ACTA ARITHMETICA
XLVII (1986)

On equations in two S-units over function fields of characteristic 0

by

JAN-HENDRIK EVERTSE (Amsterdam)

1. Introduction. Many finiteness results on the number of solutions of diophantine equations, such as the Thue–Mahler equation and linear equations in two S-units, have been proved by applying approximation techniques of Thue, Siegel and Roth. These techniques are ineffective in the sense that they do no not provide an algorithm for finding all solutions. However, these techniques can be used to find explicit upper bounds for the number of solutions of diophantine equations belonging to several classes. In [2] we derived an explicit upper bound for the number of solutions of linear equations in two S-units and, as an application of this, for the number of solutions of the Thue–Mahler equation with variables in a fixed algebraic number field. The techniques of Thue, Siegel and Roth can be extended to the case of algebraic function fields of characteristic 0. Thus analogues of the Thue–Siegel–Roth Theorem have been obtained for such function fields (cf. [7], Ch. 7, [5], [15]). In this paper we shall give analogues of the results in [2] for algebraic function fields of characteristic 0. We shall also prove a finiteness criterion for the number of solutions of the Thue equation over function fields of characteristic 0. This finiteness criterion was already proved by Mason by an effective method in case that the function field has transcendence degree 1 over its field of constants (cf. [8], Th. 2, [11], Ch. 2, Th. 2). Our results are valid for function fields with arbitrarily large transcendence degree over their fields of constants.

Our results mentioned above are in fact valid for a large class of fields, namely the fields with a product formula. Before we are able to state the results, we have to mention some facts about these fields. A field K is called a *field with a product formula* if it has characteristic 0 and if it is endowed with a set of pairwise non-equivalent (multiplicative) valuations $M_K = \{|\cdot|_v\}$ on K such that for all α in K^*: (1)

(1) If K is a field then we denote by K^* the set of non-zero elements of K.

(1) $|\alpha|_v = 1$ for all but finitely many v, the *finiteness property*;

(2) $$\prod_v |\alpha|_v = 1, \quad \text{the } \textit{product formula}.$$

If confusion can not arise, we shall not make a proper distinction between a valuation $|\cdot|_v$ and its index v. If all valuations in M_K are raised to the same power then we obtain another set of valuations satisfying (1) and (2) which is called equivalent to M_K. There are fields with inequivalent sets of valuations satisfying (1) and (2).

We shall now give three important examples of fields with a product formula.

1. *The field of rational numbers* $\boldsymbol{Q}$. Let P denote the set of prime numbers: thus every non-zero rational α can be written uniquely as

$$\alpha = \pm \prod_{p \in P} p^{w_p(\alpha)},$$

where the numbers $w_p(\alpha)$ are rational integers of which at most finitely many are non-zero. We define valuations $|\cdot|_p$ $(p \in P)$ and $|\cdot|_{p_\infty}$ by

$$|\alpha|_p = p^{-w_p(\alpha)} \quad \text{for} \quad \alpha \in \boldsymbol{Q}^*, \quad |0|_p = 0$$

for every p in P and

$$|\alpha|_{p_\infty} = |\alpha| \quad \text{for} \quad \alpha \in \boldsymbol{Q},$$

where $|\cdot|$ denotes the ordinary absolute value.

Now $\boldsymbol{Q}$ is a field with a product formula with set of valuations

$$M_{\boldsymbol{Q}} = \{|\cdot|_p\colon\ p \in P \cup \{p\}\}.$$

2. $K = \boldsymbol{k}(X_1, \ldots, X_n)$, *where* $\boldsymbol{k}$ *is a field of characteristic* 0 *and where* $X_1, \ldots, X_n$ *are independent variables*. Let I be a maximal set of pairwise non-associated irreducible polynomials in the ring $\boldsymbol{k}[X_1, \ldots, X_n]$. Every f in K^* can be written uniquely as

$$F = c \prod_{f \in I} f^{v_f(F)},$$

where $c \in \boldsymbol{k}^*$ and where the numbers $v_f(F)$ are rational integers of which at most finitely many are non-zero. If $F \in \boldsymbol{k}[X_1, \ldots, X_n]$ then we denote by $\mathrm{Deg}(F)$ the total degree of F. If $F \in K^*$ then we put $\mathrm{Deg}(F) = \mathrm{Deg}(F_1) - \mathrm{Deg}(F_2)$, where F_1, F_2 are polynomials in $\boldsymbol{k}[X_1, \ldots, X_n]$ with $F = F_1/F_2$. This is clearly well-defined. We define valuations $|\cdot|_f$ $(f \in I)$ and $|\cdot|_\infty$ by

$$|F|_f = e^{-\mathrm{Deg}(f) v_f(F)} \quad \text{for} \quad F \in K^*, \quad |0|_f = 0$$

for every f in I and

$$|F|_\infty = e^{\mathrm{Deg}(F)} \quad \text{for} \quad F \in K^*, \quad |0|_\infty = 0.$$

Then K is a field with a product formula with set of valuations

$$M_K = \{|\cdot|_v\colon\ v \in I \cup \{\infty\}\}.$$

3. *Finite extensions of fields with a product formula.* Let K be a field with a product formula with set of valuations M_K and let L be a finite extension of K of degree d. Let $v \in M_K$ and let $\overline{K_v}$ be the algebraic closure of the completion of K at v. Then there are exactly d injective homomorphisms of L in $\overline{K_v}$, $\sigma_1, \ldots, \sigma_d$, say. As is well known from valuation theory, $|\cdot|_v$ can be extended in exactly one way to $\overline{K_V}$. This extension is also denoted by $|\cdot|_v$. Every extension of $|\cdot|_v$ to L must be equal to at least one of the valuations $|\sigma_1(\cdot)|_v, \ldots, |\sigma_d(\cdot)|_v$ defined on L. A valuation $|\cdot|_V$ on L is called *normalized with respect to* $|\cdot|_v$ if

(3) $$|\alpha|_V = \prod_{i \in E_{V|v}} |\sigma_i(\alpha)|_v \quad \text{for} \quad \alpha \in L.$$

Here $E_{V|v}$ denotes the set of those i in $1, \ldots, d$ for which $|\sigma_i(\cdot)|_v$ is equivalent to $|\cdot|_V$. If $|\cdot|_V$ is normalized with respect to $|\cdot|_v$ then the restriction of $|\cdot|_V$ to K is equivalent to $|\cdot|_v$. Let $|\cdot|_{V_1}, \ldots, |\cdot|_{V_g}$ be the valuations on L which are normalized with respect to $|\cdot|_v$. Then the sets $E_{V_1|v}, \ldots, E_{V_g|v}$ are pairwise disjoint and have union $\{1, \ldots, d\}$. Hence $g \leqslant d$. Moreover, by (3),

(4) $$\prod_{i=1}^{g} |\alpha|_{V_i} = |N_{L/K}(\alpha)|_v \quad \text{for} \quad \alpha \in L.$$

Let M_L be the set of valuations on L which are normalized with respect to the valuations in M_K. We shall show that L is a field with a product formula with set of valuations M_L, in other words that M_L satisfies the finiteness property and the product formula. To this end, we need the following obvious fact:

(5) $\left\{\begin{array}{l}\text{Let } K_0 \text{ be a field with non-archimedean valuation } |\cdot|_v \text{ and let } \alpha \in K_0\\ \text{be a zero of the polynomial } f(X) = a_n X^n + \ldots + a_1 X + a_0 \text{ in } K_0[X].\\ \text{Suppose that } |a_i|_v \in \{0, 1\} \text{ for } i = 1, \ldots, n. \text{ Then } |\alpha|_v \in \{0, 1\}.\end{array}\right.$

Let $\beta \in L^*$ and let $f(X) = a_n X^n + \ldots + a_0$ be a polynomial in $K[X]$ with $f(\beta) = 0$. By applying (1) with $\alpha = 2$ we obtain that M_L contains at most finitely many archimedean valuations. Moreover, from (1) and the above arguments we infer that $|a_0|_V, \ldots, |a_n|_V \in \{0, 1\}$ for all but finitely many non-archimedean valuations $|\cdot|_V$ in M_L. Together with (5) this implies that $|\beta|_V = 1$ for all but finitely many V in M_L. By applying (4) and (2) with $\alpha = N_{L/K}(\beta)$, we obtain that $\prod_{V \in M_L} |\beta|_V = 1$.

Examples 1, 2 and 3 show that algebraic number fields and algebraic function fields of characteristic 0 in several variables are fields with a product formula. Artin and Whaples [1] showed that any field K which has a set of valuations M_K with at least one archimedean valuation satisfying the finiteness property and the product formula must be an algebraic number field. Fields with a product formula which are not algebraic number fields, namely those for which all valuations are non-archimedean, are called *F-fields*. Let K be an F-field. Then the set

$$k = \{\alpha \in K\colon |\alpha|_v \leqslant 1 \text{ for all } v \text{ in } M_K\}$$

is also a field, termed the *constant field of* K. Using (5) it is not difficult to show that $\boldsymbol{k}$ is algebraically closed in K. If L is a finite extension of K then L is also an F-field, the constant field of which is just the algebraic closure of $\boldsymbol{k}$ in L.

Let K be a field with a product formula with set of valuations M_K. For each finite subset S of M_K we denote by $\mathscr{U}_S$ the set of those α in K with the property that $|\alpha|_v = 1$ for all v in $M_K \backslash S$. $\mathscr{U}_S$ is a multiplicative group, the elements of which are called *S-units*. If K is an algebraic number field, then the groups $\mathscr{U}_S$ are finitely generated. If K is an F-field with constant field $\boldsymbol{k}$, then for each group $\mathscr{U}_S$, $\boldsymbol{k}^* \subset \mathscr{U}_S$ and $\mathscr{U}_S/\boldsymbol{k}^*$ is finitely generated.

In the theorems below, we give explicit upper bounds for the numbers of solutions of linear equations in two S-units of a field with a product formula.

THEOREM 1. *Let K be an algebraic number field of degree m with set of valuations M_K and let S be a finite subset of M_K of cardinality s, containing all archimedean valuations. For each pair λ, μ in K^*, the equation*

$$(6) \qquad \lambda x + \mu y = 1 \quad \text{in } S\text{-units } x, y$$

has at most $3 \times 7^{m+2s}$ solutions.

For the proof of this result we refer to [2]. In this paper we shall prove the following analogue for F-fields.

THEOREM 2. *Let K be an F-field with constant field $\boldsymbol{k}$ and let S be a finite subset of the set of valuations M_K. For each pair λ, μ in K^*, the equation*

$$(6) \qquad \lambda x + \mu y = 1 \quad \text{in } S\text{-units } x, y$$

has at most 2×7^{2s} solutions with $\lambda x/\mu y \notin \boldsymbol{k}$. As above, s denotes the cardinality of S.

Both Theorem 1 and Theorem 2 have several applications. By combining Theorems 1 and 2, Evertse and Győry [3] derived an upper bound for the number of solutions of the equation $\lambda x + \mu y = 1$ in elements x, y of a finitely generated multiplicative group G which is contained in a field K of characteristic 0. Unfortunately this bound depends on the choice of a

transcendence basis for K. As an application of this result, upper bounds have been obtained for the number of solutions of certain decomposable form equations over rings which are finitely generated over $\mathbf{Z}$. In [3], Evertse and Győry also derived an explicit upper bound for the number of power bases of the ring of integers of an algebraic number field. This has been generalized in several directions in [4].

In the present paper, we shall apply Theorem 2 to the Thue–Mahler equation over F-fields: the main results of these investigations are stated in the following section.

2. On the Thue–Mahler equation over F-fields. Henceforth we shall use the notation of Theorem 2: thus K denotes an F-field with set of valuations M_K and constant field $\boldsymbol{k}$ and S denotes a finite subset of M_K. Moreover, by $\mathcal{O}_S$ we shall denote the *ring of S-integers*, that is the ring of those α in K such that $|\alpha|_v \leqslant 1$ for all v in $M_K \backslash S$. Clearly, $\mathscr{U}_S$ is the unit group of $\mathcal{O}_S$. Let $F(X, Y)$ be a binary form of degree n with coefficients in $\mathcal{O}_S$, which has at least three distinct linear factors in some extension of K (such a form is termed *cubic divisible*). We shall consider the Thue–Mahler equation

$$(7) \qquad F(x, y) = \mu u \quad \text{in} \quad x, y \in \mathcal{O}_S,\ u \in \mathscr{U}_S,$$

where μ is a non-zero element of $\mathcal{O}_S$. Two solutions (x_1, y_1, u_1), (x_2, y_2, u_2) of (7) are called *dependent* if there is an α in $\mathscr{U}_S$ such that $x_2 = \alpha x_1$, $y_2 = \alpha y_1$, $u_2 = \alpha^n u_1$. Our aim is to give an upper bound for the maximal number of pairwise independent solutions of (7). This number is not always finite. We call the form F *degenerate* if there are $\alpha, \beta, \gamma, \delta$ in $\mathcal{O}_S$ with $\alpha\delta - \beta\gamma \neq 0$, an S-unit ε and a binary form $f(X, Y)$ with coefficients in $\boldsymbol{k}$ such that

$$(8) \qquad F(\alpha X + \beta Y, \gamma X + \delta Y) = \mu\varepsilon f(X, Y) \quad \text{identically in } X, Y.$$

If F satisfies (8) then we call a solution (x, y, u) of (7) *trivial* if it is dependent on a solution $(\alpha\xi + \beta\eta, \gamma\xi + \delta\eta, u')$ with $\xi, \eta \in \boldsymbol{k}$. There exists a one-to-one correspondence between the set of trivial solutions of (7) and the set of pairs $(u\xi, u\eta)$ with $u \in \mathscr{U}_S$ and $\xi, \eta \in \boldsymbol{k}$ for which $f(\xi, \eta) \neq 0$. Hence if F is degenerate then (7) has infinitely many pairwise independent solutions. We shall show that when F is degenerate, the set of trivial solutions of (7) does not depend on the choice of $\alpha, \beta, \gamma, \delta, \varepsilon, f$.

For each non-zero α in $\mathcal{O}_S$ we define $\omega(\alpha)$ to be the number of v in $M_K \backslash S$ with $|\alpha|_v \leqslant 1$. Then we have the following result.

THEOREM 3. *Let K be an F-field with set of valuations M_K and constant field $\boldsymbol{k}$ and let $F(X, Y)$ be a cubic divisible form of degree n with coefficients in $\mathcal{O}_S$, where S is a finite subset of M_K of cardinality s.*

(i) *If F is non-degenerate then (7) has at most $7^{2n^3(s+\omega(\mu))}$ pairwise independent solutions.*

(ii) *If F is degenerate then (7) has at most $7^{2n^3(s+\omega(\mu))}$ pairwise independent non-trivial solutions.*

A special case of (7) is the Thue equation

$$F(x, y) = \mu \quad \text{in} \quad x, y \in \mathcal{O}_S. \tag{9}$$

As a consequence of Theorem 3 we have

Corollary 1. *Let K, $\boldsymbol{k}$, S, μ, $F(X, Y)$ be as in Theorem 3. Then (9) has at most $n \times 7^{2n^3(s+\omega(\mu))}$ solutions, unless there are $\alpha, \beta, \gamma, \delta$ in O_S with $\alpha\delta - \beta\gamma \neq 0$ and a binary form $f(X, Y)$ with coefficients in $\boldsymbol{k}$ such that*

$$F(\alpha X+\beta Y, \gamma X+\delta Y) = \mu f(X, Y) \quad \textit{identically in } X, Y. \tag{10}$$

In that case (9) has at most $n \times 7^{2n^3(s+\omega(\mu))}$ solutions which are not of the type $(x, y) = (\alpha\xi+\beta\eta, \gamma\xi+\delta\eta)$, where ξ, η are elements of $\boldsymbol{k}$ with $f(\xi, \eta) = 1$.

This corollary can be proved by considering the solutions $(x, y, 1)$ of (7). First of all, two solutions $(x_1, y_1, 1)$, $(x_2, y_2, 1)$ of (7) are dependent if and only if $x_2 = \varrho x_1$, $y_2 = \varrho y_1$ for some nth root of unity ϱ in K. Hence there are at most n solutions $(x, y, 1)$ of (7) which are dependent on a given solution. This implies that (9) has at most $n \times 7^{2n^3(s+\omega(\mu))}$ solutions if either F is non-degenerate or F is degenerate but (7) has no trivial solutions $(x, y, 1)$. If F satisfies (8) and $(x, y, 1)$ is a trivial solution of (7) then there are $\theta \in \mathcal{U}_S$ and $\xi, \eta \in \boldsymbol{k}$ such that $x = \theta(\alpha\xi+\beta\eta)$, $y = \theta(\gamma\xi+\delta\eta)$ and $\varepsilon\theta^n f(\xi, \eta) = 1$. It is easy to check that (8) remains true with $\alpha', \beta', \gamma', \delta', \varepsilon', f'$ replacing $\alpha, \beta, \gamma, \delta, \varepsilon, f$, where $\alpha' = \theta\alpha$, $\beta' = \theta\beta$, $\gamma' = \theta\gamma$, $\delta' = \theta\delta$, $\varepsilon' = 1$, $f'(X, Y) = \varepsilon\theta^n f(X, Y)$. Hence we may assume that F satisfies (10). If F satisfies (10) then $(x, y, 1)$ is a trivial solution of (7) if and only if $x = \alpha\xi+\beta\eta$, $y = \gamma\xi+\delta\eta$ where ξ, η are elements of $\boldsymbol{k}$ with $f(\xi, \eta) = 1$. This completes the proof of Corollary 1. ∎

If $\boldsymbol{k}$ is algebraically closed then Corollary 1 implies that (9) has at most finitely many solutions if and only if F is not of the type (10). This result was already proved by Mason [8] in case that K is a function field of characteristic 0 in one variable.

We shall now consider a special case of Theorem 3. Let $\boldsymbol{k}$ be an arbitrary field of characteristic 0 and put $K = \boldsymbol{k}(X_1, \ldots, X_n)$ where $X_1, \ldots, X_n$ are independent variables. Let $\mathcal{O}$ denote the ring $\boldsymbol{k}[X_1, \ldots, X_n]$. For an ordered pair (x, y) in $\mathcal{O}^2$, we write $\gcd(x, y) = 1$ if $x = 0$, $y = 1$, or if x and y are coprime in the usual sense and the coefficient of x highest in lexicographic order is equal to 1. Thus for each pair x, y in K, not both zero, there is a unique z in K^* with $\gcd(xz, yz) = 1$.

Let $f_1, \ldots, f_t$ be irreducible polynomials in $\mathcal{O}$ such that $f_i/f_j \notin \boldsymbol{k}$ for $i \neq j$. Put $S = \{\infty, f_1, \ldots, f_t\}$ (where ∞ corresponds to the valuation $|\cdot|_\infty$ defined in example 2 of Section 1). Then

$$\mathcal{U}_S = \{cf_1^{k_1} \ldots f_t^{k_t}\colon c \in \boldsymbol{k}^*, k_1, \ldots, k_t \in \boldsymbol{Z}^*\} \text{ and } \mathcal{O}_S = \mathcal{O}\mathcal{U}_S.$$

Let $F(X, Y)$ $\mathcal{O}[X, Y]$ be a cubic divisible form of degree n. Then it is easy to show that F is degenerate (with $\mu = 1$) in the sense of (8) if and only if there are $\alpha, \beta, \gamma, \delta \in \mathcal{O}$ with $\alpha\delta - \beta\gamma \neq 0$, $\varepsilon \in \mathcal{U}_S \cap \mathcal{O}$ and a binary form $f(X, Y) \in \boldsymbol{k}[X, Y]$ such that

$$F(\alpha X+\beta Y, \gamma X+\delta Y) = \varepsilon f(X, Y) \quad \text{identically in } X, Y. \tag{11}$$

We shall consider the equation

$$F(x, y) = cf_1^{k_1} \ldots f_t^{k_t} \quad \begin{array}{l} \text{in } (x, y) \in \mathcal{O}^2,\ c \in \boldsymbol{k}^*,\ k_1, \ldots, k_t \in \boldsymbol{Z} \\ \text{with } \gcd(x, y) = 1,\ k_1 \geqslant 0, \ldots, k_t \geqslant 0. \end{array} \tag{12}$$

If F satisfies (11) then a solution $(x, y, c, k_1, \ldots, k_t)$ of (12) is called trivial if there are $\theta \in \mathcal{U}_S$ and $\xi, \eta \in \boldsymbol{k}$ such that $x = \theta(\alpha\xi+\beta\eta)$, $y = \theta(\gamma\xi+\delta\eta)$. The following result is a direct consequence of Theorem 3, with $s = t+1$ and $\omega(\mu) = 0$.

Corollary 2. (i) *If F is non-degenerate then (12) has at most $7^{2n^3(t+1)}$ solutions.*

(ii) *If F is degenerate then (12) has at most $7^{2n^3(t+1)}$ non-trivial solutions.*

3. Lemmata for the proof of Theorem 2. In [2] we reduced the equation $\lambda x+\mu y = 1$ in S-units x, y of an algebraic number field to a finite number of systems of inequalities involving differences of the type $\xi/\eta - \varrho$, where ϱ is a cube root of unity, $\xi = \lambda x - \varrho\mu y$ and $\eta = \lambda x - \varrho^2 \mu y$. To these systems of inequalities we applied an approximation method of Thue and Siegel in which certain hypergeometric polynomials were used. In the present paper, we shall instead of transforming x, y to ξ, η in the way described above, modify our hypergeometric polynomials in a proper way. This makes the proof less complicated. We mention that the method used in this paper can also be used to prove Theorem 1. We shall not work this out here.

In this section we shall develop the necessary tools for the proof of Theorem 2. Henceforth we shall use the notation of Theorem 2: thus K denotes an F-field with set of valuations M_K and constant field $\boldsymbol{k}$ and S a finite subset of M_K of cardinality s. Moreover, v_0, v_1, v_2 will denote non-zero elements of K. Let

$$\boldsymbol{P}^2(K) = \{\boldsymbol{x} = (x_0\colon x_1\colon x_2)\colon x_0, x_1, x_2 \in K, (x_0, x_1, x_2) \neq (0, 0, 0)\}$$

be the projective plane over K. If $\boldsymbol{x} \in \boldsymbol{P}^2(K)$ is given by $(x_0\colon x_1\colon x_2)$ then x_0, x_1, x_2 are called the homogeneous coordinates of $\boldsymbol{x}$. They are determined up to a common multiplicative factor in K^*. By $\boldsymbol{P}^2(\boldsymbol{k})$ we shall denote the set of points in $\boldsymbol{P}^2(K)$ of which the homogeneous coordinates can all be chosen in $\boldsymbol{k}$. Finally, let T be defined as the set of points $\boldsymbol{x} = (x_0\colon x_1\colon x_2)$ in $\boldsymbol{P}^2(K)$ satisfying

$$x_0+x_1+x_2 = 0; \quad \boldsymbol{x} \notin \boldsymbol{P}^2(\boldsymbol{k}); \quad |x_h|_v = |v_h|_v \text{ for } h \in \{0, 1, 2\} \text{ and } v \in M_K \backslash S \tag{13}$$

for some choice of x_0, x_1, x_2.

If $\lambda = -(v_0/v_2)$, $\mu = -(v_1/v_2)$ then the mapping $(x, y) \to (v_0 x\colon v_1 y\colon v_2)$

defines a one-to-one correspondence between the solutions of (6) and the elements of T. Hence the following result is equivalent to Theorem 2.

THEOREM 2′. *T has cardinality at most 2×7^{2s}.*

In fact, we shall prove Theorem 2′ instead of Theorem 2. To simplify our notation we put, for $a_1, \ldots, a_r$ in K and v in M_K,

$$\min_v(a_1, \ldots, a_r) = \min(|a_1|_v, \ldots, |a_r|_v),$$

$$\max_v(a_1, \ldots, a_r) = \max(|a_1|_v, \ldots, |a_r|_v).$$

If $\mathbf{x} = (x_0 : x_1 : x_2) \in \mathbf{P}^2(K)$ then the *height of* $\mathbf{x}$ is defined by

$$h(\mathbf{x}) = \prod_{v\in M_K} \max_v(x_0, x_1, x_2).$$

By the product formula, the height of $\mathbf{x}$ is independent of the choice of its homogeneous coordinates. By choosing one of the homogeneous coordinates equal to 1, one can show that

$$h(\mathbf{x}) \geqslant 1 \quad \text{for} \quad \mathbf{x}\in \mathbf{P}^2(K) \tag{14}$$

and

$$h(\mathbf{x}) = 1 \quad \text{if and only if} \quad \mathbf{x}\in \mathbf{P}^2(k). \tag{15}$$

In our arguments, the number

$$A = \prod_{v\notin S} \frac{(\max_v(v_0, v_1, v_2))^3}{|v_0 v_1 v_2|_v}$$

will play an important rôle. Here $v\notin S$ is an abbreviation of $v\in M_K\backslash S$.

For the proof of the following lemma we refer to [2], Lemma 4.

LEMMA 1. *Let B be a real number with $0 < B < 1$, let $q \geqslant 1$ be an integer and put $R(B) = (1-B)^{-1} B^{B/(B-1)}$. Then there exists a set $\mathscr{W}$ of cardinality at most $\max(1, (2B)^{-1}) R(B)^{q-1}$, consisting of tuples $(\Gamma_1, \ldots, \Gamma_q)$ with $\Gamma_j \geqslant 0$ for $j = 1, \ldots, q$ and $\sum_{j=1}^{q} \Gamma_j = B$ with the following property: for every set of reals $F_1, \ldots, F_q, \Lambda$ with $0 < F_j \leqslant 1$ for $j = 1, \ldots, q$ and $\prod_{j=1}^{q} F_j \leqslant \Lambda$ there exists a tuple $(\Gamma_1, \ldots, \Gamma_q)\in \mathscr{W}$ such that*

$$F_j \leqslant \Lambda^{\Gamma_j} \quad \text{for} \quad j = 1, \ldots, q.$$

In the following lemma we shall show, that each element of T satisfies one of a finite number of systems of inequalities to which we can apply our approximation method.

LEMMA 2. *Let B be a real number with $1/2 \leqslant B < 1$. Then there exists a set $\mathscr{W}_0$ of cardinality at most $3^s R(B)^{s-1}$, consisting of tuples $\big((i(v))_{v\in S}, (\Gamma_v)_{v\in S}\big)$ with $i(v)\in\{0, 1, 2\}$ and $\Gamma_v \geqslant 0$ for all v in S and $\sum_{v\in S}\Gamma_v = B$, with the following property: for every $\mathbf{x} = (x_0 : x_1 : x_2)$ in T there is a tuple $\big((i(v))_{v\in S}, (\Gamma_v)_{v\in S}\big)$ in $\mathscr{W}_0$ such that*

$$\frac{|x_{i(v)}|_v}{\max_v(x_0, x_1, x_2)} \leqslant \big(A h(\mathbf{x})^{-3}\big)^{\Gamma_v} \quad \text{for} \quad v\in S. \tag{16}$$

Proof. Let $\mathbf{x} = (x_0 : x_1 : x_2)\in T$. For each v in S, choose $i(v)$ from $\{0, 1, 2\}$ such that $|x_{i(v)}|_v = \min_v(x_0, x_1, x_2)$. Thus one obtains a tuple $(i(v))_{v\in S}$ which can be chosen from at most 3^s possibilities. If for some v in S, $i(v) = i$ and (i, j, k) is a permutation of $(0, 1, 2)$, then $|x_j|_v = |x_i + x_k|_v \leqslant |x_k|_v$ and similarly, $|x_k|_v \leqslant |x_j|_v$. Hence $|x_j|_v = |x_k|_v$. Together with the product formula this implies that

$$\prod_{v\in S} \frac{|x_{i(v)}|_v}{\max_v(x_0, x_1, x_2)} = \prod_{v\in S} \frac{|x_0 x_1 x_2|_v}{(\max_v(x_0, x_1, x_2))^3} = \Big(\prod_{v\notin S} \frac{(\max_v(x_0, x_1, x_2))^3}{|x_0 x_1 x_2|_v}\Big)\Big/ h(\mathbf{x})^3.$$

Since $\mathbf{x}\in T$, we can choose x_0, x_1, x_2 such that $|x_h|_v = |v_h|_v$ for $h = 0, 1, 2$ and $v\in M_K\backslash S$. Hence

$$\prod_{v\in S} \frac{|x_{i(v)}|_v}{\max_v(x_0, x_1, x_2)} = A h(\mathbf{x})^{-3}.$$

By Lemma 1, there is a set of tuples $(\Gamma_v)_{v\in S}$ of cardinality at most $R(B)^{s-1}$, with $\Gamma_v \geqslant 0$ for $v\in S$ and $\sum_{v\in S}\Gamma_v = B$, such that (16) holds for at least one of the tuples in that set. This proves Lemma 2. ∎

Let $B > 0$ and let $\big((i(v))_{v\in S}, (\Gamma_v)_{v\in S}\big)$ be a fixed tuple for which $i(v)\in\{0, 1, 2\}$ and $\Gamma_v \geqslant 0$ for $v\in S$ and $\sum_{v\in S}\Gamma_v = B$. We shall deal with the system of inequalities

$$\frac{|x_{i(v)}|_v}{\max_v(x_0, x_1, x_2)} \leqslant \big(A h(\mathbf{x})^{-3}\big)^{\Gamma_v} \quad \text{for} \quad v\in S \quad \text{in } \mathbf{x} = (x_0 : x_1 : x_2)\in T. \tag{17}$$

The following lemma states roughly, that large solutions of (17) can not be close together.

LEMMA 3. *Assume that $2/3 < B < 1$. Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{m+1}$ be distinct solutions of (17), ordered such that $h(\mathbf{x}_1) \leqslant h(\mathbf{x}_2) \leqslant \ldots \leqslant h(\mathbf{x}_{m+1})$. Then*

$$h(\mathbf{x}_{m+1}) \geqslant A^{(1-B)((3B-1)^m-1)/(3B-2)}\, h(\mathbf{x}_1)^{(3B-1)^m}.$$

Proof. We shall prove that for any pair of solutions $\mathbf{x}, \mathbf{y}$ of (17) with $\mathbf{x} \neq \mathbf{y}$ and $h(\mathbf{x}) \leqslant h(\mathbf{y})$,

$$h(y) \geqslant A^{1-B} h(x)^{3B-1}. \tag{18}$$

From this, Lemma 3 follows immediately by induction.

Let $x = (x_0 : x_1 : x_2)$, $y = (y_0 : y_1 : y_2)$ be distinct solutions of (17) with $h(x) \leqslant h(y)$. For all v in M_K we put

$$\Delta_v(x, y) = \frac{|x_i y_j - x_j y_i|_v}{\max_v(x_0, x_1, x_2) \cdot \max_v(y_0, y_1, y_2)},$$

where $i, j \in \{0, 1, 2\}$ and $i \neq j$. $\Delta_v(x, y)$ is independent of the choice for the homogeneous coordinates of x, y and, since $x_0 + x_1 + x_2 = y_0 + y_1 + y_2 = 0$, also independent of i, j. Moreover, by the product formula,

$$\prod_{v \in M_K} \Delta_v(x, y) = \frac{1}{h(x) h(y)}. \tag{19}$$

Firstly, let $v \in M_K \backslash S$. Choose p, q from $\{0, 1, 2\}$ such that $|x_p|_v = \min_v(x_0, x_1, x_2)$, $|y_q|_v = \min_v(y_0, y_1, y_2)$. Let k be an element of $\{0, 1, 2\}$ distinct from p, q. By a similar argument as in the proof of Lemma 2 we obtain $|x_k|_v = \max_v(x_0, x_1, x_2)$, $|y_k|_v = \max_v(y_0, y_1, y_2)$. Let i, j be integers such that (i, j, k) is a permutation of $(1, 2, 3)$. Since $x, y \in T$, we may assume that $|x_h|_v = |y_h|_v = |v_h|_v$ for $h = 0, 1, 2$. Therefore,

$$\Delta_v(x, y) \leqslant \frac{\max_v(x_i y_j, x_j y_i)}{\max_v(x_0, x_1, x_2) \cdot \max_v(y_0, y_1, y_2)} = \frac{|v_0 v_1 v_2|_v}{(\max_v(v_0, v_1, v_2))^3}.$$

By taking the product over all v in $M_K \backslash S$, we obtain

$$\prod_{v \notin S} \Delta_v(x, y) \leqslant A^{-1}. \tag{20}$$

Now let $v \in S$. Put $i = i(v)$ and let j be an element from $\{0, 1, 2\}$ distinct from i. By (17) we have

$$\begin{aligned}
\frac{|x_i y_j - x_j y_i|_v}{\max_v(x_0, x_1, x_2)\max_v(y_0, y_1, y_2)} &\leqslant \frac{\max_v(y_j x_i, x_j y_i)}{\max_v(x_0, x_1, x_2)\max_v(y_0, y_1, y_2)} \\
&\leqslant \max\left(\frac{|x_i|_v}{\max_v(x_0, x_1, x_2)}, \frac{|y_i|_v}{\max_v(y_0, y_1, y_2)}\right) \\
&\leqslant \max\left(\left(A h(x)^{-3}\right)^{\Gamma_v}, \left(A h(y)^{-3}\right)^{\Gamma_v}\right) \\
&= \left(A h(y)^{-3}\right)^{\Gamma_v}.
\end{aligned}$$

By taking the product over all v in S we obtain

$$\prod_{v \in S} \Delta_v(x, y) \leqslant \left(A h(x)^{-3}\right)^B.$$

By combining this with (19), (20) we infer that

$$\frac{1}{h(x) h(y)} \leqslant A^{-1}\left(A h(x)^{-3}\right)^B = A^{B-1} h(x)^{-3B}.$$

This proves (18). ∎

In our approximation method we shall need some auxiliary polynomials, some properties of which are stated in the lemma below.

LEMMA 4. *For each positive integer r, there exist homogeneous polynomials $F_r(X_1, Y_1, X_2, Y_2)$ of degree $3r+2$ and $W_r(X, Y)$ of degree r with rational coefficients which satisfy the following properties:*

$$\begin{aligned} F_r(X_1, Y_1, X_2, Y_2) &= Y_2 X_1^{2r+1} W_r(X_1, Y_1) \\ &\quad + (-1)^{r-1} X_2 Y_1^{2r+1} W_r(Y_1, X_1) \quad \text{for} \quad r \in \mathbf{N}; \end{aligned} \tag{21}$$

$$\begin{aligned} F_r(X_1, Y_1, X_2, Y_2) &= F_r(Y_1, -X_1 - Y_1, Y_2, -X_2 - Y_2) \\ &= F_r(-X_1 - Y_1, X_1, -X_2 - Y_2, X_2) \quad \text{for } r \in \mathbf{N}; \end{aligned} \tag{22}$$

(23) *if x_1, y_1, x_2, y_2 are elements of a field of characteristic 0 such that for some positive integer r,*

$$F_r(x_1, y_1, x_2, y_2) = F_{r+1}(x_1, y_1, x_2, y_2) = 0,$$

then both x_2, y_2 are equal to 0 or

$$x_1 y_1 (x_1 + y_1)(x_1^2 + x_1 y_1 + y_1^2) = 0.$$

Proof. In Lemma 6 of [2] we showed that for each positive integer r there are polynomials $A_r(X)$, $B_r(X)$, $V_r(X)$ of degree r with rational coefficients such that for all r,

$$A_r(X^3) - X B_r(X^3) = (1-X)^{2r+1} V_r(X);$$

$$B_r(X) = X^r A_r(1/X);$$

$$A_r(X) B_{r+1}(X) - A_{r+1}(X) B_r(X) = c_r (1-X)^{2r+1} \quad \text{with} \quad c_r \neq 0.$$

For any polynomial $f(X)$ of degree d with coefficients in some field of characteristic 0 we define the binary form $f^*(X, Y)$ by $X^d f(Y/X)$. Thus the properties of A_r, B_r, V_r mentioned above can be rewritten as

$$X A_r^*(X^3, Y^3) - Y A_r^*(Y^3, X^3) = (X-Y)^{2r+1} V_r^*(X, Y); \tag{24}$$

$$A_r^*(X, Y) A_{r+1}^*(Y, X) - A_{r+1}^*(X, Y) A_r^*(Y, X) = c_r (X-Y)^{2r+1}. \tag{25}$$

Obviously, (24) implies that

$$V_r^*(X, Y) = V_r^*(Y, X). \tag{26}$$

Let ϱ be a primitive cube root of unity. Put

$$U(X, Y) = X - \varrho Y, \quad V(X, Y) = X - \varrho^2 Y.$$

For each positive integer r we define the polynomials F_r, W_r by

$$F_r(X_1, Y_1, X_2, Y_2) = \frac{U_2 V_1 A_r^*(V_1^3, U_1^3) - V_2 U_1 A_r^*(U_1^3, V_1^3)}{(\varrho^2-\varrho)^{2r+1}},$$
$$W_r(X, Y) = V_r^*(\varrho^2 U(X, Y), \varrho V(X, Y)),$$

where U_i, V_i are abbreviations for $U(X_i, Y_i)$, $V(X_i, Y_i)$ respectively, for $i = 1, 2$. The coefficients of F_r, W_r have their coefficients in the field $\boldsymbol{Q}(\varrho)$. However, by (26) the coefficients of W_r are invariant under the $\boldsymbol{Q}$-automorphism of $\boldsymbol{Q}(\varrho)$ which maps ϱ onto ϱ^2. It is easy to check that also the coefficients of F_r are invariant under this automorphism. Hence the coefficients of F_r, W_r belong to $\boldsymbol{Q}$.

We shall first prove (21). By (26) we have

(27) $$V_r^*(U(X, Y), V(X, Y)) = V_r^*(V(X, Y), U(X, Y)) = V_r^*(X-\varrho^2 Y, X-\varrho Y) = (-1)^r V_r^*(\varrho^2 U(Y, X), \varrho V(Y, X)) = (-1)^r W_r(Y, X).$$

Moreover, by a straightforward computation,

$$F_r(X_1, Y_1, X_2, Y_2) = (\varrho^2-\varrho)^{-2r-1}\Big(\frac{U_2-V_2}{\varrho^2-\varrho}\big(\varrho^2 U_1 A_r^*(U_1^3, V_1^3) - \varrho V_1 A_r^*(V_1^3, U_1^3)\big) - \frac{\varrho^2 U_2 - \varrho V_2}{\varrho^2-\varrho}\big(U_1 A_r^*(U_1^3, V_1^3) - V_1 A_r^*(V_1^3, U_1^3)\big)\Big).$$

Together with (24) and (27) this implies that

$$F_r(X_1, Y_1, X_2, Y_2) = Y_2 X_1^{2r+1} W_r(X_1, Y_1) + (-1)^{r-1} X_2 Y_1^{2r+1} W_r(Y_1, X_1).$$

(22) follows easily from the fact that

$$U(Y, -X-Y) = \varrho U(X, Y), \qquad V(Y, -X-Y) = \varrho^2 V(X, Y),$$
$$U(-X-Y, X) = \varrho^2 U(X, Y), \qquad V(-X-Y, X) = \varrho V(X, Y).$$

We shall now prove (23). Suppose that for some positive integer r and some x_1, y_1, x_2, y_2 in some field of characteristic 0,

$$F_r(x_1, y_1, x_2, y_2) = F_{r+1}(x_1, y_1, x_2, y_2) = 0.$$

Let $u_i = U(x_i, y_i)$, $v_i = V(x_i, y_i)$ for $i = 1, 2$. Suppose that not both x_2, y_2 are zero. Then not both u_2, v_2 are zero. By assumption we have

$$v_1 A_r^*(v_1^3, u_1^3)\cdot u_2 - u_1 A_r^*(u_1^3, v_1^3)\cdot v_2 = 0,$$
$$v_1 A_{r+1}^*(v_1^3, u_1^3)\cdot u_2 - u_1 A_{r+1}^*(u_1^3, v_1^3)\cdot v_2 = 0.$$

Hence

$$v_1 u_1 \big(A_r^*(v_1^3, u_1^3) A_{r+1}^*(u_1^3, v_1^3) - A_{r+1}^*(v_1^3, u_1^3) A_r^*(u_1^3, v_1^3)\big) = 0.$$

Together with (25) this implies that

$$0 = c_r v_1 u_1 (v_1^3 - u_1^3)^{2r+1} = c_r(\varrho^2-\varrho)^{2r+1}(x_1^2 + x_1 y_1 + y_1^2)\big(x_1 y_1(x_1+y_1)\big)^{2r+1}.$$

This completes the proof of Lemma 4. ∎

In the following lemma, we shall show that (17) can not have many solutions of large height.

LEMMA 5. *Assume that* $5/6 < B < 1$. *Let* r_0 *be an integer with* $r_0 > \frac{2+2B-3B^2}{B(6B-5)}$, *let* m *be an integer with* $(3B-1)^{m+1} > 3r_0+4$ *and put*

$$f(B, r_0) = \frac{(2r_0+1)B(3B-1)+B}{3B(6B-5)r_0 - (6+6B-9B^2)},$$
$$g(B, m, r_0) = \frac{B+(B-1)(3B-1)\big((3B-1)^m-1\big)/(3B-2)}{(3B-1)^{m+1}-3r_0-4}.$$

Then (17) *has at most* m *solutions in* $\boldsymbol{x} \in T$ *with*

(28) $$h(\boldsymbol{x}) > A^{\max(f(B,r_0),\, g(B,m,r_0))}.$$

Proof. We assume that we can choose $m+1$ solutions of (17) which satisfy (28). We shall derive a contradiction from this assumption. Let $\boldsymbol{x} = \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{m+1} = \boldsymbol{y}$ be distinct solutions of (17) satisfying (28) such that $h(\boldsymbol{x}_1) \leqslant h(\boldsymbol{x}_2) \leqslant \ldots \leqslant h(\boldsymbol{x}_{m+1})$. Let $\boldsymbol{x} = (x_0 : x_1 : x_2)$, $\boldsymbol{y} = (y_0 : y_1 : y_2)$, $h_1 = h(\boldsymbol{x})$, $h_2 = h(\boldsymbol{y})$. By Lemma 3 we have

(29) $$h_2 \geqslant A^{(1-B)((3B-1)^{m-1}-1)/(3B-2)}\, h_1^{(3B-1)^m}.$$

From $h_1 > A^{g(B,m,r_0)}$ we infer that

$$A^{(1-B)(3B-1)((3B-1)^m-1)/(3B-2)}\, h_1^{(3B-1)^{m+1}} > A^B h_1^{3r_0+4}.$$

By combining this with (29), we obtain

$$h_2^{3B-1} > A^B h_1^{3r_0+4}.$$

Hence there exists an integer l with $l \geqslant r_0+1$ such that

(30) $$A^B h_1^{3l+1} < h_2^{3B-1} \leqslant A^B h_1^{3l+4}.$$

Let F_r, W_r ($r \in \boldsymbol{N}$) be the polynomials constructed in Lemma 5. Put

$$U_r = F_r(x_i, x_j, y_i, y_j) \quad \text{for} \quad r \in \boldsymbol{N},$$

where i, j are distinct elements of $\{0, 1, 2\}$. By (22) and the fact that $x_0+x_1+x_2 = y_0+y_1+y_2 = 0$, U_r is independent of i, j. Put $r = l$ if $U_l \neq 0$ and $r = l-1$ otherwise. Since $l \geqslant r_0+1$ we have

(31) $$r \geqslant r_0 \geqslant 1.$$

This shows that U_r is defined. As a consequence of (23) and the fact that $\boldsymbol{x}$, $\boldsymbol{y} \notin \boldsymbol{P}^2(k)$, U_r is non-zero. Finally we mention that, in view of (30),

$$A^B h_1^{3r+1} < h_2^{3B-1} \leqslant A^B h_1^{3r+7}. \tag{32}$$

For convenience we put $i(v) = 0$, $\Gamma_v = 0$ for all v in $M_K \backslash S$. Since $\boldsymbol{x}$, $\boldsymbol{y}$ are solutions of (17) we have

$$\frac{|x_{i(v)}|_v}{\max_v(x_0, x_1, x_2)} \leqslant (Ah_1^{-3})^{\Gamma_v}, \qquad \frac{|y_{i(v)}|_v}{\max_v(y_0, y_1, y_2)} \leqslant (Ah_2^{-3})^{\Gamma_v} \quad \text{for} \quad v \in M_K. \tag{33}$$

Let $v \in M_K$. Put $i := i(v)$ and let j be an element of $\{0, 1, 2\}$ different from i. By (21) we have

$$\begin{aligned} |U_r|_v &= |y_i x_j^{2r+1} W_r(x_i, x_j) + (-1)^{r-1} y_j x_i^{2r+1} W_r(x_j, x_i)|_v \\ &\leqslant \max_v\left(y_i x_j^{2r+1} W_r(x_i, x_j),\ y_j x_i^{2r+1} W_r(x_j, x_i)\right). \end{aligned}$$

Together with (33) and the fact that W_r has degree r, this shows that

$$\begin{aligned} |U_r|_v &\leqslant \max_v(y_0, y_1, y_2)\left(\max_v(x_0, x_1, x_2)\right)^{3r+1} \\ &\qquad \times \max\left(\frac{|y_i|_v}{\max_v(y_0, y_1, y_2)}, \left(\frac{|x_i|_v}{\max_v(x_0, x_1, x_2)}\right)^{2r+1}\right) \\ &\leqslant \max_v(y_0, y_1, y_2)\left(\max_v(x_0, x_1, x_2)\right)^{3r+1} \max\left(Ah_2^{-3}, (Ah_1^{-3})^{2r+1}\right)^{\Gamma_v}. \end{aligned}$$

Together with the fact that $U_r \neq 0$ and the product formula this implies that

$$1 = \prod_{v \in M_K} |U_r|_v \leqslant \max\left(h_1^{3r+1} h_2 (Ah_2^{-3})^B,\ h_1^{3r+1} h_2 (Ah_1^{-3})^{(2r+1)B}\right). \tag{34}$$

By the left-hand side inequality of (32) we have

$$h_1^{3r+1} h_2 (Ah_2^{-3})^B < 1. \tag{35}$$

Moreover, by the right-hand side inequality of (32),

$$\begin{aligned} &h_1^{3r+1} h_2 (Ah_1^{-3})^{(2r+1)B} \\ &\leqslant \left(A^{(2r+1)B(3B-1)+B} h_1^{3r+7+(3B-1)(3r+1-3B(2r+1))}\right)^{1/(3B-1)} = (A^{f(B,r)} h_1^{-1})^{h(B,r)}, \end{aligned}$$

where $h(B, r) = \left(3B(6B-5)r - (6+6B-9B^2)\right)/(3B-1)$. It is easy to check that $f(B, x)$ is decreasing in x for $x > (6+6B-9B^2)/3B(6B-5)$. Hence by (31) and our assumption on r_0 in the statement of the lemma, $h(B, r) > 0$ and $f(B, r) \leqslant f(B, r_0)$. Together with the inequality $h_1 > A^{f(B, r_0)}$ these facts show that

$$h_1^{3r+1} h_2 (Ah_1^{-3})^{(2r+1)B} \leqslant (A^{f(B,r)} h_1^{-1})^{h(B,r)} < 1.$$

However, this contradicts (34) and (35). Hence the assumption made at the beginning of the proof of this lemma, that (17) has at least $m+1$ solutions satisfying (28) must have been wrong. This completes the proof of Lemma 5. ∎

4. Proof of Theorem 2′. We shall now complete the proof of Theorem 2′, by combining the Lemmata 2, 3, 5: we shall henceforth choose the particular values $B = 0.846$, $r_0 = 34$ and $m = 10$. When referring to (17) or B we shall implicitly assume that $B = 0.846$.

First of all, we shall estimate the number of solutions of a fixed system (17) from above. With our choices for B, r_0, m we have $f(B, r_0) = 47.111\ldots$, $g(B, m, r_0) = -3.964\ldots$ Hence by Lemma 5 and the fact that $A \geqslant 1$, (17) has at most ten solutions in $\boldsymbol{x} \in T$ with

$$h(\boldsymbol{x}) > A^{48}.$$

We shall now give an upper bound for the number of remaining solutions of (17), namely those solutions in $\boldsymbol{x} \in T$ for which

$$1 < h(\boldsymbol{x}) \leqslant A^{48}. \tag{36}$$

If $A = 1$ then there are no such solutions. Therefore suppose that $A > 1$ and that there are t solutions $\boldsymbol{x}$ of (17) which satisfy (36), $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_t$, say, ordered such that $h(\boldsymbol{x}_1) \leqslant \ldots \leqslant h(\boldsymbol{x}_t)$. By Lemma 3 we have

$$A^{(1-B)/(3B-2)} h(\boldsymbol{x}_t) \geqslant \left(A^{(1-B)/(3B-2)} h(\boldsymbol{x}_1)\right)^{(3B-1)^{t-1}}.$$

By raising both sides of this inequality to the power $(3B-2)/(1-B)$ and using the inequalities $h(\boldsymbol{x}_1) > 1$, $h(\boldsymbol{x}_t) \leqslant A^{48}$, we obtain

$$A^{1+48(3B-2)/(1-B)} > A^{(3B-1)^{t-1}}.$$

In view of $A > 1$ and $B = 0.846$ this shows that $t < 12$. Hence (17) has at most 22 solutions.

By Lemma 2, every $\boldsymbol{x}$ in T satisfies one of at most

$$3^s R(0.846)^{s-1} \leqslant 3^s (49/3)^{s-1} \leqslant \frac{2}{22} \times 7^{2s}$$

systems of inequalities of type (17). Hence T has cardinality at most 2×7^{2s}. This completes the proof of Theorem 2′. ∎

We remark that with our method it is possible to replace the upper bound 2×7^{2s} by a bound of the type $C(B) \times \left(3R(B)\right)^s$, where B is any real number with $5/6 < B < 1$. It is easy to check that $R(B)$ is increasing in B and that $3R(5/6) = 44.789\ldots$ It is very likely that by applying a proper modification of Roth's method, one can derive a bound of the type $C(B) \times \left(3R(B)\right)^s$ for all B with $2/3 < B < 1$. This can give an improvement of our result since $3R(2/3) = 20.25$.

Let K be a function field which has transcendence degree 1 over its constant field $\boldsymbol{k}$. Then K is a finite extension of the field $\boldsymbol{k}(X)$, where X is an element of K which is transcendent over $\boldsymbol{k}$. Let M_K be the set of valuations on K which are normalized with respect to the valuations on $\boldsymbol{k}(X)$ as constructed in example 2 of Section 1. Then one can show that for each α in K and each v in M_K, $\log|\alpha_v| \in \boldsymbol{Z}$. Let g denote the genus of K and let S, s, T, ν_0, ν_1, ν_2, A have the same meaning as in Section 3. As an immediate consequence of Mason's effective result concerning S-units of a function field (cf. [8], [11], Ch. 1) one can show that for each $\boldsymbol{x}$ in T,

(37) $$\log h(\boldsymbol{x}) \leqslant 2g-2+s+s', \text{ (}^2\text{)}$$

where s' is the number of v in $M_K \backslash S$ for which the numbers $|\nu_0|_v$, $|\nu_1|_v$, $|\nu_2|_v$ are not all equal. Let v be an element of $M_K \backslash S$ such that $|\nu_0|_v$, $|\nu_1|_v$, $|\nu_2|_v$ are not equal. Suppose for convenience that $|\nu_0|_v \leqslant |\nu_1|_v \leqslant |\nu_2|_v$. Then $\log|\nu_2/\nu_0|_v \geqslant 1$. Hence

$$\log\frac{(\max_v(\nu_0, \nu_1, \nu_2))^3}{|\nu_0\nu_1\nu_2|_v} \geqslant \log\frac{\max_v(\nu_0, \nu_1, \nu_2)}{|\nu_0|_v} \geqslant 1.$$

This shows that $s' \leqslant \log A$. Together with (37) this implies that

$$\log h(\boldsymbol{x}) \leqslant 2g-2+\log A+s \quad \text{for all } \boldsymbol{x} \text{ in } T.$$

By combining this with Lemmata 2 and 3 it is possible to derive an upper bound for the cardinality of T of the type $C(B, g)\times(3R(B))^s$ for all B with $2/3 < B < 1$. If we combine Mason's effective result [12] on the equation

(38) $$\lambda_1 x_1 + \ldots + \lambda_n x_n = 1 \quad (n \geqslant 2)$$

in S-units of a function field with suitable generalizations of Lemmata 2 and 3, it might be possible to derive an upper bound for the numbers of solutions of (38) independent of $\lambda_1, \ldots, \lambda_n$.

5. Proof of Theorem 3. Let K be an F-field with set of valuations M_K and constant field $\boldsymbol{k}$. For any finite extension L of K, we denote by M_L the set of valuations on L which are normalized with respect to the valuations in M_K; thus L is an F-field with set of valuations M_L of which the constant field is just the algebraic closure of $\boldsymbol{k}$ in L. Similar to Theorem 3, S will denote a finite subset of M_K of cardinality s, μ will denote a non-zero element of $\mathcal{O}_S$ and $F(X, Y)$ will denote a cubic divisible binary form of degree n with coefficients in $\mathcal{O}_S$. We shall deal with the equation

(7) $$F(x, y) = \mu u \quad \text{in} \quad x, y \in \mathcal{O}_S,\ u \in \mathcal{U}_S.$$

(²) We mention that the valuations and heights used by Mason are just the logarithms of ours.

Let L be the splitting field of F over K, that is the smallest extension of K over which F factorizes into linear factors. Thus we have

$$F(X, Y) = aX^{n-m}\prod_{i=1}^{m}(X-\gamma_i Y),$$

where $a \in \mathcal{O}_S$, $2 \leqslant m \leqslant n$ and $\gamma_1, \ldots, \gamma_m \in L$. Suppose that (7) is soluble and let (x_0, y_0, u_0) be a fixed solution of (7). Put

$$l_1(X, Y) = F(x_0, y_0)\frac{X-\gamma_1 Y}{x_0-\gamma_1 y_0}, \quad l_i(X, Y) = \frac{X-\gamma_i Y}{x_0-\gamma_i y_0} \quad \text{for} \quad i = 2, \ldots, m,$$
$$l_i(X, Y) = X/x_0 \quad \text{for} \quad i = m+1, \ldots, n.$$

Let T be the set of valuations in M_L which are normalized with respect to the valuations in S or to the valuations $|\cdot|_v$ in $M_K \backslash S$ with $|\mu|_v < 1$. Then

(39) $$F(X, Y) = \prod_{i=1}^{n} l_i(X, Y),$$

(40) $l_i(x, y) \in \mathcal{U}_T$ for each solution (x, y, u) of (7) and for each i in $\{1, \ldots, n\}$.

(39) is obvious. In order to prove (40) we introduce the following notation: for any V in M_L and any binary form $G(X, Y) = \sum_{r=0}^{d} a_r X^{d-r} Y^r$ with coefficients in L we put

$$|G|_V = \max_V(a_0, \ldots, a_d).$$

Then we have, as a consequence of Proposition 2.1 of [7], Ch. 3, p. 55, that

(41) $$|F|_V = \prod_{i=1}^{n} |l_i|_V \quad \text{for} \quad V \in M_L.$$

Let $V \in M_L \backslash T$. By the fact that $F(X, Y)$ has its coefficients in $\mathcal{O}_T$, we have $|F|_V \leqslant 1$. On the other hand,

$$|l_i|_V \geqslant |l_i(x_0, y_0)|_V = 1 \quad \text{for} \quad i = 1, \ldots, n.$$

Together with (41) these facts show that $|l_i|_V = 1$ for $i = 1, \ldots, n$. We infer that l_i has its coefficients in $\mathcal{O}_T$ for $i = 1, \ldots, n$. Let (x, y, u) be a solution of (7). Then the numbers $l_i(x, y)$ $(i = 1, \ldots, n)$ belong to $\mathcal{O}_T$. However by (39), the product of these numbers is a unit in $\mathcal{O}_T$. Hence $l_i(x, y)$ is a unit in $\mathcal{O}_T$ for $i = 1, \ldots, n$. This proves (40).

Let (i, j, k) be a triple of integers in $1, \ldots, n$ such that the linear forms l_i, l_j, l_k are pairwise non-proportional (such a triple will be called *proper*). For each pair p, q in $1, \ldots, n$, let Δ_{pq} denote the determinant of the forms

l_p, l_q, that is the determinant of the linear transformation which maps the vector (x, y) in L'^2 onto the vector $(l_p(x, y), l_q(x, y))$. Then

(42) $\Delta_{jk} l_i(X, Y)+\Delta_{ki} l_j(X, Y)+\Delta_{ij} l_k(X, Y) = 0$ identically in X, Y.

Let L' be the smallest extension of K containing the coefficients of l_i, l_j, l_k and let T' be the set of valuations in $M_{L'}$, which are normalized with respect to the valuations in S or to the valuations $|\cdot|_v$ in $M_K \backslash S$ for which $|\mu|_v < 1$. Then the valuations in T' are, apart from equivalence, just the restrictions of the valuations in T to L'. Put $\Lambda = \Delta_{jk}/\Delta_{ji}$, $M = \Delta_{ki}/\Delta_{ji}$, $U(X, Y) = l_i(X, Y)/l_k(X, Y)$, $V(X, Y) = l_j(X, Y)/l_k(X, Y)$. For each solution (x, y, u) of (7) we have by (40) that both $U(x, y)$, $V(x, y)$ are T'-units and by (42) that

(43) $$\Lambda U(x, y)+MV(x, y) = 1.$$

It is easy to check that for any pair of solutions (x_1, y_1, u_1), (x_2, y_2, u_2) of (7), $(U(x_1, y_1), V(x_1, y_1)) = (U(x_2, y_2), V(x_2, y_2))$ if and only if (x_1, y_1, u_1) and (x_2, y_2, u_2) are dependent. Further, by the construction of the forms l_i L' has degree at most $n(n-1)(n-2)$ over K. Hence the size of T' is at most $n(n-1)(n-2)(s+\omega(\mu))$, where $\omega(\mu)$ is the number of valuations $|\cdot|_v$ in $M_K \backslash S$ with $|\mu|_v < 1$. By combining these facts with (43) and Theorem 2, we obtain that (7) has at most

$$2 \times 7^{2n(n-1)(n-2)(s+\omega(\mu))}$$

pairwise independent solutions (x, y, u) with

$$\frac{\Lambda U(x, y)}{MV(x, y)} = \frac{\Delta_{jk}}{\Delta_{ki}} \frac{l_i(x, y)}{l_j(x, y)} \notin \bar{k},$$

where $\bar{k}$ denotes the constant field of the splitting field L of F over K. Since there are at most n^3 proper triples, we conclude that (7) has at most

$$2n^3 \times 7^{2n(n-1)(n-2)(s+\omega(\mu))}$$

pairwise independent solutions which do not belong to Σ, where Σ is the set of solutions of (7) with the property that

(44) $$\frac{\Delta_{jk} l_i(x, y)}{\Delta_{ki} l_j(x, y)} \in \bar{k} \quad \text{for each proper triple } (i, j, k).$$

We shall show the following: if Σ contains at least three pairwise independent solutions of (7) then F is degenerate and moreover, Σ is just the set of trivial solutions of (7). This proves Theorem 3. For if $s+\omega(\mu) = 0$ then $\mathcal{O}_S = k$, whence F is degenerate and all solutions of (7) are trivial, while if $s+\omega(\mu) > 0$,

$$2+2n^3 \times 7^{2n(n-1)(n-2)(s+\omega(\mu))} \leqslant 7^{2n^3(s+\omega(\mu))}.$$

This shows also that when F is degenerate, the set of trivial solutions of (7) depends only on F. For if $l_1, \ldots, l_n$ are replaced by other linear forms with the property (39), then Σ does not change, hence Σ depends only on F.

Let (x_i, y_i, u_i) $(i = 1, 2, 3)$ be three pairwise independent solutions of (7) belonging to Σ. Define the linear form $l(X, Y)$ by

$$l(X, Y) = \sum_{i=1}^{n} \frac{l_i(X, Y)}{l_i(x_0, y_0)}.$$

Then

$$l(X, Y) = F(x_0, y_0)^{-1}\big(F_X(x_0, y_0)X+F_Y(x_0, y_0)Y\big),$$

where F_X, F_Y denote the partial derivatives of F at X, Y respectively. Hence l has its coefficients in K. Note that l is not identically 0, since $l(x_0, y_0) = n$. Hence there is an h in $\{1, 2\}$ such that $l(x_h, y_h) \neq 0$. For each pair i, j in $\{1, \ldots, n\}$ we have

(45) $$\frac{l_i(x_h, y_h)}{l_i(x_0, y_0)} \Big/ \frac{l_j(x_h, y_h)}{l_j(x_0, y_0)} \in \bar{k}.$$

This is obvious if l_i, l_j are proportional. If l_i, l_j are non-proportional then there exists a k such that (i, j, k) is a proper triple. (45) follows then easily from

$$\frac{\Delta_{jk} l_i(x_h, y_h)}{\Delta_{ki} l_j(x_h, y_h)} \in \bar{k}, \quad \frac{\Delta_{jk} l_i(x_0, y_0)}{\Delta_{ki} l_j(x_h, y_h)} \in \bar{k}.$$

Put

$$\lambda := l(x_h, y_h) = \sum_{i=1}^{n} \frac{l_i(x_h, y_h)}{l_i(x_0, y_0)}.$$

Then, by (45),

(46) $$\frac{l_i(x_h, y_h)}{l_i(x_0, y_0)} = c_i \lambda \quad \text{for some } c_i \in \bar{k} \quad (i = 1, \ldots, n).$$

By taking the product over all i we obtain

$$(c_1 \ldots c_n)\lambda^n = F(x_h, y_h)/F(x_0, y_0) = u_h/u_0 \in \mathcal{U}_S.$$

Moreover, since l has its coefficients in K, $\lambda \in K$. Hence $\lambda \in \mathcal{U}_S$. Put $\alpha = x_0$, $\gamma = y_0$, $\beta = x_h/\lambda$, $\delta = y_h/\lambda$, $\varepsilon = u_0$. Then $\alpha, \beta, \gamma, \delta \in \mathcal{O}_S$, $\varepsilon \in \mathcal{U}_S$ and $\alpha\delta-\beta\gamma \neq 0$. Moreover, by (46),

(47) $$\frac{l_i(\beta, \delta)}{l_i(\alpha, \gamma)} = c_i \in \bar{k} \quad \text{for} \quad i = 1, \ldots, n.$$

Let

$$f(X,\ Y) = (\mu\varepsilon)^{-1} F(\alpha X+\beta Y,\ \gamma X+\delta Y).$$

Then, by (47), (39),

$$f(X,\ Y) = F(\alpha,\ \gamma)^{-1} \prod_{i=1}^{n} l_i(\alpha X+\beta Y,\ \gamma X+\delta Y)$$
$$= \prod_{i=1}^{n} \{X+(l_i(\beta,\ \delta)/l_i(\alpha,\ \gamma))Y\} = \prod_{i=1}^{n} (X+c_i Y).$$

Hence f has its coefficients in $\bar{k}$. But obviously, f also has its coefficients in K. Thus f has its coefficients in $\boldsymbol{k}$. We conclude that F is degenerate.

For each pair p, q in $\{1, \ldots, n\}$, we have

$$\Delta_{pq} = \frac{l_p(\alpha,\ \gamma)\, l_q(\beta,\ \delta) - l_q(\alpha,\ \gamma)\, l_p(\beta,\ \delta)}{\alpha\delta-\beta\gamma} = \frac{l_p(\alpha,\ \gamma)\, l_q(\alpha,\ \gamma)}{\alpha\delta-\beta\gamma}(c_q-c_p).$$

Moreover, if (x, y, u) is a solution of (7) and if ξ, η are defined by $x = \alpha\xi+\beta\eta$, $y = \gamma\xi+\delta\eta$ then, for $r \in \{1, \ldots, n\}$,

$$l_r(X,\ Y) = l_r(\alpha,\ \gamma)(\xi+c_r\eta).$$

This shows that for each proper triple (i, j, k),

$$\frac{\Delta_{jk}\, l_i(x,\ y)}{\Delta_{ki}\, l_j(x,\ y)} = \frac{(c_k-c_j)(\xi+c_i\eta)}{(c_i-c_k)(\xi+c_j\eta)}.$$

Thus Σ is the set of solutions (x, y, u) of (7) for which

$$\frac{(c_k-c_j)(\xi+c_i\eta)}{(c_i-c_k)(\xi+c_j\eta)} \in \bar{k} \quad \text{for each proper triple } (i, j, k),$$

where $x = \alpha\xi+\beta\eta$, $y = \gamma\xi+\delta\eta$. If (x, y, u) is a trivial solution of (7) then either $\eta = 0$ or $\xi/\eta \in \boldsymbol{k}$. Hence $(x, y, u) \in \Sigma$. Now suppose that $(x, y, u) \in \Sigma$. Then either $\eta = 0$ or $\xi/\eta \in \bar{k}$. But both ξ, η belong to K. Hence either $\eta = 0$ or $\xi/\eta \in \boldsymbol{k}$. Therefore, $\xi = \theta\xi_0$, $\eta = \theta\eta_0$, where $\xi_0, \eta_0 \in \boldsymbol{k}$ and $\theta \in K$. However,

$$\mu u = F(x,\ y) = \mu\varepsilon f(\theta\xi_0,\ \theta\eta_0) = \mu\varepsilon\theta^n f(\xi_0,\ \eta_0).$$

Hence $\theta \in \mathcal{U}_S$. This shows that (x, y, u) is a trivial solution. Therefore, Σ is the set of trivial solutions of (7). ∎

References

[1] E. Artin and G. Whaples, *Axiomatic characterization of fields by the product formula for valuations*, Bull. Amer. Math. Soc. 51 (1945), pp. 469–492.

[2] J.-H. Evertse, *On equations in S-units and the Thue–Mahler equation*, Invent. Math. 75 (1984), pp. 561–584.

[3] J.-H. Evertse and K. Győry, *On unit equations and decomposable form equations*, J. Reine Angew. Math., to appear.

[4] — — *On the number of polynomials and integral elements of given discriminant*, preprint.

[5] B. P. Gill, *An analogue for algebraic functions of the Thue–Siegel theorem*, Ann. of Math. (2) 31 (1930), pp. 207–218.

[6] K. Győry, *Bounds for the solutions of norm form, discriminant form and index form equations in finitely generated integral domains*, Acta Math. Hung. 42 (1983), pp. 45–80.

[7] S. Lang, *Fundamentals of diophantine geometry*, Springer Verlag, New York etc. 1983.

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

[9] — *The hyperelliptic equation over function fields*, Math. Proc. Cambridge Phil. Soc. 93 (1983), pp. 219–230.

[10] — *Equations over function fields*, Proc. Conf. Journées Arithmétiques, Noordwijkerhout (1983), H. W. Jager, ed., Lecture Notes in Mathematics 1068, Springer Verlag, New York etc. 1984, pp. 149–157.

[11] — *Diophantine equations over function fields*, London Math. Soc. Lecture Note Series 96, Cambridge Univ. Press, 1984.

[12] — *Norm form equations*, J. Number Theory, to appear.

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

[14] — *Polynomial solutions of $F(x, y) = z^n$*, Queen's Papers in Pure Appl. Math. 54 (1980), pp. 33–65.

[15] S. Uchiyama, *Rational approximations to algebraic functions*, J. Fac. Sci. Hokkaido Univ. 15 (1961), pp. 173–192.

CENTRE FOR MATHEMATICS AND
COMPUTER SCIENCE
KRUISLAAN 413
P.O. BOX 4079
NL-1009 AB AMSTERDAM
THE NETHERLANDS

Received on 25.10.1984
and in revised form on 26.6.1985 (1467)