ACTA ARITHMETICA
133.2 (2008)

Linear equations with unknowns from a multiplicative group in a function field

by

JAN-HENDRIK EVERTSE (Leiden) and UMBERTO ZANNIER (Pisa)

To Professor Wolfgang Schmidt on his 75th birthday

1. Introduction. Let K be a field of characteristic 0, and n an integer ≥ 2. Denote by $(K^*)^n$ the n-fold direct product of the multiplicative group K^*. Thus, the group operation of $(K^*)^n$ is coordinatewise multiplication $(x_1, \ldots, x_n) \cdot (y_1, \ldots, y_n) = (x_1y_1, \ldots, x_ny_n)$. We write $(x_1, \ldots, x_n)^m := (x_1^m, \ldots, x_n^m)$ for $m \in \mathbb{Z}$. We will often denote elements of $(K^*)^n$ by bold face characters $\mathbf{x}$, $\mathbf{y}$, etc.

Evertse, Schlickewei and Schmidt [3] proved that if Γ is a subgroup of $(K^*)^n$ of finite rank r and $a_1, \ldots, a_n$ are non-zero elements of K, then the equation

$$(1.1) \qquad a_1x_1 + \cdots + a_nx_n = 1 \quad \text{in } \mathbf{x} = (x_1, \ldots, x_n) \in \Gamma$$

has at most $e^{(6n)^{3n}(r+1)}$ non-degenerate solutions, i.e., solutions with

$$(1.2) \qquad \sum_{i \in I} a_ix_i \neq 0 \quad \text{for each proper, non-empty subset } I \text{ of } \{1, \ldots, n\}.$$

In the present paper, we derive a function field analogue of this result. Thus, let k be an algebraically closed field of characteristic 0 and let K be a transcendental field extension of k, where we allow the transcendence degree to be arbitrarily large. Let Γ be a subgroup of $(K^*)^n$ such that $(k^*)^n \subset \Gamma$ and such that $\Gamma/(k^*)^n$ has finite rank. This means that there are $\mathbf{a}_1, \ldots, \mathbf{a}_r \in \Gamma$ such that for every $\mathbf{x} \in \Gamma$ there are integers $m, z_1, \ldots, z_r$ with $m > 0$ and $\xi \in (k^*)^n$ such that $\mathbf{x}^m = \xi \cdot \mathbf{a}_1^{z_1} \cdots \mathbf{a}_r^{z_r}$. If $\Gamma = (k^*)^n$ then $\Gamma/(k^*)^n$ has rank 0; otherwise, $\mathrm{rank}(\Gamma/(k^*)^n)$ is the smallest r for which there exist $\mathbf{a}_1, \ldots, \mathbf{a}_r$ as above.

2000 *Mathematics Subject Classification*: Primary 11D72.
Key words and phrases: Diophantine equations over function fields.

We deal again with equation (1.1) in solutions $(x_1, \ldots, x_n) \in \Gamma$ with coefficients $a_1, \ldots, a_n \in K^*$. We mention that in the situation we are considering now, (1.1) might have infinitely many non-degenerate solutions. But one can show that the set of non-degenerate solutions of (1.1) is contained in finitely many $(k^*)^n$-cosets, i.e., in finitely many sets of the shape $\mathbf{b} \cdot (k^*)^n = \{\mathbf{b} \cdot \xi : \xi \in (k^*)^n\}$ with $\mathbf{b} \in \Gamma$. More precisely, we prove the following:

THEOREM. *Let k be an algebraically closed field of characteristic* 0, *let K be a transcendental extension of k, let $n \geq 2$, let $a_1, \ldots, a_n \in K^*$, and let Γ be a subgroup of $(K^*)^n$ satisfying*

$$(k^*)^n \subset \Gamma, \quad \operatorname{rank}(\Gamma/(k^*)^n) = r. \tag{1.3}$$

Then the set of non-degenerate solutions of equation (1.1) *is contained in the union of not more than*

$$\sum_{i=2}^{n+1} \binom{i}{2}^r - n + 1 \tag{1.4}$$

$(k^*)^n$-cosets.

We mention that Bombieri, Mueller and Zannier [1] by means of a new approach gave a rather sharp upper bound for the number of solutions of polynomial-exponential equations in one variable over function fields. Their approach and result were extended by Zannier [5] to polynomial-exponential equations over function fields in several variables. Our proof heavily uses the arguments from this last paper.

Let us consider the case $n = 2$, that is, let us consider the equation

$$a_1x_1 + a_2x_2 = 1 \quad \text{in } (x_1, x_2) \in \Gamma, \tag{1.5}$$

where Γ, a_1, a_2 satisfy the hypotheses of the Theorem with $n = 2$. It is easy to check that all solutions (x_1, x_2) of (1.5) with $a_1x_1/a_2x_2 \in k^*$ (if any such exist) lie in the same $(k^*)^2$-coset, while any two different solutions (x_1, x_2) with $a_1x_1/a_2x_2 \notin k^*$ lie in different $(k^*)^2$-cosets. So our Theorem implies that (1.5) has at most 3^r solutions (x_1, x_2) with $a_1x_1/a_2x_2 \notin k^*$. This is a slight extension of a result by Zannier [5] who obtained the same upper bound, but for groups $\Gamma = \Gamma_1 \times \Gamma_1$ where Γ_1 is a subgroup of K^*.

The formulation of our Theorem was inspired by Mueller [4]. She proved that if S is a finite set of places of the rational function field $k(z)$, if $\Gamma = U_S^n$ is the n-fold direct product of the group of S-units in $k(z)^*$, and if $a_1, \ldots, a_n \in k(z)^*$, then the set of non-degenerate solutions of (1.1) is contained in the union of not more than $(e(n+1)!/2)^{n(2|S|+1)}$ $(k^*)^n$-cosets.

Evertse and Győry [2] also considered equation (1.1) with $\Gamma = U_S^n$, but in the more general situation that S is a finite set of places in any finite extension K of $k(z)$. They showed that if K has genus g and if $a_1, \ldots, a_n$

$\in K^*$ then the set of solutions $\mathbf{x} \in U_S^n$ of (1.1) with $(a_1x_1, \ldots, a_nx_n) \notin (k^*)^n$ is contained in the union of not more than

$$\log(g+2) \cdot (e(n+1))^{(n+1)|S|+2}$$

proper linear subspaces of K^n.

We mention that in general $\operatorname{rank} U_S^n \le n(|S|-1)$ but that in contrast to number fields, equality need not hold. From our Theorem we can deduce the following result, which removes the dependence on the genus g, and replaces the dependence on $|S|$ by one on the rank.

COROLLARY. *Let k, K, n, $a_1, \ldots, a_n$, Γ, r be as in the Theorem. Then the set of solutions of* (1.1) *with* $(a_1x_1, \ldots, a_nx_n) \notin (k^*)^n$ *is contained in the union of not more than*

$$\sum_{i=2}^{n+1} \binom{i}{2}^r + 2^n - 2n - 1 \tag{1.6}$$

proper linear subspaces of K^n.

In Section 2 we prove some auxiliary results for formal power series, in Section 3 we prove our Theorem in the case that K has transcendence degree 1 over k, in Section 4 we extend this to the general case that K is an arbitrary transcendental extension of k, and in Section 5 we deduce the Corollary.

2. Results for formal power series. Let k be an algebraically closed field of characteristic 0. Let z be an indeterminate. Denote as usual by $k[[z]]$ the ring of formal power series over k and by $k((z))$ its quotient field. Thus, $k((z))$ consists of series $\sum_{i \ge i_0} c_i z^i$ with $i_0 \in \mathbb{Z}$ and $c_i \in k$ for $i \ge i_0$. We endow $k((z))$ with a derivation $\frac{d}{dz} : \sum_{i \ge i_0} c_i z^i \mapsto \sum_{i \ge i_0} i c_i z^{i-1}$. Let $1 + zk[[z]]$ denote the set of all formal power series of the shape $1 + c_1 z + c_2 z^2 + \cdots$ with $c_1, c_2, \ldots \in k$. Clearly, $1 + zk[[z]]$ is a multiplicative group. For $f \in 1 + zk[[z]]$, $u \in k$ we define

$$f^u := \sum_{i=0}^{\infty} \binom{u}{i} (f-1)^i, \tag{2.1}$$

where $\binom{u}{0} = 1$ and $\binom{u}{i} = u(u-1)\cdots(u-i+1)/i!$ for $i > 0$. Thus, f^u is a well-defined element of $1 + zk[[z]]$. This definition of f^u coincides with the usual one for $u = 0, 1, 2, \ldots$. We have $\frac{d}{dz} f^u = u f^{u-1} \frac{d}{dz} f$, and moreover,

$$\begin{cases} (fg)^u = f^u g^u & \text{for } f, g \in 1 + zk[[z]],\ u \in k; \\ f^{u+v} = f^u f^v \text{ and } (f^u)^v = f^{uv} & \text{for } f \in 1 + zk[[z]],\ u, v \in k. \end{cases} \tag{2.2}$$

(One may verify (2.2) by taking logarithmic derivatives and recalling that two series in $1 + zk[[z]]$ are equal if and only if their logarithmic derivatives are equal.) We endow $(1+zk[[z]])^r$ with the usual coordinatewise multiplication. Given $\mathbf{B} = (b_1, \dots, b_r) \in (1+zk[[z]])^r$, we define $\mathbf{B}^u := (b_1^u, \dots, b_r^u)$ for $u \in k$ and $\mathbf{B}^{\mathbf{u}} := b_1^{u_1} \cdots b_r^{u_r}$ for $\mathbf{u} = (u_1, \dots, u_r) \in k^r$. Thus, $\mathbf{B}^u \in (1+zk[[z]])^r$ and $\mathbf{B}^{\mathbf{u}} \in 1+zk[[z]]$.

Let h, r be integers with $h \geq 2$, $r \geq 1$. Further, let $a_1, \dots, a_h$ be elements of $k[[z]]$ which are algebraic over the field of rational functions $k(z)$ and which are not divisible by z, and let α_{ij} $(i = 1, \dots, h,\ j = 1, \dots, r)$ be elements of $1+zk[[z]]$ which are algebraic over $k(z)$. Put $\mathbf{A}_i := (\alpha_{i1}, \dots, \alpha_{ir})$ $(i = 1, \dots, h)$. Define

$$R := \{\mathbf{u} \in k^r : a_1\mathbf{A}_1^{\mathbf{u}}, \dots, a_h\mathbf{A}_h^{\mathbf{u}} \text{ are linearly dependent over } k\}.$$

By a *class* we mean a set $R' \subset k^r$ with the property that there are a subset J of $\{1, \dots, h\}$ and $\mathbf{u}_0 \in \mathbb{Q}^r$ such that for every $\mathbf{u} \in R'$,

$$\begin{cases} a_i\mathbf{A}_i^{\mathbf{u}} \ (i \in J) \text{ are linearly dependent over } k; \\ (\mathbf{A}_i\mathbf{A}_j^{-1})^{\mathbf{u}-\mathbf{u}_0} = 1 \text{ for all } i, j \in J. \end{cases} \tag{2.3}$$

LEMMA 1. *R is the union of finitely many classes.*

Proof. This is basically a special case of [5, Lemma 1]. In the proof of that lemma, it was assumed that $k = \mathbb{C}$, and that the a_i and α_{ij} are holomorphic functions in the variable z which are algebraic over $\mathbb{C}(z)$ and which are defined and have no zeros on a simply connected open subset Ω of $\mathbb{C}$. It was shown that provided $k = \mathbb{C}$, this was no loss of generality. The argument remains precisely the same if one allows k to be an arbitrary algebraically closed field of characteristic 0 and if one takes for the a_i power series from $k[[z]]$ which are algebraic over $k(z)$ and which are not divisible by z, and for the α_{ij} power series from $1+zk[[z]]$ which are algebraic over $k(z)$.

We mention that in [5] the definition of a class is slightly different from (2.3), allowing $(\mathbf{A}_i\mathbf{A}_j^{-1})^{\mathbf{u}-\mathbf{u}_0} \in k^*$ for all $i, j \in J$. But in our situation this implies automatically that $(\mathbf{A}_i\mathbf{A}_j^{-1})^{\mathbf{u}-\mathbf{u}_0} = 1$ since $(\mathbf{A}_i\mathbf{A}_j^{-1})^{\mathbf{u}-\mathbf{u}_0} \in 1+zk[[z]]$. ■

We now impose some further restriction on the α_{ij} and prove a more precise result. Namely, we assume that

$$\{\mathbf{u} \in k^r : (\mathbf{A}_i \cdot \mathbf{A}_h^{-1})^{\mathbf{u}} = 1 \text{ for } i = 1, \dots, h\} = \{\mathbf{0}\}. \tag{2.4}$$

Let S be the set of $\mathbf{u} \in k^r$ such that there are $\xi_1, \dots, \xi_h \in k$ with

$$\sum_{i=1}^{h} \xi_i a_i \mathbf{A}_i^{\mathbf{u}} = 0; \tag{2.5}$$

$$\sum_{i\in I} \xi_i a_i \mathbf{A}_i^{\mathbf{u}} \neq 0 \quad \text{for each proper, non-empty subset } I \text{ of } \{1, \dots, h\}. \tag{2.6}$$

LEMMA 2. *Assume (2.4). Then S is finite.*

Proof. We prove a slightly stronger statement. We partition $\{1,\ldots,h\}$ into subsets $I_1,\ldots,I_s$ such that $\mathbf{A}_i=\mathbf{A}_j$ if and only if i,j belong to the same set I_l for some $l\in\{1,\ldots,s\}$. Let $\widetilde{S}$ be the set of $\mathbf{u}\in k^r$ for which there exist $\xi_1,\ldots,\xi_n\in k$ such that (2.5) holds and, instead of (2.6),

$$\sum_{i\in I}\xi_i a_i\mathbf{A}_i^{\mathbf{u}}\neq 0 \tag{2.7}$$

holds for each proper, non-empty subset I of $\{1,\ldots,h\}$ which is a union of some of the sets $I_1,\ldots,I_s$. We prove that $\widetilde{S}$ is finite. This clearly suffices.

We proceed by induction on $p:=h+s$. Notice that from assumption (2.4) it follows that $h\geq 2$ and $s\geq 2$. First let $h=2$, $s=2$, i.e., $p=4$. Thus, $\widetilde{S}$ is the set of $\mathbf{u}\in k^r$ for which there are non-zero $\xi_1,\xi_2\in k$ with $\xi_1a_1\mathbf{A}_1^{\mathbf{u}}+\xi_2a_2\mathbf{A}_2^{\mathbf{u}}=0$. Then for $\mathbf{u}\in\widetilde{S}$ we have

$$(\mathbf{A}_1\cdot\mathbf{A}_2^{-1})^{\mathbf{u}}=\xi(a_2a_1^{-1})$$

with $\xi\in k^*$. Consequently, $(\mathbf{A}_1\cdot\mathbf{A}_2^{-1})^{\mathbf{u}_2-\mathbf{u}_1}\in k^*$ for any $\mathbf{u}_1,\mathbf{u}_2\in\widetilde{S}$. But then for $\mathbf{u}_1,\mathbf{u}_2\in\widetilde{S}$ we must have $(\mathbf{A}_1\cdot\mathbf{A}_2^{-1})^{\mathbf{u}_2-\mathbf{u}_1}=1$ since $(\mathbf{A}_1\cdot\mathbf{A}_2^{-1})^{\mathbf{u}_2-\mathbf{u}_1}\in 1+zk[[z]]$. In view of assumption (2.4) this implies that $\widetilde{S}$ consists of at most one element.

Now let $p>4$ and assume Lemma 2 to be true for all pairs (h,s) with $h\geq 2$, $s\geq 2$ and $h+s<p$. We apply Lemma 1 above. Clearly, $\widetilde{S}$ is contained in the set R dealt with in Lemma 1, and therefore, $\widetilde{S}$ is the union of finitely many sets $\widetilde{S}\cap R'$ where R' is a class as in (2.3). So we have to show that each such set $\widetilde{S}\cap R'$ is finite.

Thus let $S':=\widetilde{S}\cap R'$, where R' is a typical one among these sets. Let J be the corresponding subset of $\{1,\ldots,h\}$, and $\mathbf{u}_0\in\mathbb{Q}^r$ the corresponding vector, such that (2.3) holds. We distinguish two cases. First suppose that J is contained in some set I_l. Then the elements a_j $(j\in J)$ are linearly dependent over k. There is a proper subset J' of J such that a_j $(j\in J')$ are linearly independent over k and such that each a_j with $j\in J\setminus J'$ can be expressed as a linear combination over k of the a_j with $j\in J'$. By substituting these linear combinations into (2.5), (2.7), we obtain similar conditions, but with I_l replaced by the smaller set obtained by removing from I_l the elements from $J\setminus J'$. This reduces the number h of terms. Further, condition (2.4) remains valid. Thus we may apply the induction hypothesis, and conclude that S' is finite.

Now assume that J is not contained in one of the sets I_l. We transform our present situation into a new one with instead of $I_1,\ldots,I_s$ a partition of $\{1,\ldots,h\}$ into fewer than s sets. Then again, the induction hypothesis is applicable.

There are $i, j \in J$ with $\mathbf{A}_i \neq \mathbf{A}_j$, say $i \in I_{l_1}$ and $j \in I_{l_2}$. Further, there is $\mathbf{u}_0 \in \mathbb{Q}^r$ such that $(\mathbf{A}_i \mathbf{A}_j^{-1})^{\mathbf{u}-\mathbf{u}_0} = 1$ for $\mathbf{u} \in S'$. According to an argument in the proof of Lemma 1 of [5], the set of $\mathbf{u} \in k^r$ with $(\mathbf{A}_i \mathbf{A}_j^{-1})^{\mathbf{u}} = 1$ is a linear subspace V of k^r which is defined over $\mathbb{Q}$. Let $\mathbf{v}_1, \ldots, \mathbf{v}_{r'}$ be a basis of V contained in $\mathbb{Z}^r$. Thus, each $\mathbf{u} \in S'$ can be expressed uniquely as

$$\mathbf{u}_0 + w_1 \mathbf{v}_1 + \cdots + w_{r'} \mathbf{v}_{r'} \quad \text{with } \mathbf{w} = (w_1, \ldots, w_{r'}) \in k^{r'}. \tag{2.8}$$

Now define

$$b_q := a_q \mathbf{A}_q^{\mathbf{u}_0}, \quad \mathbf{B}_q := (\mathbf{A}_q^{\mathbf{v}_1}, \ldots, \mathbf{A}_q^{\mathbf{v}_{r'}}) \quad (q = 1, \ldots, h).$$

Thus, for $\mathbf{u} \in S'$ we have

$$a_q \mathbf{A}_q^{\mathbf{u}} = b_q \mathbf{B}_q^{\mathbf{w}} \quad \text{for } q = 1, \ldots, h. \tag{2.9}$$

Clearly, $b_q \in k[[z]]$ and the coordinates of $\mathbf{B}_q$ belong to $1 + zk[[z]]$, for $q = 1, \ldots, h$. Further, b_q and the coordinates of $\mathbf{B}_q$ $(q = 1, \ldots, h)$ are algebraic over $k(z)$ since $\mathbf{u}_0 \in \mathbb{Q}^r$ and since $\mathbf{v}_1, \ldots, \mathbf{v}_{r'} \in \mathbb{Z}^r$.

From the definition of $\mathbf{B}_q$ $(q = 1, \ldots, h)$ it follows that if $(\mathbf{B}_q \mathbf{B}_h^{-1})^{\mathbf{w}} = 1$ for $q = 1, \ldots, h$, then $(\mathbf{A}_q \mathbf{A}_h^{-1})^{\sum_j w_j \mathbf{v}_j} = 1$ for $q = 1, \ldots, h$, which by (2.4) implies $\sum_j w_j \mathbf{v}_j = \mathbf{0}$ and so $\mathbf{w} = \mathbf{0}$. Therefore, condition (2.4) remains valid if we replace $\mathbf{A}_q$ by $\mathbf{B}_q$ for $q = 1, \ldots, h$.

It is important to notice that $\mathbf{B}_{q_1} = \mathbf{B}_{q_2}$ for any $q_1, q_2 \in I_{l_1} \cup I_{l_2}$. Further, for each $l \neq l_1, l_2$, we have $\mathbf{B}_{q_1} = \mathbf{B}_{q_2}$ for any $q_1, q_2 \in I_l$.

Lastly, if $\mathbf{u} \in S'$ then by substituting (2.9) into (2.5), (2.7), we deduce that there are $\xi_1, \ldots, \xi_h \in k^*$ such that $\sum_{q=1}^h \xi_q b_q \mathbf{B}_q^{\mathbf{w}} = 0$ and $\sum_{q \in I} \xi_q b_q \mathbf{B}_q^{\mathbf{w}} \neq 0$ for each proper subset I of $\{1, \ldots, h\}$ which is a union of some of the sets from $I_{l_1} \cup I_{l_2}$, I_l $(l = 1, \ldots, s,\ l \neq l_1, l_2)$. Thus, each $\mathbf{u} \in S'$ corresponds by means of (2.8) to $\mathbf{w} \in k^{r'}$ which satisfies similar conditions as $\mathbf{u}$, but with the partition $I_1, \ldots, I_s$ of $\{1, \ldots, h\}$ replaced by a partition consisting of only $s - 1$ sets. Now by the induction hypothesis, the set of $\mathbf{w}$ is finite, and therefore, S' is finite. This proves Lemma 2. ∎

We now proceed to estimate the cardinality of S. We need a few auxiliary results. For any subset A of $k[[z]]$, we denote by $\mathrm{rank}_k A$ the cardinality of a maximal k-linearly independent subset of A. For each subset I of $\{1, \ldots, h\}$ and each integer t with $1 \leq t \leq h - 1$, we define the set

$$V(I, t) = \{\mathbf{u} \in k^r : \mathrm{rank}_k \{a_i \mathbf{A}_i^{\mathbf{u}} : i \in I\} \leq t\}. \tag{2.10}$$

Clearly, $V(I, t) = k^r$ if $t \geq |I|$.

Lemma 3. *Let I, t be as above and assume that $t < |I|$. Then $V(I, t)$ is the set of common zeros in k^r of a system of polynomials in $k[X_1, \ldots, X_r]$, each of total degree at most $\binom{t+1}{2}$.*

Proof. The vector $\mathbf{u}$ belongs to $V(I, t)$ if and only if each $t + 1$-tuple among the functions $a_i \mathbf{A}_i^{\mathbf{u}}$ $(i \in I)$ is linearly dependent over k, that is, if and only if for each subset $J = \{i_0, \ldots, i_t\}$ of I of cardinality $t + 1$, the Wronskian determinant

$$\det\left(\left(\frac{d}{dz}\right)^i a_{i_j}\mathbf{A}_{i_j}^{\mathbf{u}}\right)_{i,j=0,\ldots,t}$$

is identically 0 as a function of z. By an argument completely similar to that in the proof of Proposition 1 of [5], one shows that the latter condition is equivalent to $\mathbf{u}$ being a common zero of some finite set of polynomials of degree $\leq \binom{t+1}{2}$. This proves Lemma 3. ∎

Lemma 4. *We have* $\mathbf{u} \in S$ *if and only if*

$$\text{(2.11)} \qquad \operatorname{rank}_k\{a_i\mathbf{A}_i^{\mathbf{u}} : i \in I\} + \operatorname{rank}_k\{a_i\mathbf{A}_i^{\mathbf{u}} : i \notin I\} > \operatorname{rank}_k\{a_i\mathbf{A}_i^{\mathbf{u}} : i = 1, \ldots, h\}$$

for each proper, *non-empty subset* I *of* $\{1, \ldots, h\}$.

Proof. First let $\mathbf{u} \in S$. Take a proper, non-empty subset I of $\{1, \ldots, h\}$. From (2.5), (2.6) it follows that there are $\xi_1, \ldots, \xi_h \in k$ such that

$$\sum_{i\in I} \xi_i a_i \mathbf{A}_i^{\mathbf{u}} = -\sum_{i\notin I} \xi_i a_i \mathbf{A}_i^{\mathbf{u}} \neq 0$$

and therefore the k-vector spaces spanned by $\{a_i\mathbf{A}_i^{\mathbf{u}} : i \in I\}$, $\{a_i\mathbf{A}_i^{\mathbf{u}} : i \notin I\}$, respectively, have non-trivial intersection. This implies (2.11).

Now let $\mathbf{u} \in k^r$ be such that (2.11) holds for every proper, non-empty subset I of $\{1, \ldots, h\}$. Let W be the vector space of $\xi = (\xi_1, \ldots, \xi_h)$ in k^h with $\sum_{i=1}^h \xi_i a_i \mathbf{A}_i^{\mathbf{u}} = 0$. Further, for a proper, non-empty subset I of $\{1, \ldots, h\}$, let $W(I)$ be the vector space of $\xi = (\xi_1, \ldots, \xi_h) \in k^h$ with $\sum_{i\in I} \xi_i a_i \mathbf{A}_i^{\mathbf{u}} = 0$ and $\sum_{i\notin I} \xi_i a_i \mathbf{A}_i^{\mathbf{u}} = 0$. Given a proper, non-empty subset I of $\{1, \ldots, h\}$, it follows from (2.11) that there are $\xi_1, \ldots, \xi_h \in k$ with $\sum_{i\in I} \xi_i a_i \mathbf{A}_i^{\mathbf{u}} = -\sum_{i\notin I} \xi_i a_i \mathbf{A}_i^{\mathbf{u}} \neq 0$; hence $W(I)$ is a proper linear subspace of W. It follows that there is $\xi \in W$ with $\xi \notin W(I)$ for each proper, non-empty subset I of $\{1, \ldots, h\}$. This means precisely that $\mathbf{u} \in S$. ∎

Proposition. *Assume* (2.4). *Then* $|S| \leq \sum_{p=2}^h \binom{p}{2}^r - h + 2$.

Proof. For $t = 1, \ldots, h - 1$, let $T_t = V(\{1, \ldots, h\}, t)$ (that is, the set of $\mathbf{u} \in k^r$ with $\operatorname{rank}_k\{a_i\mathbf{A}_i^{\mathbf{u}} : i = 1, \ldots, h\} \leq t$) and let S_t be the set of $\mathbf{u} \in S$ such that $\operatorname{rank}_k\{a_i\mathbf{A}_i^{\mathbf{u}} : i = 1, \ldots, h\} = t$. By (2.11), $\operatorname{rank}_k\{a_i\mathbf{A}_i^{\mathbf{u}} : i = 1, \ldots, h\} < h$, so $S = S_1 \cup \cdots \cup S_{h-1}$. We show by induction on $t = 1, \ldots, h-1$ that

$$\text{(2.12)} \qquad |S_1 \cup \cdots \cup S_t| \leq \sum_{p=1}^t \binom{p+1}{2}^r - t + 1.$$

Taking $t = h - 1$ proves the Proposition.

First let $t=1$. Let $\mathbf{u}_1,\mathbf{u}_2\in S_1$. Then $(a_ia_h^{-1})(\mathbf{A}_i\mathbf{A}_h^{-1})^{\mathbf{u}_j}\in k^*$ for $i=1,\dots,h$, $j=1,2$, which implies $(\mathbf{A}_i\mathbf{A}_h^{-1})^{\mathbf{u}_1-\mathbf{u}_2}\in k^*$ for $i=1,\dots,h$. But then $(\mathbf{A}_i\mathbf{A}_h^{-1})^{\mathbf{u}_1-\mathbf{u}_2}=1$ since $(\mathbf{A}_i\mathbf{A}_h^{-1})^{\mathbf{u}_1-\mathbf{u}_2}\in 1+zk[[z]]$ for $i=1,\dots,h$. Now assumption (2.4) gives $\mathbf{u}_1=\mathbf{u}_2$. So $|S_1|=1$, which implies (2.12) for $t=1$.

Now assume that $2\le t\le h-1$ and that (2.12) is true with t replaced by any number t' with $1\le t'<t$. By Lemma 3, T_t is an algebraic subvariety of k^r, being the set of common zeros of a system of polynomials of degree not exceeding $\binom{t+1}{2}$. By the last part of the proof of Proposition 1 of [5], T_t has at most $\binom{t+1}{2}^r$ irreducible components.

We first show that $T_t\setminus S_t$ is a finite union of proper algebraic subvarieties of T_t. Notice that $\mathbf{u}\in T_t\setminus S_t$ if and only if either $\mathrm{rank}_k\{a_i\mathbf{A}_i^{\mathbf{u}}:i=1,\dots,h\}\le t-1$ or (by Lemma 4) there are a proper, non-empty subset I of $\{1,\dots,h\}$ and an integer q with $1\le q\le t-1$ such that $\mathrm{rank}_k\{a_i\mathbf{A}_i^{\mathbf{u}}:i\in I\}\le q$ and $\mathrm{rank}_k\{a_i\mathbf{A}_i^{\mathbf{u}}:i\notin I\}\le t-q$. This means that $T_t\setminus S_t$ is equal to the union of T_{t-1} and of all sets $V(I,q)\cap V(\{1,\dots,h\}\setminus I,t-q)$ with I running through the proper, non-empty subsets of $\{1,\dots,h\}$ and q running through the integers with $1\le q\le t-1$. By Lemma 3 these sets are all subvarieties of T_t.

Now, by Lemma 2, S_t is finite, hence each element of S_t is an irreducible component (in fact an isolated point) of T_t. So $|S_t|\le\binom{t+1}{2}^r$. Now two cases may occur.

If $T_t=S_t$ then $S_{t'}=\emptyset$ for $t'=1,\dots,t-1$ and so $|S_1\cup\dots\cup S_t|=|S_t|\le\binom{t+1}{2}^r$. This certainly implies (2.12).

If S_t is strictly smaller than T_t then $T_t\setminus S_t$ has at least one irreducible component. But then $|S_t|\le\binom{t+1}{2}^r-1$. In conjunction with the induction hypothesis this gives

$$\begin{aligned}|S_1\cup\dots\cup S_t|&=|S_1\cup\dots\cup S_{t-1}|+|S_t|\\&\le\sum_{p=1}^{t-1}\binom{p+1}{2}^r-t+2+\binom{t+1}{2}^r-1,\end{aligned}$$

which again implies (2.12).

This completes the proof of our induction step, hence of our Proposition. ∎

3. Proof of the Theorem for transcendence degree 1. We prove the Theorem in the special case that K has transcendence degree 1 over k. For convenience we put $N:=\sum_{i=2}^{n+1}\binom{i}{2}^r-n+1$.

We start with some reductions. There are $\mathbf{a}_j=(\alpha_{1j},\dots,\alpha_{nj})\in\Gamma$ ($j=1,\dots,r$) such that for each $\mathbf{x}\in\Gamma$ there are integers $m,w_1,\dots,w_r$ with $m>0$, and $\xi=(\xi_1,\dots,\xi_n)\in(k^*)^n$ such that $\mathbf{x}^m=\xi\cdot\mathbf{a}_1^{w_1}\cdots\mathbf{a}_r^{w_r}$. Let L be the extension of k generated by $a_1,\dots,a_n$ and the α_{ij} ($i=1,\dots,n$,

$j = 1, \ldots, r$). Then L is the function field of a smooth projective algebraic curve C defined over k. Choose $z \in L$, $z \notin k$, such that the map $z : C \to \mathbb{P}_1(k) = k \cup \{\infty\}$ is unramified at 0 and such that none of the functions a_i, α_{ij} has a zero or a pole in any of the points from $z^{-1}(0)$. Thus, L can be embedded into $k((z))$, and the a_i and α_{ij} may be viewed as elements of $k[[z]]$ not divisible by z. By multiplying the α_{ij} with appropriate constants from k^*, which we are free to do, we may assume without loss of generality that the α_{ij} belong to $1 + zk[[z]]$.

Making the assumptions for the a_i and α_{ij} just mentioned, we can apply our Proposition. The functions α_{ij}^u ($u \in k$) are defined uniquely by means of (2.1). Therefore, we can express each $\mathbf{x} \in \Gamma$ as $\xi \cdot \mathbf{a}_1^{u_1} \cdots \mathbf{a}_r^{u_r}$ with $u_1, \ldots, u_r$ in $\mathbb{Q}$ and with $\xi = (\xi_1, \ldots, \xi_n) \in (k^*)^n$. Putting $\mathbf{A}_i := (\alpha_{i1}, \ldots, \alpha_{ir})$ ($i = 1, \ldots, n$), we can rewrite this as

$$\mathbf{x} = (\xi_1 \mathbf{A}_1^{\mathbf{u}}, \ldots, \xi_n \mathbf{A}_n^{\mathbf{u}}) \tag{3.1}$$

with $\xi_1, \ldots, \xi_n \in k^*$, $\mathbf{u} = (u_1, \ldots, u_r) \in \mathbb{Q}^r$. Putting in addition $h := n+1$, $\mathbf{A}_h := (1, \ldots, 1)$ (r times), $a_h := -1$, $\xi_h := 1$ we deduce that if $\mathbf{x} \in \Gamma$ is a non-degenerate solution of (1.1) then

$$\sum_{i=1}^{h} \xi_i a_i \mathbf{A}_i^{\mathbf{u}} = 0, \tag{3.2}$$

$$\sum_{i \in I} \xi_i a_i \mathbf{A}_i^{\mathbf{u}} \neq 0 \quad \text{for each proper, non-empty subset } I \text{ of } \{1, \ldots, h\}. \tag{3.3}$$

It remains to verify condition (2.4). According to an argument in the proof of Lemma 1 of [5], the set of $\mathbf{u} \in k^r$ such that $(\mathbf{A}_i \mathbf{A}_h^{-1})^{\mathbf{u}} = 1$ for $i = 1, \ldots, h$ is a linear subspace of k^r, say V, which is defined over $\mathbb{Q}$. Now if $\mathbf{u} = (u_1, \ldots, u_r) \in V \cap \mathbb{Q}^r$, then $\mathbf{A}_i^{\mathbf{u}} = 1$ for $i = 1, \ldots, n$ since $\mathbf{A}_h = (1, \ldots, 1)$, and therefore $\mathbf{a}_1^{u_1} \cdots \mathbf{a}_r^{u_r} = (1, \ldots, 1)$. This implies $\mathbf{u} = \mathbf{0}$, since otherwise $\operatorname{rank}(\Gamma/(k^*)^n)$ would be smaller than r. Hence $V \cap \mathbb{Q}^r = \{\mathbf{0}\}$, and therefore $V = \{\mathbf{0}\}$ since V is defined over $\mathbb{Q}$. This implies (2.4).

As observed above, if $\mathbf{x} \in \Gamma$ is a non-degenerate solution of (1.1), then $\mathbf{u}$ satisfies (3.2), (3.3), which means that $\mathbf{u}$ belongs to the set S given by (2.5), (2.6). So by the Proposition, we have at most N possibilities for $\mathbf{u}$. Then according to (3.1), the non-degenerate solutions $\mathbf{x}$ of (1.1) lie in at most N $(k^*)^n$-cosets. This completes the proof of our Theorem in the special case where K has transcendence degree 1 over k. ∎

4. Proof of the Theorem in the general case. We now prove our Theorem in the general case, i.e., when the field K is an arbitrary transcendental extension of k. As before, we define $N := \sum_{i=2}^{n+1} \binom{i}{2}^r - n + 1$.

There is of course no loss of generality in assuming that K is generated over k by the coefficients $a_1, \ldots, a_n$ and the coordinates of all elements of Γ.

Since Γ is assumed to have rank r, there are $\mathbf{a}_1, \ldots, \mathbf{a}_r \in \Gamma$ such that for every $\mathbf{x} \in \Gamma$ there are integers $m, z_1, \ldots, z_r$ with $m > 0$ and $\xi \in (k^*)^n$ such that $\mathbf{x}^m = \xi \cdot \mathbf{a}_1^{z_1} \cdots \mathbf{a}_r^{z_r}$. Hence K is algebraic over the extension of k generated by $a_1, \ldots, a_n$ and the coordinates of $\mathbf{a}_1, \ldots, \mathbf{a}_r$. Therefore, K has finite transcendence degree over k. We will prove by induction on $d := \mathrm{trdeg}(K/k)$ that for any group Γ with $\mathrm{rank}(\Gamma/(k^*)^n) \le r$, the non-degenerate solutions $\mathbf{x} \in \Gamma$ of (1.3) lie in not more than N $(k^*)^n$-cosets. The case $d = 0$ is trivial since in that case $\Gamma = (k^*)^n$ and all solutions lie in a single $(k^*)^n$-coset. Further, the case $d = 1$ has been taken care of in the previous section. So we assume $d > 1$ and that the above assertion is true up to $d - 1$.

We suppose by contradiction that (1.1) has at least $N + 1$ non-degenerate solutions, denoted $\mathbf{x}_1, \ldots, \mathbf{x}_{N+1} \in \Gamma$, falling into pairwise distinct $(k^*)^n$-cosets. For each such solution $\mathbf{x}_j =: (x_{1j}, \ldots, x_{nj})$ and for each non-empty subset I of $\{1, \ldots, n\}$ let us consider the corresponding subsum $\sum_{i \in I} a_i x_{ij}$, which we denote by $\sigma_{(j,I)}$. In this way we obtain finitely many elements $\sigma_{(j,I)} \in K$, none of which vanishes, since the solutions are non-degenerate.

Further, let $\mathbf{x}_u, \mathbf{x}_v$ be distinct solutions, with $1 \le u \ne v \le N + 1$. Since the solutions lie in distinct $(k^*)^n$-cosets, for some $i \in \{1, \ldots, n\}$ the ratio x_{iu}/x_{iv} does not lie in k. For each pair (u, v) as above let us pick one such index $i = i(u, v)$ and let us put $\tau_{(u,v)} := x_{iu}/x_{iv} \in K^* \setminus k^*$.

We are going to "specialize" such elements of K, getting corresponding elements of a field with smaller transcendence degree and obtaining eventually a contradiction. We shall formulate the specialization argument in geometric terms.

Let $\widetilde{K}$ be the extension of k generated by $a_1, \ldots, a_n$ and by the coordinates of $\mathbf{x}_1, \ldots, \mathbf{x}_{N+1}$. Thus $\widetilde{K}$ is finitely generated over k. Further, let $\widetilde{\Gamma}$ be the group containing $(k^*)^n$ and generated over it by $\mathbf{x}_1, \ldots, \mathbf{x}_{N+1}$. Then $\widetilde{\Gamma}$ is a subgroup of $\Gamma \cap (\widetilde{K}^*)^n$, and so $\mathrm{rank}(\widetilde{\Gamma}) \le r$. Now (1.1) has at least $N+1$ non-degenerate solutions in $\widetilde{\Gamma}$ lying in different $(k^*)^n$-cosets. By the induction hypothesis this is impossible if $\mathrm{trdeg}(\widetilde{K}/k) < d$. So $\mathrm{trdeg}(\widetilde{K}/k) = d$.

The finitely generated extension $\widetilde{K}/k$ may be viewed as the function field of an irreducible affine algebraic variety V over k, with $d = \dim V$. Then each element of $\widetilde{K}$ represents a rational function on V. Let us consider irreducible closed subvarieties W of V, with function field denoted $L := k(W)$, with the following properties:

(A) $\dim W = d - 1$.
(B) There exists a point $P \in W(k)$ such that each of the (finitely many) elements a_i, x_{ij} and $\sigma_{(j,I)}$, $\tau_{(u,v)}$ constructed above is defined and non-zero at P; so the elements induce by restriction non-zero rational functions a'_i, x'_{ij}, $\sigma'_{(j,I)}$ and $\tau'_{(u,v)}$ in $L^* = k(W)^*$.

(C) None of the elements $\tau'_{(u,v)}$ lies in k^*.

We shall construct W as an irreducible component of a suitable hyperplane section of V.

To start with, (A) follows from the well-known fact that any irreducible component W of any hyperplane section of V has dimension $d-1$.

Let us analyze (B). Each of the elements of $\widetilde{K}^*$ mentioned in (B) may be expressed as a ratio of non-zero polynomials in the affine coordinates of V; since these elements are defined and non-zero by assumption, none of these polynomials vanishes identically on V, so each such polynomial defines in V a proper (possibly reducible) closed subvariety. Take now a point $P \in V(k)$ outside the union of these finitely many proper subvarieties. For (B) to be satisfied it then plainly suffices that W contains P.

Finally, let us look at (C). For each $u, v \in \{1, \ldots, n\}$, $u \neq v$, let $Z(u,v)$ be the variety defined in V by the equation $\tau_{(u,v)} = \tau_{(u,v)}(P)$. Since $\tau_{(u,v)}$ is not constant on V, each component of $Z(u,v)$ is a subvariety of V of dimension $d-1$. Choose now W as an irreducible component through P of the intersection of V with a hyperplane π going through P, such that W is not contained in any of the finitely many $Z(u,v)$. It suffices e.g. that the hyperplane π does not contain any irreducible component of any $Z(u,v)$, and there are plenty of choices for that. (For example, for each of the relevant finitely many varieties, each of dimension $d-1 \geq 1$, take a point $Q \neq P$ in it and let π be a hyperplane through P and not containing any of the Q's. Note that here we use the fact that $d \geq 2$.) Since $P \in W(k)$ and $\tau_{(u,v)}$ is not constantly equal to $\tau_{(u,v)}(P)$ on all of W by construction, the restriction $\tau'_{(u,v)}$ is not constant, as required.

Consider now the elements $\mathbf{x}'_j := (x'_{1j}, \ldots, x'_{nj}) \in L^n$, $j = 1, \ldots, N+1$, where the prime denotes, as before, the restriction to W (which by (B) is well-defined for all the functions in question). Notice that the restriction to W is a homomorphism from the local ring of V at P to the local ring of W at P which is contained in L. This homomorphism maps $\widetilde{\Gamma}$ to the group Γ' containing $(k^*)^n$, generated over it by the elements $\mathbf{x}'_1, \ldots, \mathbf{x}'_{N+1}$. Thus, $a'_1, \ldots, a'_n$ and the coordinates of the elements from Γ' lie in L. Further, Γ' is a homomorphic image of $\widetilde{\Gamma}$ which was in turn a subgroup of Γ; therefore $\operatorname{rank}(\Gamma'/(k^*)^n) \leq r$. Since the $\mathbf{x}_j$ are solutions of (1.1) in $\widetilde{\Gamma}$, the elements $\mathbf{x}'_j$ are solutions of $a'_1 x_1 + \cdots + a'_n x_n = 1$ in Γ'. Again by (B), we see that none of the (non-empty) subsums $\sigma'_{(j,I)} = \sum_{i \in I} a'_i x'_{ij}$ vanishes, so these solutions are non-degenerate. Finally, by (C), no two solutions $\mathbf{x}'_u, \mathbf{x}'_v$, $1 \leq u \neq v \leq N+1$, lie in a same $(k^*)^n$-coset of $(L^*)^n$. Since by (A) the field L has transcendence degree $d-1$ over k, this contradicts the inductive assumption, concluding the induction step and the proof. ∎

5. Proof of the Corollary. We keep the notation and assumptions from Section 1. We consider the non-degenerate solutions $(x_1, \dots, x_n) \in \Gamma$ of (1.1) such that

$$(a_1x_1, \dots, a_nx_n) \notin (k^*)^n. \tag{5.1}$$

We first show that each $(k^*)^n$-coset of such solutions is contained in a proper linear subspace of K^n. Fix a non-degenerate solution $\mathbf{x} = (x_1, \dots, x_n)$ of (1.1) with (5.1). Any other solution of (1.1) in the same $(k^*)^n$-coset as $\mathbf{x}$ can be expressed as $\mathbf{x} \cdot \xi = (x_1\xi_1, \dots, x_n\xi_n)$ with $\xi = (\xi_1, \dots, \xi_n) \in (k^*)^n$ and $a_1x_1\xi_1 + \dots + a_nx_n\xi_n = 1$. Now the points $\xi \in k^n$ satisfying the latter equation lie in a proper linear subspace of k^n, since otherwise $(a_1x_1, \dots, a_nx_n)$ would be the unique solution of a system of n linearly independent linear equations with coefficients from k, hence $a_1x_1, \dots, a_nx_n \in k$, violating (5.1). But this implies that indeed the $(k^*)^n$-coset $\{\mathbf{x} \cdot \xi : \xi \in (k^*)^n\}$ is contained in a proper linear subspace of K^n.

Now our Theorem implies that the non-degenerate solutions of (1.1) with (5.1) lie in at most $\sum_{i=2}^{n+1} \binom{i}{2}^r - n + 1$ proper linear subspaces of K^n. Further, the degenerate solutions of (1.1) lie in at most $2^n - n - 2$ proper linear subspaces of K^n, each given by $\sum_{i \in I} a_i x_i = 0$, where I is a subset of $\{1, \dots, n\}$ of cardinality $\neq 0, 1, n$. By adding these two bounds our Corollary follows. ■

References

[1] E. Bombieri, J. Mueller and U. Zannier, *Equations in one variable over function fields*, Acta Arith. 99 (2001), 27–39.

[2] J.-H. Evertse and K. Győry, *On the numbers of solutions of weighted unit equations*, Compos. Math. 66 (1988), 329–354.

[3] J.-H. Evertse, H. P. Schlickewei and W. M. Schmidt, *Linear equations in variables which lie in a multiplicative group*, Ann. of Math. 155 (2002), 807–836.

[4] J. Mueller, *S-unit equations in function fields via the abc-theorem*, Bull. London Math. Soc. 32 (2000), 163–170.

[5] U. Zannier, *On the integer solutions of exponential equations in function fields*, Ann. Inst. Fourier (Grenoble) 54 (2004), 849–874.

Mathematisch Instituut
Universiteit Leiden
Postbus 9512
2300 RA Leiden, The Netherlands
E-mail: evertse@math.leidenuniv.nl

Scuola Normale Superiore
Piazza dei Cavalieri 7
56126 Pisa, Italy
E-mail: u.zannier@sns.it

Received on 4.10.2007 (5546)