_k|_{\infty_k} = ld_l + \sum_{k=l+1}^{d-1} d_k$. Therefore,

$$\begin{pmatrix} -d_1 & -d_2 & -d_3 & \dots & -d_{d-1} \\ d_1 + \sum_{k=2}^{d-1} d_k & -d_2 & -d_3 & \dots & -d_{d-1} \\ -d_2 & 2d_2 + \sum_{k=3}^{d-1} d_k & -d_3 & \dots & -d_{d-1} \\ -d_3 & -d_3 & 3d_3 + \sum_{k=4}^{d-1} d_k & \dots & -d_{d-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -d_{d-2} & -d_{d-2} & -d_{d-2} & \dots & -d_{d-1} \end{pmatrix}$$

is the matrix which consists of the vectors $\log(\alpha), \log(\alpha - p_1), \dots,$ $\log(\alpha - p_{d-2})$. After a 180° rotation and Gaussian elimination, we obtain the upper triangular matrix

$$\begin{pmatrix} d_{d-1} & & \dots & \\ 0 & (d-1)d_{d-2} + d_{d-1} & \dots & \ddots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 2d_1 + \sum_{k=2}^{d-1} d_k \end{pmatrix}.$$

This yields the lemma. □

Lemma 5. *Let* $\Sigma = \sum_{k=1}^{d-1} d_k = -d_0$. *Then we have*

$$R_{\mathcal{E}} \leq \left(\left(2 - \frac{2\sqrt{d-2}}{d-1} \right) \Sigma \right)^{d-1} < (2\Sigma)^{d-1}.$$

Proof. By Lemma 4, we have

$$\begin{aligned} R_{\mathcal{E}} &= (2d_1 + d_2 + \cdots d_{d-1})(3d_2 + d_2 + \cdots + d_{d-1}) \cdots \\ &\quad \times ((d-1)d_{d-2} + d_{d-1})d_{d-1} \\ &= (\Sigma + d_1)(\Sigma + 2d_2 - d_1) \cdots (\Sigma + (d-2)d_{d-2} - d_1 - \cdots - d_{d-3}) \\ &\quad \times (\Sigma - d_1 - \cdots - d_{d-2}) \\ &\leq \left(\frac{(\Sigma + d_1) + \cdots + (\Sigma - d_1 - \cdots - d_{d-2})}{d-1} \right)^{d-1} \\ &= \left(\frac{d_1 + 3d_2 + 5d_3 + \cdots + (2d-5)d_{d-2} + (d-1)d_{d-1}}{d-1} \right)^{d-1}. \end{aligned}$$

The inequality is due to the arithmetic-geometric mean. We have to maximize

$$d_1 + 3d_2 + \cdots + (2d-5)d_{d-2} + (d-1)d_{d-1} \tag{5}$$

under the condition

$$0 \leq d_1 \leq d_2 \leq \cdots \leq d_{d-1} \leq \Sigma - d_1 - d_2 - \cdots - d_{d-2}. \tag{6}$$

Since this describes a linear program, the maximum lies in a corner of the polytop defined by (6). Let us consider the corner $d_1 = d_2 = \cdots = d_k = 0$ and $d_{k+1} = \cdots = d_{d-1} = \Sigma/(d-1-k)$ for some $k = 0, \ldots, d-1$. For this corner (assume $k < d-1$) function (5) turns into

$$\begin{aligned} &\frac{\Sigma}{d-1-k}\Big(-(d-2) + \sum_{j=k+1}^{d-1}(2j-1)\Big) \\ &\qquad = \frac{\Sigma((d-1)^2 - k^2)}{d-k-1} - \frac{\Sigma(d-2)}{d-1-k} \\ &\qquad = \Sigma\Big(d-1+k-\frac{d-2}{d-1-k}\Big) \leq \Sigma(2(d-1) - 2\sqrt{d-2}). \end{aligned}$$

For $k = d-1$ the quantity (5) is zero and yields no maximum. □

Let U be the group of units of $k[T,\alpha]$ and $\widetilde{U}$ the group of units generated by $\mathcal{E}$, then we know that $I := [U : \widetilde{U}] = R_{\mathcal{E}}/R$, where R is the regulator of U. Therefore we have to find a lower bound for R in order to find an upper bound for I. Before we determine such an upper bound we prove the following lemma.

Lemma 6. *There is a symmetric convex region $\mathcal{S}$ with $\lambda(\mathcal{S}) \geq \Sigma^{d-1}2^{d-1}/(d-1)!$ around 0 such that $\log(U) \cap \mathcal{S} = \{0\}$ and λ denotes the $d-1$-dimensional Lebesgue measure.*

Proof. By Proposition 3 for any unit $\varepsilon \in U \setminus k$ we have $H_K(\varepsilon) \geq \Sigma$. Let $\log(\varepsilon) = (e_1, \ldots, e_{d-1})$, then

$$\max\Big\{\sum \max\{0, e_j\}, -\sum \min\{0, e_j\}\Big\} \geq \Sigma.$$

We consider the set $\mathcal{S} = \{x = (x_0, \dots, x_{d-1}) \in \mathbf{R}^{d-1} : \sum |x_i| \leq \Sigma\}$. Then $\mathcal{S}$ consists of 2^{d-1} simplices of volume $\Sigma^{d-1}/(d-1)!$ and fulfills $\log(U) \cap \mathcal{S} = \{0\}$. □

Proposition 4. *We have*

$$I \leq (d-1)! 4^{d-1}.$$

Proof. We apply Minkowski's convex body theorem (see [**1**, Theorem II]) to the region $\mathcal{S}$ of Lemma 6. This yields $R \geq (\Sigma/2)^{d-1}/(d-1)!$. On the other hand we have by Lemma 5 the inequality $R_{\mathcal{E}} < (2\Sigma)^{d-1}$; hence,

$$I = \frac{R_{\mathcal{E}}}{R} \leq (d-1)! 4^{d-1}. \qquad \square$$

4. Application of Mason's fundamental lemma. Let (X, Y) be a solution to (2). In the classical approach to Thue equations, the terms $X - \alpha_i Y$ and $(\alpha_k - \alpha_l)(X - \alpha_i Y)$ are denoted by β_i and $\gamma_{k,l,i}$, respectively. Then Siegel's identity can be written as

$$\gamma_{k,l,i} + \gamma_{l,i,k} + \gamma_{i,k,l} = 0.$$

We apply Mason's fundamental lemma (Lemma 1) in order to find an upper bound for $H_L(\gamma_{k,l,i}/\gamma_{l,i,k})$. Since $L/k(T)$ is Galois and k is of characteristic zero we have by equation (3)

$$\delta_{\mathfrak{O}_L/k[T]} = \prod_{a \in k} (T-a)^{(e_a - 1) g_a}.$$

Let $\delta_{k[T,\alpha]/k[T]}$ be the discriminant of $k[T, \alpha]$ over $k[T]$. Since

$$\delta_{k[T,\alpha]/k[T]} = \prod_{i<j} (\alpha_i - \alpha_j)^2 \in k[T],$$

we have $\deg(\delta_{k[T,\alpha]/k[T]}) = 2 \sum_{i<j} \max\{d_i, d_j\} = 2 \sum_{j=1}^{d-1} j d_j$. Moreover, algebraic number theory shows that $\delta_{k[T,\alpha]/k[T]}$ and $\delta_{\mathfrak{O}_L/k[T]}$ differ

only by a square factor, say R^2. Let $\deg(R) = r$. By Hurwitz's formula we obtain

$$\begin{aligned} 2g_L - 2 &= \deg L(2g_{k(T)} - 2) + \sum_a (e_a - 1) g_a \\ &= -2\deg L + \sum_a (e_a - 1) g_a. \end{aligned}$$

By Mason's fundamental lemma we get

$$\begin{aligned} H_L\left(\frac{\gamma_{k,l,i}}{\gamma_{l,i,k}}\right) &\leq -2\deg L + \sum_a (e_a - 1) g_a + \sharp\mathcal{V} + \sharp\mathcal{V}_\infty \\ &= -\deg L + \sum_a (e_a - 1) g_a + \sharp\{\mathfrak{p} \triangleleft \mathfrak{O}_L : \mathfrak{p}|\delta\} \\ &\leq -\deg L + \sum_a (e_a - 1) g_a + \sum_a g_a + r\deg L \\ &= \deg L\left(r - 1 + \sum_a 1\right). \end{aligned}$$

All the sums are taken over all a such that $(T-a)|\delta_{\mathfrak{O}_L/k[T]}$. Moreover, $\mathcal{V}$ denotes the set of all valuations ω of L such that $\omega(\alpha_i - \alpha_j) \neq 0$ for some i, j and $\mathcal{V}_\infty$ denotes the set of all infinite valuations of L. This enables us to compute a bound for $H_K(\beta)$:

Lemma 7.

$$H_K(\beta) \leq (d-1)\left(r - 1 + \sum_a 1\right) + \frac{d-1}{d} d_{d-1} + \frac{4(d-1)}{27} d_{d-2}.$$

Proof. Let $b_0 \geq b_1 \geq \cdots \geq b_{d-1}$ be the infinite valuations of $\beta = X - \alpha Y$ over K. Since β is a unit, we have $\sum_{j=0}^{d-1} b_j = 0$. Moreover, we have $\nu(X - \alpha_i Y) \leq 0$ unless $\deg X \neq \deg Y + d_i$; hence, $H_K(\beta) = b_0$. Indeed there is at most one positive valuation of β, which is by definition b_0. Furthermore, we have

$$\begin{aligned} H_L\left(\frac{\beta_i(\alpha_k - \alpha_l)}{\beta_k(\alpha_l - \alpha_i)}\right) &\geq H_L^{(\infty)}\left(\frac{\beta_i(\alpha_k - \alpha_l)}{\beta_k(\alpha_l - \alpha_i)}\right) \\ &= H_L\left(\frac{\beta_i}{\beta_k}\right) - H_L^{(\infty)}\left(\frac{\alpha_k - \alpha_l}{\alpha_l - \alpha_i}\right). \end{aligned}$$

Let H be a system of representatives of $S_d/\mathrm{Gal}\,(L/k(T))$. Note that the symmetric group S_d acts on L by permuting the conjugates of α. We obtain

(7)
$$\begin{aligned}\max_{i,j,k} H_L\left(\frac{\gamma_{k,l,i}}{\gamma_{l,i,k}}\right) &\geq \frac{|G|}{|S_d|}\left(\sum_{\rho\in H} H_L\left(\rho\frac{\beta_1}{\beta_2}\right) - \sum_{\rho\in H} H_L\left(\rho\frac{\alpha_2-\alpha_3}{\alpha_3-\alpha_1}\right)\right)\\ &= \frac{|G|}{|S_d|}\left(\sum_{\sigma\in S_d}\max\left(0,\nu\left(\sigma\frac{\beta_1}{\beta_2}\right)\right)\right.\\ &\qquad\left.-\sum_{\sigma\in S_d}\max\left(0,\nu\left(\sigma\frac{\alpha_2-\alpha_3}{\alpha_3-\alpha_1}\right)\right)\right).\end{aligned}$$

Let us consider the first sum of (7):

$$\begin{aligned}&\frac{|G|}{|S_d|}\sum_{\sigma\in S_d}\max\left(0,\nu\left(\sigma\frac{\beta_1}{\beta_2}\right)\right)\\ &\qquad= \frac{\deg L}{d!}(d-2)!\sum_{i<j}(b_i-b_j)\\ &\qquad= \frac{\deg L}{d(d-1)}[(d-1)b_0+(d-3)b_1+\cdots+(-d+1)b_{d-1}]\\ &\qquad= \frac{\deg L}{d(d-1)}[-(2d-2)b_{d-1}-\cdots-2b_1]\\ &\qquad\geq \frac{\deg L}{d(d-1)}db_0 = \frac{\deg L}{d-1}H_K(\beta).\end{aligned}$$

The last inequality is true, since $-(2d-2)b_{d-1}-\cdots-2b_1$ with $0\geq b_1\geq\cdots\geq b_{d-1}\geq -b_0=\sum_{j>0}b_j$ takes its minimum for $b_1=\cdots=b_{d-1}=-b_0/(d-1)$.

Now we investigate the second sum of (7). Let us remark that

$$\max\left(0,\frac{\alpha_k-\alpha_l}{\alpha_l-\alpha_i}\right)=\begin{cases}0 & \text{if } l>k \text{ or } i>k,\\ d_k-d_i & \text{if } k>i>l,\\ d_k-d_l & \text{if } k>l>i.\end{cases}$$

This yields

$$\begin{aligned}
&\frac{|G|}{|S_d|}\sum_{\sigma\in S_d}\max\left(0,\nu\left(\sigma\frac{\alpha_2-\alpha_3}{\alpha_3-\alpha_1}\right)\right)\\
&\quad=\frac{\deg L}{d!}2(d-3)!\sum_{i=2}^{d-1}\sum_{j=1}^{i-1}j(d_i-d_j)\\
&\quad=\frac{\deg L}{d(d-1)(d-2)}\left(\sum_{i=1}^{d-1}i(i-1)d_i-2\sum_{i=2}^{d-1}\sum_{j=1}^{i-1}jd_j\right)\\
&\quad=\frac{\deg L}{d(d-1)(d-2)}\sum_{i=1}^{d-1}(i(i-1)-2i(d-i-1))d_i\\
&\quad=\frac{\deg L}{d(d-1)(d-2)}\sum_{i=1}^{d-1}(d-i)(d-3i+1)d_{d-i}\\
&\quad\le\frac{\deg L}{d}d_{d-1}+\frac{\deg L}{d(d-1)(d-2)}d_{d-2}\sum_{i=2}^{[(d+1)/3]}(d-i)(d-3i+1)\\
&\quad\le\frac{\deg L}{d}d_{d-1}+\frac{4\deg L}{27}d_{d-2}.
\end{aligned}$$

Combining these estimates with (7), we get the desired result. □

Before we establish an upper bound for $\deg Y$, let us note that we may exclude the case $X=0$. In this case equation (2) turns into $Y^d=\xi$, which yields a trivial solution. Therefore, we may assume $\deg X\ge 0$.

Let us note $\nu(X-\alpha_jY)<0$ unless $\deg X=\deg Y+d_j$. If β is not a constant there exists an index j such that $\nu(X-\alpha_jY)>0$. Let us fix this index j, and let $b_0\ge\cdots\ge b_{d-1}$ be the infinite valuations of β. Then we have $b_1=\cdots=b_j=-\deg X$, $b_k=-\deg Y-d_k$ for $k>j$ and

$$b_0=j\deg X+(d-1-j)\deg Y+\sum_{j<k\le d-1}d_k.$$

Hence,

$$H_K(\beta)=b_0=(d-1)\deg Y+jd_j+\sum_{j<k\le d-1}d_k\ge(d-1)\deg Y+\sum_{k=1}^{d-1}d_k.$$

Together with Lemma 7 we obtain

$$\deg Y \leq r - 1 + \sum_a 1 - \sum_{j=1}^{d-1} \frac{d_j}{d-1} + \frac{d_{d-1}}{d} + \frac{4d_{d-2}}{27}.$$

Counting the number of ramifications yields:

Proposition 5. *We have*

$$\deg Y \leq \left(\sum_{j=1}^{d-1} \left(j - \frac{1}{d-1} \right) d_j \right) - 1 + \frac{d_{d-1}}{d} + \frac{4d_{d-2}}{27}.$$

5. A lower bound for $\deg Y$**.** First, we exclude the case $Y \in k$.

Lemma 8. *If* $Y \in k$*, then* (X, Y) *is a trivial solution.*

Proof. Let us write $Y = \zeta \in k$ and assume $X \neq \zeta p_j$ for all $1 \leq j \leq d-1$ and $X \neq 0$. Then the righthand side of (2) has degree at least $\sum_{j=1}^{d-1} d_j = -d_0 > 0$ which is a contradiction. Therefore, $X = \zeta p_j$ for some j or $X = 0$. This yields $Y^d = \zeta^d = \xi$; hence, (X, Y) is a trivial solution. □

As mentioned above there is some index j such that $\nu(\beta_j) = \nu(X - \alpha_j Y) > 0$. Let us fix this index j. Then we have, for $k \neq j$,

$$\begin{aligned} \nu(\beta_k) = \nu(X - Y\alpha_k) &= \nu\left(Y(\alpha_j - \alpha_k) \left(1 + \frac{\beta_j}{Y(\alpha_k - \alpha_j)} \right) \right) \\ &= \nu(\alpha_j - \alpha_k) - \deg Y. \end{aligned}$$

Note that $\nu(\beta_j/(Y(\alpha_k - \alpha_j))) > 0$, hence $\nu(1 + (\beta_j)/(Y(\alpha_k - \alpha_j))) = 0$. On the other hand, β_k is a unit and therefore

$$\beta_i = (\alpha_k)^{B_0/I} (\alpha_k - p_1)^{B_1/I} \cdots (\alpha_k - p_{d-2})^{B_{d-2}/I}$$

where $B_0, B_1, \ldots, B_{d-2} \in \mathbf{Z}$ and $I \leq 4^{d-1}(d-1)$. This yields

(8)

$$\nu(\alpha_k) \frac{B_0}{I} + \nu(\alpha_k - p_1) \frac{B_1}{I} + \cdots + \nu(\alpha_k - p_{d-2}) \frac{B_{d-2}}{I} = \nu(\alpha_j - \alpha_k) - \deg Y$$

for $k \neq j$. Solving system (8) with $k \neq j$ for B_i we obtain

$$\frac{B_k}{I} = \frac{v_k - u_k \deg Y}{R}, \tag{9}$$

where

$$u_k = \det(\nu(\alpha_i), \ldots, \nu(\alpha_i - p_{k-1}), 1, \nu(\alpha_i - p_{k+1}) \ldots, \nu(\alpha_i - p_{d-2}))_{i \neq j},$$
$$v_k = \det(\nu(\alpha_i), \ldots , \nu(\alpha_i - p_{k-1}), \nu(\alpha_i - \alpha_j), \nu(\alpha_i - p_{k+1}),$$
$$\ldots , \nu(\alpha_i - p_{d-2}))_{i \neq j},$$
$$R = \det(-\nu(\alpha_i), -\nu(\alpha_i - p_1), \ldots , -\nu(\alpha_i - p_{d-2}))_{i \neq j}.$$

Lemma 9. *We have*

$$v_k = \begin{cases} R & k = j \neq d-1; \\ -R & k = d-1; \\ 0 & \textit{otherwise}; \end{cases}$$

with $R = (-1)^{d+j} R_{\mathcal{E}}$.

Proof. Let $k = j \neq d-1$. Then the matrices which determine v_k and R are identical. Moreover, the matrix corresponding to R can be transformed to the matrix corresponding to the regulator by summing up all lines in the last line, multiplying the last line by -1 and exchanging $d-1-k$ lines.

Now let us assume that $j \neq k$ and $k < d-1$. The jth, respectively kth, column of the matrix corresponding to v_k are $(-\nu(\alpha_j), -\nu(\alpha_j - p_1), \ldots , -\nu(\alpha_j - p_{d-1}))^T$ and $(-\nu(\alpha_j - \alpha_0), \ldots , -\nu(\alpha_j - \alpha_{d-1}))^T$, where we omit the $j+1$th entry. Since these two columns are equal we deduce $v_k = 0$.

In the case of $k = d-1$ we sum up all columns of the corresponding matrix to v_k in the kth column. We multiply this column by -1 and obtain the matrix which corresponds to R. $\square$

Lemma 10. *We have*

$$\frac{u_{d-1}}{R} = \begin{cases} \frac{d_{d-2} - d_{d-1}}{d_{d-1}((d-1)d_{d-2} + d_{d-1})} & \textit{if } j < d-2, \\ \frac{d_{d-2} + (d-1)d_{d-1}}{d_{d-1}((d-1)d_{d-2} + d_{d-1})} & \textit{if } j = d-2, \\ -\frac{1}{d_{d-1}} & \textit{if } j = d-1, \end{cases}$$

and

$$\frac{u_{d-2}}{R} = \begin{cases} \frac{U_d}{d_{d-1}((d-2)d_{d-3}+d_{d-2}+d_{d-1})((d-1)d_{d-2}+d_{d-1})} & \textit{if } j < d-3, \\ \frac{U'_d}{d_{d-1}((d-2)d_{d-3}+d_{d-2}+d_{d-1})((d-1)d_{d-2}+d_{d-1})} & \textit{if } j = d-3, \\ \frac{d_{d-2}-d_{d-1}}{d_{d-1}((d-1)d_{d-2}+d_{d-1})} & \textit{if } j = d-2, \\ -\frac{1}{d_{d-1}} & \textit{if } j = d-1, \end{cases}$$

where

$$U_d = (d-2)d_{d-3}d_{d-2} + d_{d-2}^2 + 2d_{d-3}d_{d-1} - dd_{d-2}d_{d-1} - d_{d-1}^2$$

and

$$\begin{aligned} U'_d = &(d_{d-2} - d_{d-3})(d_{d-2} - d_{d-1}) \\ &+ ((d-1)d_{d-2} + d_{d-1})(d_{d-3} + (d-1)d_{d-1}). \end{aligned}$$

Moreover, $u_k = -R/d_{d-1}$ *if* $j = d-1$*.*

Proof. We consider only the cases $k = d-1$ and $j < d-2$ since the other cases run analogously. Let us write $\Sigma_0 = \Sigma = d_1 + \cdots + d_{d-1}$, $\Sigma_i = id_i + d_{i+1} + \cdots + d_{d-1}$ and $\Sigma'_i = (i+1)d_i + d_{i+1} + \cdots + d_{d-1}$ for $i = 1, \ldots d-1$. We note that $R = (-1)^{d+j}\Sigma'_1 \cdots \Sigma'_{d-2} d_{d-1}$. In order to prove the lemma we have to compute the determinant of

$$\begin{pmatrix} -\Sigma_0 & d_1 & \cdots & d_{j-1} & d_j & d_{j+1} & \cdots & d_{d-3} & 1 \\ d_1 & -\Sigma_1 & \cdots & d_{j-1} & d_j & d_{j+1} & \cdots & d_{d-3} & 1 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ d_{j-1} & d_{j-1} & \cdots & -\Sigma_{j-1} & d_j & d_{j+1} & \cdots & d_{d-3} & 1 \\ d_{j+1} & d_{j+1} & \cdots & d_{j+1} & d_{j+1} & -\Sigma_{j+1} & \cdots & d_{d-3} & 1 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ d_{d-3} & d_{d-3} & \cdots & d_{d-3} & d_{d-3} & d_{d-3} & \cdots & -\Sigma_{d-3} & 1 \\ d_{d-2} & d_{d-2} & \cdots & d_{d-2} & d_{d-2} & d_{d-2} & \cdots & d_{d-2} & 1 \\ d_{d-1} & d_{d-1} & \cdots & d_{d-1} & d_{d-1} & d_{d-1} & \cdots & d_{d-1} & 1 \end{pmatrix},$$

which is $(-1)^{d-1}u_{d-1}$. This is done by Gaussian elimination. Let us assume $j \neq 0$. First, we subtract from every row (except the first row) the first row and obtain the matrix

$$\begin{pmatrix} -\Sigma_0 & d_1 & \cdots & d_{j-1} & d_j & d_{j+1} & \cdots & 1 \\ \Sigma + d_1 & -\Sigma'_1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \Sigma + d_{j-1} & d_{j-1} - d_1 & \cdots & -\Sigma'_{j-1} & 0 & 0 & \cdots & 0 \\ \Sigma + d_{j+1} & d_{j+1} - d_1 & \cdots & d_{j+1} - d_{j-1} & d_{j+1} - d_j & -\Sigma'_{j+1} & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \Sigma + d_{d-1} & d_{d-1} - d_1 & \cdots & d_{d-1} - d_{j-1} & d_{d-1} - d_j & d_{d-1} - d_{j+1} & \cdots & 0 \end{pmatrix}.$$

Next, we sum up all columns and write this sum instead of the first column. Then the first column is of the form

$$(1 - d_{d-2} - d_{d-1}, 0, \cdots, 0, (d-1)d_{d-2} + d_{d-1}, d_{d-2} + (d-1)d_{d-1})^T.$$

This yields $u_{d-1} = (-1)^j \Sigma'_1 \cdots \Sigma'_{j-1} \det M$, where

$$M = \begin{pmatrix} d_{j+1} - d_j & -\Sigma'_{j+1} & \cdots & 0 & 0 \\ d_{j+2} - d_j & d_{j+1} - d_j & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ d_{d-3} - d_j & d_{d-3} - d_{j+1} & \cdots & -\Sigma'_{d-3} & 0 \\ d_{d-2} - d_j & d_{d-2} - d_{j+1} & \cdots & d_{d-2} - d_{d-3} & \Sigma'_{d-2} \\ d_{d-1} - d_j & d_{d-1} - d_{j+1} & \cdots & d_{d-1} - d_{d-3} & d_{d-2} + (d-1)d_{d-1} \end{pmatrix}.$$

We obtain M from the previous matrix if we place the first column behind the last column and delete the first j rows and columns. Next, we subtract from the last row the second to last row and then from the second to last row the third to last row and so on. By transposing this new matrix, we obtain

$$\begin{pmatrix} d_{j+1} - d_j & d_{j+2} - d_{j+1} & \cdots & d_{d-2} - d_{d-3} & d_{d-1} - d_{d-2} \\ -\Sigma'_{j+1} & \Sigma_{j+1} + d_{j+2} & \cdots & d_{d-2} - d_{d-3} & d_{d-1} - d_{d-2} \\ 0 & -\Sigma'_{j+2} & \cdots & d_{d-2} - d_{d-3} & d_{d-1} - d_{d-2} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & \Sigma_{d-3} + d_{d-2} & d_{d-1} - d_{d-2} \\ 0 & 0 & \cdots & \Sigma'_{d-2} & (d-2)(d_{d-1} - d_{d-2}) \end{pmatrix}.$$

We multiply the last row by -1 and add the last row to the second to last. Then we add the second to last row to the third to last row and so on. This yields

$$\begin{pmatrix} -\Sigma_{j+1}-d_j & j(d_{j+1}-d_{j+2}) & \cdots & j(d_{d-2}-d_{d-1}) \\ -\Sigma'_{j+1} & (j+1)(d_{j+1}-d_{j+2}) & \cdots & (j+1)(d_{d-2}-d_{d-1}) \\ 0 & -\Sigma'_{j+2} & \cdots & (j+2)(d_{d-2}-d_{d-1}) \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & (d-3)(d_{d-2}-d_{d-1}) \\ 0 & 0 & \cdots & (d-2)(d_{d-2}-d_{d-1}) \end{pmatrix}.$$

We multiply the second row by j and subtract $j+1$ times the first. Then we divide the first line by j. Therefore, the new matrix has the same determinant but the first two rows are of the form

$$\begin{pmatrix} -(\Sigma_{j+1}-d_j)/j & d_{j+1}-d_{j+2} & \cdots & d_{d-2}-d_{d-1} \\ -\Sigma'_j & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \end{pmatrix}.$$

Now we eliminate the other rows and obtain

$$\begin{pmatrix} -(\Sigma_{j+1}-d_j)/j & d_{j+1}-d_{j+2} & \cdots & d_{d-3}-d_{d-2} & d_{d-2}-d_{d-1} \\ -\Sigma'_j & 0 & \cdots & 0 & 0 \\ 0 & -\Sigma'_{j+1} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -\Sigma'_{d-3} & 0 \end{pmatrix},$$

which yields the lemma in this case.

In the case $j=0$, we have to compute the determinant of the matrix

$$\begin{pmatrix} d_1 & -\Sigma_1 & d_2 & \cdots & d_{d-3} & 1 \\ d_2 & d_2 & -\Sigma_2 & \cdots & d_{d-3} & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ d_{d-3} & d_{d-3} & d_{d-3} & \cdots & -\Sigma_{d-3} & 1 \\ d_{d-2} & d_{d-2} & d_{d-2} & \cdots & d_{d-2} & 1 \\ d_{d-1} & d_{d-1} & d_{d-1} & \cdots & d_{d-1} & 1 \end{pmatrix}.$$

Subtracting the first column from all other columns except the last column yields the matrix

$$\begin{pmatrix} d_1 & -\Sigma_1' & d_2-d_1 & \cdots & d_{d-3}-d_1 & 1 \\ d_2 & 0 & -\Sigma_2' & \cdots & d_{d-3}-d_2 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ d_{d-3} & 0 & 0 & \cdots & -\Sigma_{d-3}' & 1 \\ d_{d-2} & 0 & 0 & \cdots & 0 & 1 \\ d_{d-1} & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

whose determinant is $\Sigma_1' \cdots \Sigma_{d-3}'(d_{d-1} - d_{d-2})$.

The last statement of the lemma can easily be deduced. Multiply the $j+1$th column by $-d_{d-1}$ and add all other columns to this column. This yields the matrix corresponding to R. □

As indicated by Heuberger et al. (see [**11**]) we want to find a linear combination of the equations (9) such that we get a lower bound for $\deg Y$. Due to Lemmas 9 and 10 we have, for $j < d-3$,

$$\frac{B_{d-2} - B_{d-1}}{I} = \frac{d(d_{d-2} - d_{d-3})\deg Y}{\Sigma_{d-3}'\Sigma_{d-2}'},$$

for $j = d-3$

$$\begin{aligned} &\frac{1-(d-1)B_{d-1} - B_{d-2}}{I} \\ &\qquad = \frac{d((d-2)d_{d-3}d_{d-2} + d_{d-2}^2}{d_{d-1}\Sigma_{d-3}'\Sigma_{d-2}'} \\ &\qquad\quad \frac{-(d-3)d_{d-3}d_{d-1} + (d-2)d_{d-2}d_{d-1})\deg Y}{d_{d-1}\Sigma_{d-3}'\Sigma_{d-2}'} \end{aligned}$$

and, for $j = d-2$,

$$\frac{1 - B_{d-1} - (d-1)B_{d-2}}{I} = \frac{d d_{d-2}\deg Y}{\Sigma_{d-2}' d_{d-1}}.$$

For $j = d-1$, we find

$$\frac{B_0}{I} = \cdots = \frac{B_{d-1}}{I} = -\frac{d_{d-1} + \deg Y}{d_{d-1}}. \tag{10}$$

We note that the righthand sides of the equations above are > 0. Since the numerator on the left side is an integer, the righthand sides are at least $1/I$. Hence, we get lower bounds for $\deg Y$:

Proposition 6. *Let $j < d-1$ and $cd_{d-2} \leq d_{d-1}$ for some constant c. Then we have*

$$\deg Y \geq \begin{cases} \dfrac{cd_{d-1}}{2dI} & \text{if } j < d-3, \\ \dfrac{c^2 d_{d-1}}{dI(d-1+c(d-2))} & \text{if } j = d-3, \\ \dfrac{cd_{d-1}}{dI} & \text{if } j = d-2. \end{cases}$$

6. Proof of Theorem 1. We have to consider two cases: $j < d-1$ and $j = d-1$. The first case is solved by comparing the bounds for $\deg Y$ given by Propositions 5 and 6. Let us assume $cd_{d-2} \leq d_{d-1}$ for some constant c. Then we get by Proposition 5,

$$\deg Y \leq (d-1)d_{d-1} + \frac{((d-2)(d-1)+8/27)d_{d-1}}{2c},$$

and, on the other hand, we have

$$\deg Y \geq \frac{c}{dI(2d-3)},$$

by Proposition 6. This yields

$$\frac{c}{dI(2d-3)} \leq d-1+\frac{(d-2)(d-1)+8/27}{2c}.$$

But this inequality holds only for $c \leq c_d$ with

$$\begin{aligned} c_d = & \frac{9d(d-1)(2d-3)I}{18} \\ & + \frac{\sqrt{3dI(2(2d-3)(62+27(d-3)d)+27d(2d^2-5d+3)^2 I)}}{18} \\ \leq & \, 1.031 d(d-1)(2d-3)I, \end{aligned}$$

where $I \leq (d-1)!4^{d-1}$ and $d \geq 3$. Therefore, Theorem 1 is proved for $j < d-1$.

Now let us assume $j = d-1$. By (10) we know that $B_0 = \cdots = B_{d-2} =: -B$, i.e.,

$$\beta = \xi\alpha^{-B/I}(\alpha - p_1)^{-B/I}\cdots(\alpha - p_{d-2})^{-B/I} = (\alpha - p_{d-1})^{-B/I}$$

with $\xi \in k$, which yields

$$\nu(\beta) = \frac{B}{I}\nu(\alpha_{d-1} - p_{d-1}) = -\frac{B}{I}(d-1)d_{d-1} < 0,$$

hence $B > 0$.

Let us consider the case $d = 3$. By Lemma 7 and $r + \sum_a 1 = \sum_{j>0} jd_j$, we have

$$H_K(\beta) = H_K\left((\alpha - p_2)^{B/I}\right) = \frac{2Bd_2}{I} \leq 2\left(\sum_{j>0} jd_j - 1\right) < 6d_2,$$

i.e., $B/I < 3$. On the other hand. we know $I = 1$ for $d = 3$ (see [**25**]). Hence, we have $B = 1, 2$. The case $B = 1$ yields a trivial solution and in the case $B = 2$ we obtain $\beta = \xi(\alpha^2 - 2\alpha p_2 + p_2^2)$ which is not a solution.

Therefore, we may assume $d \geq 4$. We compute

$$(\alpha_k - p_{d-1})^{B/I} = p_{d-1}^{B/I} - \frac{B}{I}\alpha_k p_{d-1}^{B/I-1} + \frac{B(B-I)}{2I^2}\alpha_k^2 p_{d-1}^{B/I-2} + \cdots,$$

where the remaining terms have lower valuations provided $k < d-1$. Assume $\beta = (\alpha - p_{d-1})^{B/I}$ yields a solution. Then by Siegel's identity we have

$$\begin{aligned}
0 &= (\alpha_0 - \alpha_1)(\alpha_2 - p_{d-1})^{B/I} + (\alpha_1 - \alpha_2)(\alpha_0 - p_{d-1})^{B/I} \\
&\quad + (\alpha_2 - \alpha_0)(\alpha_1 - p_{d-1})^{B/I} \\
&= (\alpha_0 - \alpha_1)\left(p_{d-1}^{B/I} - \frac{B}{I}\alpha_2 p_{d-1}^{B/I-1} + \frac{B(B-I)}{2I^2}\alpha_2^2 p_{d-1}^{B/I-2} + \cdots\right) \\
&\quad + (\alpha_1 - \alpha_2)\left(p_{d-1}^{B/I} - \frac{B}{I}\alpha_0 p_{d-1}^{B/I-1} + \frac{B(B-I)}{2I^2}\alpha_0^2 p_{d-1}^{B/I-2} + \cdots\right) \\
&\quad + (\alpha_2 - \alpha_0)\left(p_{d-1}^{B/I} - \frac{B}{I}\alpha_1 p_{d-1}^{B/I-1} + \frac{B(B-I)}{2I^2}\alpha_1^2 p_{d-1}^{B/I-2} + \cdots\right) \\
&= \frac{B(B-I)p_{d-1}^{B-2}}{2I^2}\left(\alpha_0\alpha_2^2 - \alpha_1\alpha_2^2 + \alpha_1\alpha_0^2 - \alpha_2\alpha_0^2 + \alpha_2\alpha_1^2 - \alpha_0\alpha_1^2\right) \\
&\quad + \cdots,
\end{aligned}$$

where the remaining terms have lower valuations. Therefore, some cancelation occurs in the main term. Since $\nu(\alpha_2) < \nu(\alpha_1) < \nu(\alpha_0)$, we deduce that $B(B-I) = 0$, i.e., $B = 0$ or $B = I$. In both cases we obtain only trivial solutions. Therefore, Theorem 1 is proved.

Remark 1. Note that in the case of $d = 3$ we have $I = 1$. Therefore, we find $c_3 = 18 \cdot 1.031 = 18.558$. This improves the bound $c_3 = 34$ found by Ziegler [**25**, Theorem 1].

Acknowledgments. The author wants to thank Clemens Fuchs, who draw the attention of the author to this problem.

REFERENCES

1. J.W.S. Cassels, *An introduction to the geometry of numbers*, Springer-Verlag, Berlin, 1971. Second printing, corrected, *Die Grundlehren der mathematischen Wissenschaften*, Band 99.

2. C. Fuchs and B. Jadrijević, *On a parametric family of Thue inequalities over function fields*, Math. Proc. Cambridge Philos. Soc., to appear.

3. C. Fuchs and V. Ziegler, *On a family of Thue equations over function fields*, Monatsh. Math. **147** (2006), 11–23.

4. ———, *Thomas' family of Thue equations over function fields*, Quart. J. Math. **57** (2006), 81–91.

5. I. Gaál and M. Pohst, *Diophantine equations over global function fields.* I. *The Thue equation*, J. Number Theory **119** (2006), 49–65.

6. ———, *Diophantine equations over global function fields.* II. *R-integral solutions of Thue equations*, Experiment. Math. **15** (2006), 1–6.

7. B.P. Gill, *An analogue for algebraic functions of the Thue-Siegel theorem*, Annals Math. **31** (1930), 207–218.

8. F. Halter-Koch, G. Lettl, A. Pethő, and R.F. Tichy, *Thue equations associated with Ankeny-Brauer-Chowla number fields*, J. London Math. Soc. **60** (1999), 1–20.

9. C. Heuberger, *On a family of quintic Thue equations*, J. Symbolic Comput. **26** (1998), 173–185.

10. ———, *On a conjecture of E. Thomas concerning parametrized Thue equations*, Acta. Arith. **98** (2001), 375–394.

11. C. Heuberger, A. Togbé, and V. Ziegler, *Automatic solution of families of Thue equations and an example of degree* 8, J. Symbolic Computation **38** (2004), 1145–1163.

12. E. Lee, *Studies on Diophantine equations*, PhD thesis, Cambridge University, 1992.

13. G. Lettl, *Thue equations over algebraic function fields*, Acta Arith. **117** (2005), 107–123.

14. R.C. Mason, *On Thue's equation over function fields*, J. London Math. Soc. **24** (1981), 414–426.

15. ———, *Diophantine equations over function fields*, London Math. Soc. Lect. Note Series **96**, Cambridge University Press, Cambridge, 1984.

16. M. Mignotte and N. Tzanakis, *On a family of cubics*, J. Number Theory **39** (1991), 41–49.

17. C.F. Osgood, *An effective lower bound on the "diophantine approximation" of algebraic functions by rational functions*, Mathematika **20** (1973), 4–15.

18. A. Pethő, *Complete solutions to families of quartic Thue equations*, Math. Comp. **57** (1991), 777–798.

19. A. Pethő and R.F. Tichy, *On two-parametric quartic families of Diophantine problems*, J. Symbolic Comput. **26** (1998), 151–171.

20. M. Rosen, *Number theory in function fields*, Grad. Texts Math. **210**, Springer-Verlag, New York, 2002.

21. W.M. Schmidt, *Thue's equation over function fields*, J. Austral. Math. Soc. **25** (1978), 385–422.

22. E. Thomas, *Solutions to certain families of Thue equations*, J. Number Theory bf43 (1993), 319–369.

23. A. Thue, *Über Annäherungswerte algebraischer Zahlen*, J. Reine angew. Math. **135** (1909), 284–305.

24. I. Wakabayashi, *On a family of quartic Thue inequalities.* I, J. Number Theory **66** (1997), 70–84.

25. V. Ziegler, *Thomas' conjecture over function fields*, J. Théor. Nombres Bordeaux **19** (2007), 289–309.

Graz University of Technology, Institute for analysis and computational number theory, Steyrergasse 30, A-8010 Graz, Austria
Email address: ziegler@finanz.math.tugraz.at