JOURNAL OF NUMBER THEORY 39, 339–350 (1991)

The Generalized Fermat Equation in Function Fields

E. BOMBIERI AND J. MUELLER*

School of Mathematics, Institute for Advanced Study, Princeton, New Jersey 08540; and Department of Mathematics, Fordham University, Bronx, New York 10458

Communicated by Hans Zassenhaus

Received January 19, 1990; revised June 7, 1990

We show that if $r > n!(n!-2)$ the set of solutions $x_j \in \mathbf{C}(t)$ of a Fermat equation $\sum_1^n a_j x_j^r = 0$, $a_j \in \mathbf{C}(t)$, is the union of at most $n!^{n!}$ families with an explicitly given simple structure. In particular, the number of projective solutions, up to rth roots of unity, of such an equation is either at most $n!^{n!}$ or infinite. The proof uses the function field version of the abc-conjecture due to Mason, Voloch, and Brownawell and Masser.

I. INTRODUCTION

The purpose of this paper is to study the structure of the set of solutions of the generalized Fermat equation

$$a_1 z_1^r + a_2 z_2^r + \cdots + a_n z_n^r = 0$$

over a field $K = k(t)$ of rational functions of one variable, with constant field k algebraically closed of characteristic 0. In practice, it is more convenient to deal with the linear equation

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0$$

with coefficients $a_i \in K$, not all zero, and ask for solutions $\mathbf{x}$ such that every coordinate x_i is an rth power.

Questions of this type have been studied by several authors. Bounds for the height and number of solutions as a function of the degree r and of the heights of the coefficients a_i of the basic Fermat equation can be found, for instance, in papers by Silverman [Sil] and Voloch [V]. Both authors rely on the so-called abc-conjecture of Masser and Oesterlé in function fields,

* Supported in part by NSF Grant DMS-8808398.

641/39/3-7

0022-314X/91 $3.00

proved by Mason [Ma], Voloch [V], and Brownawell and Masser [Br–M]; some of these ideas also appear in the book by Shafarevich [Sh, pp. 7–8], in the analysis of the classical Fermat equation in the function field case.

More recently, one of the authors of this paper [M] obtained results independent of the coefficients a_i in the case $n=3$, by noting that if there are sufficiently many solutions then one can eliminate the coefficients. The resulting vanishing determinant in the solutions is a Fermat equation with coefficients ± 1, in a higher number of variables, and the abc-inequality can be used again. This idea of eliminating the coefficients appears explicitly in a short paper by Chowla [Ch] in 1964, although its origin may be earlier and possibly goes back to Siegel.

As a consequence, Mueller proved that a "non-degenerate" equation $ax+by=c$ has at most two non-zero solutions $x, y \in K^r$ if $r>30$, which is clearly best possible except for the fact that the range for r can probably be diminished somewhat. She also deals with the case in which the function field K is a function field of one variable of positive genus, with rather similar results.

It is clear that in the study of the general Fermat equation one must allow infinite families of solutions. In fact the equation may split into smaller Fermat equations and moreover a Fermat equation, after a change of variable, may become equivalent to the equation $z_1^r + \cdots + z_n^r = 0$, which we can solve easily with $z_i \in k$, since k is algebraically closed.

Our main result is that if $r > r(n)$ then solutions fall into finitely many families with the simple structure described before and moreover there is a bound on the number of such families, as a function of n only.

THEOREM. *Suppose that $r > n!(n!-2)$. Then the solutions of*

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0 \tag{1}$$

with all x_i's rth *powers in K forming a finite number of families. Each family of solutions determines a decomposition*

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = \sum_{\mathscr{R}} \left(\sum_{j \in \mathscr{R}} a_jx_j \right),$$

where

$$\sum_{j \in \mathscr{R}} a_jx_j = 0$$

for every $\mathscr{R}$. The set of solutions of the latter equation belonging to the given family consists of all points

$$x_j = \eta_j^r S^r v_j,$$

where the η_j's are suitable fixed elements of K, S is an arbitrary element of K, and the vector $\mathbf{v}_{\mathscr{R}} = \{v_j \mid j \in \mathscr{R}\}$ runs over all elements of some vector subspace of the euclidean space $\mathbf{E}^{\mathrm{Card}(\mathscr{R})}(k)$.

Moreover, the number of such families is at most $n!^{n!}$.

COROLLARY. *Suppose that $r > n!(n!-2)$. Then the number of equivalence classes, up to rth roots of unity, of K-rational points of the projective Fermat variety*

$$a_1 z_1^r + \cdots + a_n z_n^r = 0$$

is either at most $n!^{n!}$ or infinite.

Thus our result may be considered an extension of Mueller's theorem to a Fermat equation is several variables. There is no question that our theorem admits an extension to the case of a function of positive genus, and it is very likely that it may be generalized to deal with a large class of "fewnomial" equations. It is also clear that the method of proof will extend, with minor changes, to the case of number fields as soon as the appropriate form of the *abc*-conjecture is available.

We leave these generalizations aside for further study and limit ourselves here, mainly for the sake of simplicity, to the case of the generalized Fermat equation in dimension n over the function field $K = k(t)$ with k algebraically closed of characteristic 0.

The authors thank A. Granville for pointing out an improvement of our original lower bound for r and W. M. Schmidt for some important suggestions.

II. The Transformed Equation

Let us consider the basic equation

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0 \tag{1}$$

to be solved for $x_i \in K^r$ and its set $\mathscr{X}$ of solutions $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots$. We may also consider (1) as a set of linear equations for the quantities $a_1, a_2, \ldots$, one for each solution $\mathbf{x}^{(i)}$. Since not all a_i are zero, we must have

$$\mathrm{rank}(\mathbf{x}^{(i)})_{i=1,2,\ldots} = m < n.$$

By reordering the solutions, we may also assume that

$$\mathrm{rank}(\mathbf{x}^{(i)})_{i=1,2,\ldots,m} = m < n$$

and thus

$$\operatorname{rank}(\mathbf{x}^{(i)})_{i=1,2,\dots,m,\nu} = m < n \tag{2}$$

for every ν.

Let J be any subset of $\{1, \dots, n\}$ of cardinality $|J| = m+1$. Then Eq. (2) yields

$$\det(\mathbf{x}_J^{(i)})_{i=1,2,\dots,m,\nu} = 0,$$

where, for a vector $\mathbf{x}$, we write $\mathbf{x}_J = \{x_j \mid j \in J\}$. If we expand the determinant in full we obtain

$$\sum_\sigma \varepsilon(\sigma)\, x_{\sigma_1}^{(1)} \cdots x_{\sigma_m}^{(m)} x_{\sigma_{m+1}}^{(\nu)} = 0, \tag{3}$$

where σ runs over the permutations of J and where $\varepsilon(\sigma)$ is ± 1 according to the parity of σ. We abbreviate

$$m_\sigma(\mathbf{x}) = \varepsilon(\sigma)\, x_{\sigma_1}^{(1)} \cdots x_{\sigma_m}^{(m)} x_{\sigma_{m+1}};$$

denote by L_J the linear forms in $x_1, \dots, x_n$,

$$L_J(x) = \sum_\sigma m_\sigma(\mathbf{x});$$

and denote by $\mathcal{M} = \mathcal{M}(J)$ the set of permutations σ of J.

Some of the linear forms L_J, but not all, may be identically 0.

Our first and easy remark is that the system $L_J(\mathbf{x}) = 0$ for varying J but fixed $x^{(i)}$ is equivalent to Eq. (1).

LEMMA 1. *A vector* $\mathbf{x}$ *with* $x_i \in K^r$ *satisfies the Fermat equation*

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0$$

if and only if it satisfies the system of equations

$$L_J(\mathbf{x}) = 0$$

for every subset $J \in \{1, \dots, n\}$ *of cardinality* $m+1$.

Proof. We have already shown that the solutions of the generalized Fermat equation (1) satisfy

$$L_J(\mathbf{x}) = 0$$

for every subset $J \in \{1, \dots, n\}$ of cardinality $m+1$.

In order to prove the implication in the other direction, we proceed as follows. Let V be the K-vector space of solutions of Eq. (1) but without the condition $x_i \in K^r$. Then V has dimension $n-1$ with a basis $\mathbf{x}^{(1)}, ..., \mathbf{x}^{(m)}, \mathbf{z}^{(1)}, ..., \mathbf{z}^{(n-1-m)}$ for suitable vectors $\mathbf{z}^{(h)}$. Now the a_i's are Grassmann coordinates for V and therefore they are proportional to the cofactors of the matrix of vectors of the basis. This means that there is $\alpha \in K$ such that

$$a_1 x_1 + a_2 x_2 + a_n x_n = \alpha \det \begin{pmatrix} \mathbf{x}^{(i)} \\ \mathbf{z}^{(h)} \\ \mathbf{x} \end{pmatrix}_{i=1,...,m;\, h=1,...,n-m-1}$$

We expand the right-hand side of this equation according to the Laplace expansion for the block of rows containing the auxiliary vectors $\mathbf{z}^{(h)}$, thus expressing $a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$ as a linear combination of the $L_J(\mathbf{x})$; this completes the proof of Lemma 1.

III. Application of the *abc*-Inequality

We are interested in which subsums in the left-hand side of (3) may vanish. Let

$$\sum \left(\sum_{\mathcal{N}} m_\sigma(\mathbf{x}^{(\nu)}) \right) = 0$$

be a decomposition of (3) into vanishing subsums. This decomposition always exists, though it need not be unique. We may and do assume that every component

$$\sum_{\mathcal{N}} m_\sigma(\mathbf{x}^{(\nu)})$$

of the above sum has the property that no proper subsum of it vanishes.

For every decomposition as above there is a corresponding partition $\pi = \bigcup \mathcal{N}$ of $\mathcal{M}$ into subsets, and we group together solutions which give rise to the same partition. We also write $\mathcal{M}(J)$ and $\pi(J)$ if we want to emphasize the fact that these objects are associated with the set J.

Definition. The set $\mathcal{X}(\pi)$ is the set of solutions $\mathbf{x} \in (K^r)^n$ such that

$$\sum_{\mathcal{N}} m_\sigma(\mathbf{x}) = 0 \tag{4}$$

for every component $\mathcal{N}$ of the partition π and such that no proper of the above sum vanishes.

Now we can state

LEMMA 2. *Let J and π be as above. If $r > n!(n!-2)$ then for every solution $\mathbf{x} \in \mathcal{X}(\pi(J))$ and every $\sigma \in \mathcal{N}$ we have*

$$m_\sigma(\mathbf{x}) = \varepsilon(\sigma)\, x^{(1)}_{\sigma_1} \cdots x^{(m)}_{\sigma_m} x_{\sigma_{m+1}} = u_\sigma S^r, \tag{5}$$

where $u_\sigma \in k$ depends on σ, $\mathbf{x}$, $\mathcal{N}$, while $S \in K$ depends only on $\mathbf{x}$ and $\mathcal{N}$ but not on σ. Moreover we have

$$\sum_{\mathcal{N}} u_\sigma = 0 \tag{6}$$

and no proper subsum of the u_σ's vanishes.

Conversely, suppose we are given the solutions $\mathbf{x}^{(i)}$, $i = 1, ..., m$, and, for each J, let us choose a partition π of the set $\mathcal{M}$ of the permutations of J into subsets $\mathcal{N}$. Suppose that for each J, π, $\mathcal{N}$ and $\sigma \in \mathcal{N}$ the point $\mathbf{x}$ satisfies (5) *and* (6) *for some $u_\sigma \in k$ and some $S \in K$, independent of σ. Then $\mathbf{x} \in \mathcal{X}$.*

Proof. Let $\mathbf{x} \in \mathcal{X}(\pi)$. We analyze (4) for each $\mathcal{N}$; we need only consider the case in which $N = \operatorname{card}(\mathcal{N}) > 1$. Each monomial $m_\sigma(\mathbf{x})$ is a rational function of t and we can write

$$m_\sigma(\mathbf{x}) = P_\sigma R$$

for $\sigma \in \mathcal{N}$, where R is a non-zero rational function and where the P_σ's are polynomials without common factor. Thus we have

$$\sum_{\sigma \in \mathcal{N}} P_\sigma = 0, \tag{7}$$

the elements of (7) are polynomials without common factor, and no subsum of (7) vanishes.

Let $D_\sigma = \deg P_\sigma$ and $D = \max D_\sigma$ and let us choose σ_0 such that $D_{\sigma_0} = D$. We define $f_\sigma = P_\sigma / P_{\sigma_0}$ and apply Theorem B of [Br–M], i.e., the *abc*-inequality in function fields, to the sum $\sum f_\sigma$. We deduce

$$D \leqq \sum_v (\gamma_N - \gamma_{m_v}),$$

where $\gamma_p = (p-1)(p-2)/2$ and where, for each place v of K/k,

$$m_v = \#\{\sigma \mid f_\sigma \text{ is a unit at } v\}.$$

We write $m_v = N - p_v$ and obtain a fortiori

$$D \leqq \sum_v ((2N-3)\, p_v - p_v^2)/2 \leqq (N-2) \sum_v p_v.$$

Suppose first that $v \neq \infty$. Then f_σ is a unit at v unless either P_σ or P_{σ_0} vanishes at v. Thus

$$\sum_{v \neq \infty} p_v \leqslant \sum_\sigma \# \{\text{roots of } P_\sigma \text{ counted without multiplicity}\}.$$

For the place at $v = \infty$ we note that f_σ is a unit unless $D_\sigma < D$. Thus

$$p_\infty \leqq \# \{\sigma \mid D_\sigma < D\}.$$

On the other hand, the coordinates of each element of $\mathscr{X}$ are rth powers; therefore so are the monomials m_σ and the polynomials P_σ and the common factor R. In particular $R = S^r$ for some $S \in K$.

Since every root of P_σ has multiplicity at least r, we have

$$\sum_\sigma \# \{\text{roots of } P_\sigma \text{ counted without multiplicity}\} + \# \{\sigma \mid D_\sigma < D\} \leqq ND/r.$$

We combine this inequality with the two preceding inequalities and the bound on D and obtain

$$\max \deg P_\sigma \leqq \frac{N(N-2)}{r} \max \deg P_\sigma.$$

Also $N \leqslant (m+1)! \leqq n!$. Now we have two possibilities. If $\max \deg P_\sigma > 0$ we deduce $r \leqslant N(N-2)$, contrary to our hypothesis bounding r from below. Therefore $\max \deg P_\sigma = 0$, and we let each u_σ be the non-zero constant P_σ, which proves (5). It is also clear that (6) is equivalent to (4).

It remains to prove the last part of Lemma 2. The hypothesis $\sum_{\mathscr{N}} u_\sigma = 0$ implies $\sum_{\mathscr{N}} m_\sigma(\mathbf{x}) = 0$ for every component $\mathscr{N}$ of the partition of $\mathscr{M}$, and therefore we also have $L_J(\mathbf{x}) = 0$. Since this holds for every J, the result follows from Lemma 1.

IV. Parametrization of Solutions and Conclusion of Proof

In this section we simplify the parametrization of solutions obtained in Lemma 2 on the assumption that $r > n!(n!-2)$.

According to Lemma 2, the solutions of (1) are obtained as follows. We start by choosing once and for all a basic set $\mathbf{x}^{(1)}, ..., \mathbf{x}^{(m)}$ of independent solutions. Then we choose, for every subset $J \in \{1, ..., n\}$ of cardinality

$m+1$, a partition $\pi=\pi(J)$ of the set $\mathscr{M}$ of permutations of J. Now suppose that for every component $\mathscr{N}$ of the partition we have $S\in K$ and $u_\sigma\in k$ such that

$$m_\sigma(\mathbf{x})=\varepsilon(\sigma)\,x_{\sigma_1}^{(1)}\cdots x_{\sigma_m}^{(m)}x_{\sigma_{m+1}}=u_\sigma S^r, \tag{5}$$

and moreover

$$\sum_{\mathscr{N}} u_\sigma=0, \tag{6}$$

where also no proper subsum of (6) vanishes, and in particular, $u_\sigma\neq 0$ if $\mathrm{Card}(\mathscr{N})>1$. Then $\mathbf{x}\in\mathscr{X}$ and every solution of (1) with all coordinates rth powers arises in this fashion for some collection of partitions $\{J,\pi(J)\}$.

The trouble with this parametrization is that different σ's, perhaps associated to different components $\mathscr{N}$ and different J's, may involve the same $x_{\sigma_{m+1}}$, and our final step consists in analyzing how this can happen.

We begin with the case in which there are two or more σ's with the same last coordinate σ_{m+1} and belonging to the same component $\mathscr{N}$. Let σ and τ be two such elements and let us write for simplicity $j=\sigma_{m+1}=\tau_{m+1}$. Then (5) yields

$$\varepsilon(\sigma)\,x_{\sigma_1}^{(1)}\cdots x_{\sigma_m}^{(m)}x_j/u_\sigma=\varepsilon(\tau)\,x_{\tau_1}^{(1)}\cdots x_{\tau_m}^{(m)}x_j/u_\tau,$$

and thus, after division by x_j,

$$\frac{u_\tau}{u_\sigma}=\frac{\varepsilon(\tau)\,x_{\tau_1}^{(1)}\cdots x_{\tau_m}^{(m)}}{\varepsilon(\sigma)\,x_{\sigma_1}^{(1)}\cdots x_{\sigma_m}^{(m)}}.$$

The right-hand side of this equation is independent of $\mathbf{x}$, while the left-hand side is an element of k. Therefore their common value must be a constant in k and we can write

$$u_\tau=\gamma_{\sigma,\tau}u_\sigma \tag{8}$$

for a constant $\gamma_{\sigma,\tau}\in k$, independent of $\mathbf{x}$.

Let $p(\mathscr{N})$ be the image of the projection of $\mathscr{N}$ into J obtained by means of the last coordinate, i.e., $\sigma\mapsto\sigma_{m+1}$. We have shown:

LEMMA 3. *The set of Eqs.* (5) *and* (6) *for* $\sigma\in\mathscr{N}$ *is equivalent to*

$$x_j=\eta_j^r S^r v_j \tag{9}$$

for $j\in p(\mathscr{N})$, *where the* η_j's *are independent of* $\mathbf{x}$ *and where the* v_j's *are elements of k satisfying a suitable linear condition*

$$\sum_{p(\mathscr{N})} c_j v_j=0. \tag{10}$$

Proof. If $\operatorname{card}(\mathcal{N})=1$, the result follows by taking $c_j=1$ if $x_{\sigma_1}^{(1)}\cdots x_{\sigma_m}^{(m)}\neq 0$, $c_j=0$ otherwise, and $\eta_j=1$. If instead $\operatorname{card}(\mathcal{N})>1$, it suffices to select one representative σ for each $j\in p(\mathcal{N})$ and write $v_j=u_\sigma$, $\eta_j^r=(\varepsilon(\sigma)\,x_{\sigma_1}^{(1)}\cdots x_{\sigma_m}^{(m)})^{-1}$. Now (8) shows that the linear condition (6) is equivalent to (10) for some constants $c_j\in k$, which concludes the proof.

We deal with the case of different components $\mathcal{N}$, $\mathcal{N}'$ with $p(\mathcal{N})\cap p(\mathcal{N}')\neq\varnothing$ in a similar fashion.

Let us say that $p(\mathcal{N})$ and $p(\mathcal{N}')$ are connected if they are not disjoint. Then the collection $\{J,\pi(J)\}$ determines a partition $\rho=\bigcup\mathcal{R}$ of $\{1,\ldots,n\}$ into maximal connected components of $\bigcup\bigcup_{\pi(J)}p(\mathcal{N})$.

LEMMA 4. *Let $\{J,\pi(J)\}$ and $\rho=\bigcup\mathcal{R}$ be as before. Then for every $\mathcal{R}$ we can find elements $\eta_j\in K$, $j\in\mathcal{R}$ and a system of linear equations*

$$\sum_{j\in\mathcal{R}} c_{hj}v_j=0 \tag{11}$$

with $h=1,\ldots,h(\mathcal{R})$ and $c_{hj}\in k$, such that every $\mathbf{x}\in\bigcap\mathcal{X}(\pi(J))$ is written as

$$x_j=\eta_j^r S^r v_j,\qquad j\in\mathcal{R}, \tag{12}$$

for suitable $S\in K$ and $v_j\in k$ satisfying (11).

Conversely, if $\mathbf{x}$ satisfies (11) and (12) with a vector $\mathbf{v}_{\mathcal{R}}=\{v_j\mid j\in\mathcal{R}\}$, $v_j\in k$, and an element $S_{\mathcal{R}}\in K$, for all components $\mathcal{R}$ of the partition ρ, then we have $\mathbf{x}\in\mathcal{X}$.

Proof. Let $\mathcal{R}$ be a minimal connected component of the partition ρ and let $\mathrm{R}=\bigcup p(\mathcal{N}_\alpha)$.

We order the sets $p(\mathcal{N}_\alpha)$ so that each set $p(\mathcal{N}_\alpha)$ is connected with at least one set preceding it and denote by $\mathcal{R}_h$ the union of the first h elements of this ordering.

By Lemma 3, on each $p(\mathcal{N}_\alpha)$ we have

$$x_j=\eta_{j\alpha}^r S_\alpha^r v_{j\alpha},\qquad j\in p(\mathcal{N}_\alpha),$$

with $\eta_{j\alpha}$ independent of $\mathbf{x}$ and with the $v_{j\alpha}$ satisfying a linear relation

$$\sum_{p(\mathcal{N}_\alpha)} c_j(\mathcal{N}_\alpha)\,v_{j\alpha}=0,$$

for suitable constants $c_j(\mathcal{N}_\alpha)\in k$.

We claim that the set of such equations corresponding to the first h elements of the ordering is equivalent to a system of relations

$$x_j=\eta_j^r S^r v_j,\qquad j\in\mathcal{R}_h,$$

with η_j independent of $\mathbf{x}$ and with the $v_j \in k$ satisfying h linear relations

$$\sum c_{ij} v_j = 0$$

for $i = 1, ..., h$. The statement of Lemma 4 then is the statement corresponding to the last $\mathscr{R}_h$ in the construction.

We prove our claim by induction on h. If $h = 1$, this is the conclusion of Lemma 3. Now suppose that $h > 1$ and that the statement is true for $h - 1$, and hence

$$x_j = \eta_j^r S^r v_j, \qquad j \in \mathscr{R}_{h-1}, \tag{13}$$

and

$$\sum c_{ij} v_j = 0$$

for $i = 1, ..., h - 1$. To obtain the new system for $\mathscr{R}_h$ we must add the new relations

$$x_j = \eta_{j\alpha}^r S_\alpha^r v_{j\alpha}, \qquad j \in p(\mathscr{N}_\alpha), \tag{14}$$

and

$$\sum_{p(\mathscr{N}_\alpha)} c_j(\mathscr{N}_\alpha)\, v_{j\alpha} = 0 \tag{15}$$

associated to the hth element $p(\mathscr{N}_\alpha)$ of the ordering.

By assymption, there is $l \in \mathscr{R}_{h-1} \cap p(\mathscr{N}_\alpha)$, and we choose l once and for all. Consider first the case in which $\operatorname{card}(\mathscr{N}_\alpha) = 1$. Then Lemma 3 shows that either $c_j(\mathscr{N}_\alpha) = 0$, in which case (14) simply expresses the fact that x_j is an rth power, or $c_j(\mathscr{N}_\alpha) \neq 0$, in which case $x_j = 0$ and therefore (15) can be expressed by the single equation $v_j = 0$. This shows that we need consider only the case in which $\operatorname{card}(\mathscr{N}_\alpha) > 1$.

In this case, (13) and (14) yield

$$\eta_l^r S^r v_l = \eta_{l\alpha}^r S_\alpha^r v_{l\alpha}$$

and therefore

$$S_\alpha^r = \left(\frac{\eta_l}{\eta_{l\alpha}}\right)^r S^r \frac{v_l}{v_{l\alpha}}.$$

We substitute in (14), which then becomes

$$x_j = \left(\frac{\eta_l}{\eta_{l\alpha}}\eta_{j\alpha}\right)^r S^r \frac{v_l}{v_{l\alpha}} v_{j\alpha}, \qquad j \in p(\mathscr{N}_\alpha). \tag{16}$$

Suppose first that $j \in p(\mathcal{N}_\alpha)$ but not $j \notin \mathcal{R}_{h-1}$. Now we simply define η_j and v_j by means of

$$\eta_{j\alpha} = \frac{\eta_{l\alpha}}{\eta_l} \eta_j, \qquad v_{j\alpha} = \frac{v_{l\alpha}}{v_l} v_j; \tag{17}$$

then (16) extends (13) to the whole of $\mathcal{R}_h$.

If instead, $j \in p(\mathcal{N}_\alpha) \cap \mathcal{R}_{h-1}$, we analyze (16) as follows. By (13) and (16) we have

$$\eta_j^r S^r v_j = \left(\frac{\eta_l}{\eta_{l\alpha}} n_{j\alpha}\right)^r S^r \frac{v_l}{v_{l\alpha}} v_{j\alpha}$$

and therefore

$$\frac{v_l v_{j\alpha}}{v_j v_{l\alpha}} = \left(\frac{\eta_j \eta_{l\alpha}}{\eta_l \eta_{j\alpha}}\right)^r.$$

The right-hand side of this equation is independent of **x**, while the left-hand side is in k. Thus their common value is a constant $\gamma_{j\alpha} \in k$. This means that

$$v_{j\alpha} = \gamma_{j\alpha} \frac{v_{l\alpha}}{v_l} v_j \tag{18}$$

and

$$\eta_{j\alpha} = \gamma_{j\alpha}^{-1/r} \frac{\eta_{l\alpha}}{\eta_l} \eta_j. \tag{19}$$

By (17), (18), (19) we see that if we define $\gamma_{j\alpha}$ to be 1 for $j \in p(\mathcal{N}_\alpha)$ but with $j \notin \mathcal{R}_{h-1}$, then (18) and (19) hold for every $j \in p(\mathcal{N}_\alpha)$. Now Eq. (19) expresses compatibility relations which need to hold among the $\eta_{j\alpha}$'s. We substitute (18) into (15) and obtain

$$\sum_{p(\mathcal{N}_\alpha)} c_j(\mathcal{N}_\alpha)\, \gamma_{j\alpha} \frac{v_{l\alpha}}{v_l} v_j = 0,$$

which, after division by $v_{l\alpha}/v_l$, can be written as

$$\sum_{\mathcal{R}} c_{hj} v_j = 0$$

with $c_{hj} = c_j(\mathcal{N}_\alpha)\, \gamma_{j\alpha}$ for $j \in p(\mathcal{N}_\alpha)$ and $c_{hj} = 0$ elsewhere. This completes the induction step. The last part of Lemma 4 is also clear, since we take into account relations (10) for all pairs $\{J, \mathcal{N}\}$, and thus Lemma 4 is established.

The proof of our theorem follows immediately from Lemma 4. In fact, each choice of a collection $\{J, \pi(J)\}$ determines a family $\bigcap_J \mathscr{X}(\pi(J))$ of solutions, which by Lemma 4 is parametrized as we want. Since

$$\mathscr{X} = \bigcap \bigcap_J \mathscr{X}(\pi(J))$$

with $\bigcup$ over all possible collections of partitions $\{J, \pi(J)\}$, a bound for the number of families is obtained by counting the number of choices for $\{J, \pi(J)\}$. For each J we have $(m+1)!^{(m+1)!}$ choices of π, and we have $\binom{n}{m+1}$ possibilities for J. Thus the number of families of solutions is at most

$$\#\{\text{families}\} \leqq ((m+1)!^{(m+1)!})^{\binom{n}{m+1}} \leqq n!^{n!} \leqq n^{n^n}.$$

References

[Br–M] W. D. Brownawell and D. W. Masser, Vanishing sums in function fields, *Math. Proc. Cambridge Philos. Soc.* **100** (1986), 427–434.

[Ch] S. Chowla, Remarks on two theorems of Siegel, *Arcta Arith.* **9** (1964), 417–418.

[M] J. Mueller, Binomial Thue's equation over function fields, *Compositio Math.* **73** (1990), 189–197.

[Ma] R. C. Mason, Diophantine equations over function fields, *in* "London Math. Society Lecture Notes," Vol. 96, Cambridge Univ. Press, London/New York, 1984.

[Sh] I. R. Shafarevich, "Basic Algebraic Geometry," Springer-Verlag, Berlin/Heidelberg/New York, 1977.

[Sil] J. H. Silverman, The Catalan equation over function fields, *Trans. Amer. Math. Soc.* **273** (1982), 201–205.

[V] J. F. Voloch, Diagonal equations over function fields, *Bol. Soc. Brasil. Math.* **16** (1985), 29–39.