120017341_AGB31_02_R1_120502

COMMUNICATIONS IN ALGEBRA®
Vol. 31, No. 2, pp. 1–23, 2003

Vanishing Polynomial Sums

L. N. Vaserstein[1,*] and E. R. Wheland[2]

[1]Department of Mathematics, The Pennsylvania State University, University Park, Pennsylvania, USA
[2]Department of Mathematics and Computer Science, The University of Akron, Akron, Ohio, USA

1. INTRODUCTION

AQ1

The Euler-Fermat equations are

$$x_1^n + \cdots + x_m^n = x_0^n \tag{1.1}$$

with unknown x_i. In 1769 Euler conjectured that (1.1) has no positive integral solutions when $m < n$. This generalizes both his result for $n = 3$ and Fermat's Last Theorem. However his conjecture was refuted for $m + 1 = n = 5$ (Lander and Parkin, 1966) and then for $m + 1 = n = 4$ (Elkies, 1988). It is still unknown whether (1.1) has solutions with $m + 1 = n \geq 6$ or with $m = 3 = n - 2$. On the other hand, (1.1) has positive

*Correspondence: L. N. Vaserstein, Department of Mathematics, The Pennsylvania State University, University Park, PA 16802, USA. E-mail: vstein@math. psu.edu.

DOI: 10.1081/AGB-120017341 0092-7872 (Print); 1532-4125 (Online)

www.dekker.com

integral solutions for $m \geq G(n)$ where $G(n)$ is a number related with the Waring problem, whose best known bound for large n is $G(n) \leq n(\log n + \log\log n + O(1))$ (Wooley, 1992).

In this paper, we are looking for solutions in $F[t]$ rather than in the integers. Here F is a field of characteristic $p \geq 0$ and $F[t]$ is the polynomial ring in a variable t with coefficients in F.

When $\mathrm{char}(F) = 0$, it is known that (1.1) has no nontrivial solutions in $F[t]$ when $m < (n/8)^{1/2}$ (Newman and Slater, 1979). *Nontrivial* here means primitive, indecomposable, nonconstant. A solution to (1.1) (or to the more general equation (1.2) below) is called primitive if $\gcd(x_0, x_1, \ldots, x_m) = 1$. We call it indecomposable if no partial sum on the left hand side is 0. We call it nonconstant if not all x_i belong to F.

When $p = \mathrm{char}(F) \neq 0$, given any solution of (1.1) we can obtain solutions of a similar equation with the same m but n replaced by np, by taking the p-th power. So in the case of finite characteristic we can have nontrivial solutions of (1.1) with a small m and arbitrarily large n. Moreover, it is known that (1.1) may have nontrivial solutions with $m = 3$ and all n (including all n coprime with p) (Paley, 1933). Furthermore, even for any fixed x_0, (1.1) with $m = 3$ may have nontrivial solutions with arbitrarily large n coprime with p (Vaserstein, 1991). See Sec. 5.

Returning to the case $p = 0$, we consider equations more general than (1.1), the generalized Fermat equations

$$c_0 x_0^{n_0} + \cdots + c_m x_m^{n_m} = 0, \tag{1.2}$$

in $m+1$ unknowns x_i with given $n_i \geq 2$ and nonzero $c_i \in F$. Our goal is to prove that (1.2) has no nontrivial solutions in $F[t]$ when n_i are "large" in a certain sense.

It is stated in (Newman and Slater, 1979, p. 486) where $F = \mathbf{C}$, the complex numbers, that when $m \leq (n/8)^{1/2}$, the equation (1.1) has no solutions in $F[t]$ with nonconstant x_i which are pairwise coprime. Actually, the method of Newman and Slater (1979) works when we relax the above condition to the following condition: the solution must be be primitive, indecomposable and nonconstant. Moreover, this method, involving the Wronskian of x_i, works to prove the following:

Theorem 1.3. *Let* $\mathrm{char}(F) = 0$. *Then* (1.2) *has no nontrivial solutions in* $F[t]$ *when* $\sum 1/n_j \leq 1/(m-1)$ *and* $0 \neq c_j \in F$. *In particular,* (1.1) *has no nontrivial solutions in* $F[t]$ *when* $n \geq m^2 - 1$.

Thus, this relaxes the condition $m \leq (n/8)^{1/2}$ of Newman and Slater (1979) to the condition $m \leq (n+1)^{1/2}$. Proposition 5.2 of Darmon and

Granville (1995) is our Theorem 1.3 with the redundant condition that all x_i are pairwise coprime.

On the other hand, it is shown in (Newman and Slater, 1979, p. 486) (with attribution to Molluzzo) that (1.1) has a primitive nonconstant solution with an arbitrary x_0 when $m = [(4n+1)^{1/2}]$ and $F = \mathbf{C}$. Moreover, this solution works for any algebraically closed field F of any characteristic. It is still unknown even for $F = \mathbf{C}$ whether (1.1) may have nontrivial solutions when $(n+1)^{1/2} < m < [(4n+1)^{1/2}]$.

We prove Theorem 1.3 after improving upon results of Voloch (1985, 1998); Brownwell and Masser (1986), and Zannier (1993) generalizing the (abc)-theorem (see the next section). An interesting question outside the scope of the present paper is to establish a bound on the number of nontrivial solutions of (1.2) with given $c_i \in F[t]$ for 'large' n_i. Results in this direction are known when char(F) = 0 (see Beukers, 1998; Beukers and Schlickewei, 1996 and Mueller, 1993, 2000) or char(F) $\neq 0$ and $m = 2$ (Voloch, 1998).

2. THE (ABC)-THEOREM AND ITS GENERALIZATIONS

For any nonzero polynomial $f = f(t) \in F[t]$ let $\nu(f)$ denote the number of $\alpha \in \bar{F}$, an algebraic closure of F, whose multiplicity in f is not divisible by p. When $p = \text{char}(F) = 0$, $\nu(f)$ is just the number of distinct zeros of f.

Theorem 2.1. *If $a + b = c$ in $A = F[t]$ with $p = 0$, $gcd(a, b) = 1$ and $(abc)' \neq 0$ then* $\deg(c) \leq \nu(abc) - 1$.

This well-known theorem (see Mason, 1986 and Lang, 1999) implies that the Fermat equation $x^n + y^n = z^n$ has no nontrivial solutions in $A = F[t]$ when $p = 0$ and $n \geq 3$ (this was known before Newman and Slater (1979), but in Newman and Slater (1979) it was proved for $n \geq 32$). It also has other applications (Mueller, 1993).

For a, b, c as in Theorem 2.1, one can ask whether there are restrictions on $N = \nu(abc)$, $s = \deg(c)$ other than the restrictions $s \leq N - 1$ given by the theorem and the obvious restrictions $0 \leq s$ and $N \geq 2$. It can be shown by examples that the answer is "no" when char(F) = 0, while the complete description of all possible pairs (s, N) in the case of $p \neq 0$ seems to be a difficult open problem.

Since we can switch a or b with -c in Theorem 2.1, its conclusion can be written as follows: max(deg(a), deg(b), deg(c)) $\le \nu$(abc) -1. It follows from our Theorem 3.1 that min(deg(a), deg(b), deg(c)) $\le \nu(abc)$ -2. On the other hand, it is clear that $\nu(abc) \le$ deg(a) +deg(b) +deg(c)) ≤ 3 max(deg(a), deg(b), deg(c)).

When char$(F) = 0$, it can be shown by examples that for any given integers M, s, N such that $0 \le s \le M \le N-1$ and $s \le N-2 \le s+2M-2$ there are a, b, $c \in F[t]$ such that deg$(a) = s$, deg$(b) =$ deg$(c) = M$, $a+b=c$, and $\nu(abc) = \nu(a) + \nu(b) + \nu(c) = N$. A more general open problem is to describe all possible 6-tuples (deg(a), $\nu(a)$, deg(b), $\nu(b)$, deg(c), $\nu(c)$). This seems to be a difficult problem even when char$(F) = 0$.

Here is our generalization of Theorem 2.1 in the case char$(F) = 0$:

Theorem 2.2. *Let $m \ge 2$, $y_j \in A = F[t]$, $y_1 + \cdots + y_m = y_0$, gcd$(y_1, \ldots, y_m) = 1$, char$(F) = 0$, not all y_j are constants, and no nonempty subsum of $y_1, \ldots, y_m$ vanishes. Then:*

(a) deg$(y_0) < (m-1)(\sum_{j=0}^{m} \nu(y_j))$.
(b) deg$(y_0) \le (\nu(y_0 \ldots y_m) - 1)m(m-1)/2$.

We will obtain the theorem in Secs. 3 and 4 below as a corollary of more precise results. Namely, we will prove the theorem first in the case when $y_1, \ldots, y_m$ are linearly independent over F (see Theorem 3.1 below which is an improved version of Theorem 2.2 in this case) and then in the case when $y_1, \ldots, y_m$ are linearly dependent over F (see Theorem 4.1 below which is a more precise version of Theorem 2.2(b) and Corollary 4.8 below which is a more precise version of Theorem 2.2(a)).

Part (b) was first stated in Mason, 1986, first proved in Brownwell and Masser (1986, Theorem B) (with $\nu(y_0 \ldots y_m)$ instead of $\nu(y_0 \ldots y_m) - 1$) and then proved in a more precise form in Zannier (1993). In the case when $y_1, \ldots, y_m$ are linearly independent over F, Theorem 2.2 is a particular case (with $g = 0$) of results of Brownwell and Masser (1986) and Voloch (1985) (although part(a) is not stated there explicitly). It is unknown whether the number $m(m-1)/2$ in (b) can be replaced by a smaller number when $m \ge 4$. It is proved in Browkin and Brzeziński (1994) that the correct number must be at least $2m-3$ and conjectured (and proved for $m = 3$) that the number is $2m-3$.

In the case $m = 2$ then both (a) and (b) coincide with Theorem 2.1 with $p = 0$, because in this case we have $\nu(y_0 y_1 y_2) = \nu(y_0) + \nu(y_1) + \nu(y_2)$ and $1 = m-1 = m(m-1)/2$.

Note also that (a) implies Theorem 1.3. Indeed, by obvious symmetry of the equation $y_1 + \cdots + y_m = y_0$, the number $\deg(y_0)$ in Theorem 2.2 can be replaced by $\deg(y_j)$ for any j and hence by $M = \max_j \deg(y_j)$. Dividing (a) by M, we obtain that

$$1 < (m-1)\sum_{i=0}^{m} \nu(y_i)/M \leq (m-1)\sum_{j=0}^{m} \nu(y_j)/\deg(y_j).$$

The number $\deg(y_j)/\nu(y_j)$ is the arithmetic mean of multiplicities of zeros of y_i while $(m+1)/\sum_{i=0}^{m}\nu(y_i)/\deg(y_i)$ is the harmonic average of $\deg(y_j)/\nu(y_j)$.

Corollary 2.3. *Under the conditions of Theorem* 2.2, *the harmonic mean of* $\deg(y_j)/\nu(y_j)$ *is less than* $m^2 - 1$.

This corollary generalizes Theorem 1.3. (Applying the corollary to the case $y_0 = -c_0 x^{n_0}$, $y_i = c_j x_j^{n_j}$ for $j \geq 1$, we obtain Theorem 1.3.)

3. IMPROVING UPON THEOREM 2.2 IN THE LINEARLY INDEPENDENT CASE

In this section we assume that $\operatorname{char}(F) = 0$. Our main goal is to prove Theorem 2.2 in the case when $y_1, \ldots, y_m$ are linearly independent over F. In fact we state and prove a more precise version of Theorem 2.2, Theorem 3.1, involving the orders $\operatorname{ord}_\alpha(y_j)$. We also obtain some known results as corollaries of Theorem 3.1.

First of all, for any integer $d \geq 1$ and any sequence $k_1, \ldots, k_d$ of integers we will define its "diversity" $\Delta(k_1, \ldots, k_d)$, which is an integer between 0 and $(d-1)d/2$. Namely, we consider the set X of all sequences $l_1, \ldots, l_d$ of distinct integers such that $l_j \geq k_j$ for all js. Then $\Delta(k_1, \ldots, k_d)$, our diversity of $k_1, \ldots, k_d$, is the maximal value of $(d-1)d/2 - \sum_{j=1}^{d}(l_j - k_j)$. In other words, we compute the minimal total increase to make all k_j distinct and then subtract it from its maximal possible value, $(d-1)d/2$. It is clear that $0 \leq \Delta(k_1, \ldots, k_d) \leq (d-1)d/2$. (To see the first inequality, note that $\{k_j + j\} \in X$ when $\{k_j\}$ is nondecreasing.) It is also clear that $\Delta(k_1, \ldots, k_d) = 0$ if and only if $\{k_j\}$ is a constant sequence and that $\Delta(k_1, \ldots, k_d) = (d-1)d/2$ if and only if all k_j are distinct. Here are two other easy upper bounds for $\Delta(k_1, \ldots, k_d)$:

$$\Delta(k_1,\ldots,k_d) \le \sum_{j=1}^{d} \min(j-1, k_j - k_1)$$

when $k_1 \le \cdots \le k_d$ and

$$\Delta(k_1,\ldots,k_d) \le (d-1)d/2 - \sum_{s} (m_s - 1)m_s/2,$$

where m_s is the total number of js such that $k_j = s$.

Now let $y_1,\ldots,y_d \in F[t]$ be linearly independent over F, and we set $y_0 = y_1 + \cdots + y_d$. For every $\alpha \in \bar{F}$ let $k_0 \le \cdots \le k_d$ be the numbers $\mathrm{ord}_\alpha(y_j)$ $(0 \le j \le d)$ written in nondecreasing order. We set

$$\Delta_\alpha(y_1,\ldots,y_d) = \Delta(k_1,\ldots,k_d).$$

Similarly, let $k_0 \le \cdots \le k_d$ be the numbers $\{\mathrm{ord}_\infty(y_j) = -\deg(y_j) : 0 \le j \le d\}$ written in nondecreasing order. We set

$$\Delta_\infty(y_1,\ldots,y_d) = \Delta(k_1,\ldots,k_d).$$

Thus, $\Delta_\alpha(y_1,\ldots,y_d)$ is defined for all $\alpha \in P^1(\bar{F}) = \bar{F} \cup \{\infty\}$. For almost all α (all but finitely many) $\mathrm{ord}_\alpha(y_j) = 0$ hence $\Delta_\alpha(y_1,\ldots,y_d) = 0$.

Theorem 3.1. *Let $m \ge 2$, $y_j \in A = F[t]$, $y_1 + \cdots + y_m = y_0$, $\gcd(y_1,\ldots,y_m) = 1$, $char(F) = 0$, and assume that $y_1, \ldots, y_m$ are linearly independent over F. Then*

$$\begin{aligned}\deg(y_0) &\le -(m-1)m + \sum_{\alpha \in P^1(\bar{F})} \Delta_\alpha(y_1,\ldots,y_m)\\ &\le -(m-1)m/2 + (m-1)\min\left(\nu(y_0\ldots y_m)m/2, \sum_{j=0}^{m} \nu(y_j)\right).\end{aligned}$$

Note that $\Delta_\alpha(y_1,\ldots,y_m) \ne 0$ if and only if either α is a zero of $y_0 \ldots y_m$ (recall that $\gcd(y_1,\ldots,y_m) = 1$) or $\alpha = \infty$ and not all $\deg(y_j)$ are the same. Therefore

$$\deg(y_0) \le (\mathrm{card}(S) - 2)m(m-1)/2$$

when S is a finite subset in $P^1(\bar{\mathrm{F}})$ and all y_j are units outside S. Thus, our Theorem 3.1 contains Theorem 4 of Voloch (1985, 1998) (which is essentially the same as Corollary 1 of Brownwell and Masser (1986) with $g=0$. Moreover $\Delta_\alpha(y_1,\ldots,y_m) \le (m-1)m/2-(l-2)(l-1)/2$ where l is the number of js with $\mathrm{ord}_\alpha(y_j)$ taking a fixed value. (In the case when this value is not minimal, $(l-2)(l-1)/2$ can be replaced by $(l-1)l/2$.) Therefore our Theorem 3.1 contains Theorem A of Brownwell and Masser (1986) with $g=0$. Moreover, the number S of places α where not all $\mathrm{ord}_\alpha(y_j)$ are 0 in Brownwell and Masser (1986) and Voloch (1985, 1998) can be replaced by the number $S' \le S$ of places α where not all $\mathrm{ord}_\alpha(y_j)$ are the same.

Theorem 3.1 gives Theorem 2.2 in the case when the polynomials $y_1,\ldots,y_m$ in Theorem 2.2 are linearly independent (just use that $\Delta_\alpha(y_1,\ldots,y_m) \le m(m-1)/2$ for all α to get Theorem 2.2 (b) and that $\Delta_\alpha(y_1,\ldots,y_m) \le (m-1)\{\#j : \mathrm{ord}_\alpha(y_j) \ne 0\}$).

After Theorem 3.1 is proved we will obtain some known results as its consequences, and then, in the next section, we use it to prove an improved version of Theorem 2.2, Theorem 4.1.

To prove Theorem 3.1, we introduce orders of finite-dimensional F-subspaces of $F[t]$.

Let V be an F-vector subspace of $A=F[t]$ of dimension d. Assume that $1 \le d < \infty$. For any $\alpha \in \bar{F}$, we can choose a basis $v_1,\ldots,v_d$ of V such that $\mathrm{ord}_\alpha(v_1) < \cdots < \mathrm{ord}_\alpha(v_d)$. These numbers $\mathrm{ord}_\alpha(v_i)$ are independent of the choice of such a basis. We define $\mathrm{ord}_\alpha(V)$ by

$$\mathrm{ord}_\alpha(V) = -(d-1)d/2 + \sum_{i=1}^{d} \mathrm{ord}_\alpha(v_j).$$

Clearly, $\mathrm{ord}_\alpha(V) \ge 0$. For any basis $y_1,\ldots,y_d$ of V we have

$$\mathrm{ord}_\alpha(V) \ge -(d-1)d/2 + \sum_{i=1}^{d} \mathrm{ord}_\alpha(y_i).$$

This is because by addition operations with the polynomials y_j we can make all $\mathrm{ord}_\alpha(y_j)$ distinct while increasing them or keeping them the same. To get a better bound, we can replace one of y_j with minimal $\mathrm{ord}_\alpha(y_j)$ by $y_0 = y_1 + \cdots + y_d$. This proves a more precise inequality:

Lemma 3.2. *Let V be as above and $y_1,\ldots,y_d$ a basis of V over F. Then, for all $\alpha \in \bar{F}$, $\mathrm{ord}_\alpha(V) \ge -\Delta(\mathrm{ord}_\alpha(y_1),\ldots,\mathrm{ord}_\alpha(y_d)) + \sum_{j=1}^{d} \mathrm{ord}_\alpha(y_j)$ and $\mathrm{ord}_\alpha(V) \ge -\Delta_\alpha(y_1),\ldots,y_d) + \sum_{j=1}^{d} \mathrm{ord}_\alpha(y_j) - \min(\mathrm{ord}_\alpha(y_0),\ldots,\mathrm{ord}_\alpha(y_d))$.*

Also we can choose a basis $v_1, \ldots, v_d$ of V such that $\deg(v_1) < \cdots < \deg(v_d)$. The numbers $\deg(v_i)$ are independent of the choice of such a basis. We set

$$\deg(V) = -d(d-1)/2 + \sum_{i=1}^{d} \deg(v_i)$$

and

$$\mathrm{ord}_\infty(V) = -d(d-1)/2 + \sum_{i=1}^{d} \mathrm{ord}_\infty(v_i),$$

where $\mathrm{ord}_\infty(v_i) = -\deg(v_i)$.

Clearly, $\deg(V) \geq 0$ and $\deg(V) + \mathrm{ord}_\infty(V) = -d(d-1)$. For any basis $v_1, \ldots, v_d$ of V we have $\deg(V) \leq -d(d-1)/2 + \sum_{i=1}^{d} \deg(v_i)$. Moreover, we have the following analogue of Lemma 3.2:

Lemma 3.3. *The conclusions of Lemma* 3.2 *also hold for* $\alpha = \infty$.

Proof. It is similar to that of Lemma 3.2. Also Lemma 3.3 can be reduced to Lemma 3.2 by a change of variable. □

Proposition 3.4. $\deg(V) = \sum_{\alpha \in \bar{F}} \mathrm{ord}_\alpha(V)$.

Proof. We pick a basis $v_1, \ldots, v_d$ of V and consider the Wronskian $D = \det(v_i^{(j)})$. It is a nonzero polynomial in $F[t]$, and it is determined by V up to a nonzero scalar factor. Choosing v_i as in the definition of $\deg(V)$, we see that $\deg(D) = \deg(V)$. Indeed, writing the determinant D as a sum of $d!$ terms, we see that every nonzero term has degree $\deg(V)$ and that the sum of coefficients in this degree is, up to a nonzero scalar factor equal to the product of leading terms in v_i, the determinant of the matrix $\left(\binom{k_i}{j-1}\right)_{1 \leq i,j \leq d}$ which is, up to the factor $(1! \ldots (d-1)!)^d$, the Vandermonde determinant $\det(k_i^j) = \prod_{i<j}(k_i - k_j) \neq 0$ where $k_i = \deg(v_i)$. □

On the other hand, choosing v_i as in the definition of $\mathrm{ord}_\alpha(V)$, we see that $\mathrm{ord}_\alpha(D) = \mathrm{ord}_\alpha(V)$ for all $\alpha \in \bar{F}$. Indeed writing the determinant D as a sum of $d!$ terms, we see that every term is divisible by $(x-\alpha)^{\mathrm{ord}_\alpha(V)}$ and that the sum of the terms of order $\mathrm{ord}_\alpha(V)$ is, up to a factor of order 0 equal to the product of $v_i/(x-\alpha)^{\mathrm{ord}_\alpha(v_i)}$, the determinant of the matrix $\left(\binom{k_i}{j-1}\right)_{1 \leq i,j \leq d}$ which is, up to the factor $(1! \ldots (d-1)!)^d$, the Vandermonde determinant $\det(k_i^j) = \prod_{i<j}(k_i - k_j)$ where $k_i = \deg(v_i)$.

Now $\deg(V) = \deg(D) = \prod_{\alpha\in\bar{F}} \mathrm{ord}_\alpha(D) = \prod_\alpha(V)$.

Remark. When $d = 1$, the proposition says that $\deg(D) = \sum_{\alpha\in\bar{F}} \mathrm{ord}_\alpha(D)$ for every nonzero polynomial $D \in F[t]$.

Remark. As a corollary of the proposition, $\sum_\alpha \mathrm{ord}_\alpha(V) < \infty$.

Now we are ready to prove Theorem 3.1. Let $d_0 \le d_1 \le \cdots \le d_m$ be the degrees of $y_j\,(0 \le j \le m)$ written in nondecreasing order. Then $d_{m-1} = d_m \ge m$. By Lemma 3.3 with $d = m$,

$$\deg(V) = -m(m-1) - \mathrm{ord}_\infty(V) \le -m(m-1) + \Delta_\infty(y_1, \ldots, y_m) - d_m + \sum_{j=0}^{m} d_j.$$

On the other hand, since $\gcd(y_1, \ldots, y_m) = 1$, for any $\alpha \in \bar{F}$ we have

$$\min(\mathrm{ord}_\alpha(y_1), \ldots, \mathrm{ord}_\alpha(y_m)) = 0$$

and, by Lemma 3.2 with $d = m$,

$$\mathrm{ord}_\alpha(V) \ge -\Delta_\alpha(y_1, \ldots, y_m) + \sum_{j=1}^{m} \mathrm{ord}_\alpha(y_j).$$

Switching, if neccessary, $-y_0$ with a polynomial y_j such that $y_j(\alpha) \ne 0$, we have

$$\mathrm{ord}_\alpha(V) \ge -\Delta_\alpha(y_1, \ldots, y_m) + \sum_{j=0}^{m} \mathrm{ord}_\alpha(y_j).$$

Taking the sum over all $\alpha \in \bar{F}$ and using Proposition 3.1 and the fact that $\deg(y_j) = \sum \mathrm{ord}_\alpha(y_j)$ for each j, we obtain $\deg(V) = \sum_{\alpha\in\bar{F}} \mathrm{ord}_\alpha(V) \ge -\sum_{\alpha\in\bar{F}} \Delta_\alpha(y_1, \ldots, y_m) + \sum_{j=0}^{m} d_j$.

Comparing these upper and lower bounds for $\deg(V)$ and cancelling $\sum_{j=0}^{m} d_j$, we obtain the conclusion of Theorem 3.1.

Here are some consequences of Theorem 3.1.

Corollary 3.5 (cf. Theorem on p. 480 of Newman and Star (1979)). *Let $n \ge 3$ and $\sum_{j=1}^{m} x_j^n = t$ in $F[t]$ with $\mathrm{char}(F) = 0$. Then $m^2 - m > n$.*

Proof. Without loss of generality, we can assume that F is algebraically closed and that $x_1^n, \ldots, x_m^n$ are linearly independent (otherwise, t is a sum of a smaller number of n-th powers). Also $\gcd(x_1^n, \ldots, x_m^n) = 1$ because this is an nth power dividing t and $n \geq 2$. Let $M = \max(\deg(x_1), \ldots, \deg(x_m))$. By Theorem 3.1,

$$
\begin{aligned}
Mn &\leq -(m-1)m/2 + (m-1)\left(\nu(t) + \sum_{j=1}^{m} \nu(x_j)\right) \\
&\leq -(m-1)m/2 + (m-1)(1 + Mm) < Mm(m-1),
\end{aligned}
$$

because $\nu(y_j) = \nu(x_j) \leq \deg(x_j) \leq M$, $\nu(t) = 1$, and $m \geq 3$ when $n \geq 3$. □

Remark. The condition $n \geq 3$ is necessary, and it is implicit in the theorem on p. 480 of Newman and Slater (1979), see p. 479 or 481 in Newman and Slater (1979).

Corollary 3.6. *Let $n \geq 2$ and $\sum_{j=1}^{m} x_j^n = tx_0^n$ in $F[t]$ with $x_0 \neq 0$ and* $\mathrm{char}(F) = 0$. *Then $m^2 > n + 1$.*

Proof. Without loss of generality, we can assume that F is algebraically closed, that $x_1^n, \ldots, x_m^n$ are linearly independent, and that $\gcd(x_1, \ldots, x_m) = 1$. Let

$$M = \max(\deg(x_1), \ldots, \deg(x_m)).$$

By Theorem 3.1, $Mn < (M(m+1+1))(m-1) - m(m-1)/2$, hence $n < m^2 - 1$ when $n \geq 3$. When $n = 2$, notice that $m \geq 2$. □

Remark. This improves upon the theorem on p. 483 of Newman and Slater (1979) which asserts that $m^2 > n/8$.

Theorem 3.7. *Let $n \geq 2$ and $x_1^n + \cdots + x_m^n = x_0$ in $F[t]$ with char$(F) = 0$. Assume that $m \geq 1$ and $x_1^n, \ldots, x_m^n$ are linearly independent over F. Then*

$$\deg(x_0) \geq M(n - m^2 + m) + (m-1)m/2,$$

where $M := \max(\deg(x_1), \ldots, \deg(x_m))$.

Proof. Without loss of generality, we can assume that F is algebraically closed, that $m \geq 2$ (the case $m = 1$ is trivial), that $\gcd(x_1, \ldots, x_m) = 1$ (the general case follows easily from the primitive one), that $n > m(m-1)/2$ (otherwise there is nothing to prove), that $x_1^n, \ldots, x_m^n$ are linearly independent over F (otherwise the number m can be reduced), and that $M = \deg(x_m)$. Using Theorem 3.1 with $y_0 = -x_m^n$, $y_m = -x_0$, $y_j = x_j^n$ for $j \leq m-1$ and the fact that

$$\Delta_\alpha(y_1, \ldots, y_m) \leq (m-1)\left(\sum_{j=0}^{m-1} \operatorname{sign}(\operatorname{ord}_\alpha(y_j)\right) + \operatorname{ord}^\alpha(ym)$$

for all $\alpha \in \bar{F}$, we obtain that

$$\begin{aligned} nM &= \deg(y_0) \leq \deg(x_0) + \sum_{j=0}^{m-1} \nu(y_j) - m(m-1)/2 \\ &\leq \deg(x_0) + (m-1)mM - m(m-1)/2, \end{aligned}$$

hence $\deg(x_0) \geq M(n - m^2 + m) + m(m-1)/2$.

Remark. Our theorem is close to the theorem on p. 481 of Newman and Slate (1979) and has the same conclusion as the second conclusion of that theorem. The condition that $n \geq 2$ is implicit in Newman and Slate (1979) (without this condition the theorem is not true). We relaxed the condition of Newman and State (1979) that all x_j are nonconstant and nonlinear to the condition that not all of them are constant. We imposed an additional condition that $x_1^n, \ldots, x_m^n$ are linearly inde pendent over F to prevent counterexamples to the theorem of [MS] AQ2 like the following two:

(i) $m = 2$, any $n \geq 2$, any $M \geq 1$, $x_1 = t^M$, $x_2 = -t^M$, $x_0 = 0$.
(ii) $m = 2$, any $n \geq 3$, any $M \geq 1$, $x_1 = t^M$, $x_2 = t^M$, $x_0 = 2t^M$.

The condition is missing in Newman and Slate (1979) by mistake. When $F = \bar{F}$, this condition is weaker than the condition that x_0 is not a sum of a smaller than m number of nth powers in $F[t]$. It could also be replaced by the condition that x_0 is not divisible by any nonconstant n-th power in $F[t]$.

The first conclusion of the theorem of Newman and Slate (1979) is that $\deg(x_0) \geq n - (m-1)m/2$. However, this conclusion is stated twice on the same page as $\deg(x_0) \geq n - (m-1)m$, which follows from our theorem because either $M = 0$ and $m = 1$ or $M \geq 1$. Also the theorem implies that $\deg(x_0) \geq n - (m-1)m$ when either $n \geq m^2 - m$ or $M = 1$. By the way, when $M = 1$ it is easy to prove that in fact $\deg(x_0) \geq n - m + 1$, and that this bound cannot be improved when $n \geq m - 1$ and $F = \bar{F}$.

Theorem 2.2(a) with linearly independent $y_1, \ldots, y_m$ (which we have obtained from Theorem 3.1) allows us now to obtain Theorem 1.3 in the following two cases (in general, Theorem 1.3 follows from Theorem 2.2 which we will obtain from Theorem 4.4 below).

Corollary 3.8. *The equation* (1.1) *has no nonconstant indecomposable primitive solutions when* $n \geq m^2 - 1$.

Proof. Consider such a solution of (1.1). Let $M = \max(\deg(x_1), \ldots, \deg(x_m))$. If $x_1^n, \ldots, x_m^n$ are linearly independent, then Theorem 3.1 via Theorem 2.2 (a) gives that $Mn < (m+1)M(m-1)$.

If $x_1^n, \ldots, x_m^n$ are linearly dependent, we consider a linear relation of minimal length. Permuting terms, we can assume that this relation is $c_1 x_1^n + \cdots + c_{d+1} x_{d+1}^n = 0$. If not all terms here are colinear, we divide this equality by the gcd and pass to $\bar{F}$ to conclude that $n < d^2 - 1 < m^2 - 1$. Otherwise, we also pass to $\bar{F}$ and replace all terms by one term of the form x'^n and after dividing by the gcd we obtain an equation of the form (1.1) with not all terms constant and with the same n but m replaced by $m - d < m$, hence we are done by induction on m.

Corollary 3.9. *The equation* (1.2) *has no nonconstant indecomposable solutions with pairwise coprime terms when* $\sum 1/n_j \leq 1(m-1)$.

Proof. Again either the terms are linearly independent and we are done by Theorem 3.1 or we can obtain a shorter equation of the same form (passing, if necessary, to $\bar{F}$). Here we cannot divide by the gcd because this produces nonconstant coefficients, but we do not need to do this because the terms are pairwise coprime.

4. IMPROVING THEOREM 2.2 IN GENERAL

Now we drop the condition about linear independence and prove a more precise version of Theorem 2.2 involving the dimension $d \leq m$ of $Fy_1 + \cdots + Fy_m$ and another number $d' \leq d$ which is defined below.

First we introduce two definitions which make sense for any sequence (family, multiset) $y_0, \ldots, y_m$ of vectors in an F-vector space.

Following Brownwell and Masser (1986), a nonempty subset $I \subset \{0, 1, \ldots, m\}$ is called *minimal* if the corresponding family $\{y_j\colon j \in I\}$ of vectors is linearly dependent over F but $\{y_j\colon j \in J\}$ is linearly independent over F for any proper subset $J \subset I$. A linear relation is called *minimal* if the corresponding family of vectors is minimal.

Clearly, $\mathrm{card}(I) \leq d+1$ for any minimal set I. We define $d'+1$ to be the maximal cardinality of the minimal sets. Thus, $d' \leq d = \dim(Fy_0 + \cdots + Fy_m)$.

In general not every index j needs to be an element of a minimal set. The condition $y_0 = y_1 + \cdots + y_m$ in Theorem 2.2 implies that every j is contained in a minimal set. Namely, we can take a basis not involving j, write y_j with respect to this basis, and drop zero terms to obtain such a set. This condition, together with the condition of no vanishing subsums of Theorems 2.2 and 4.1, implies that all linear relations between $y_0, y_1, \ldots, y_m$ follow from minimal relations (Lemma 4 of Brownell and Masser, 1986).

We call a family (multiset) $y_0, \ldots, y_m$ of vectors *irreducible* if for any partition $I \cup I' = \{0, 1, \ldots, m\}$ into nonempty parts, the corresponding vector subspaces $\sum_{j \in I} Fy_j$ and $\sum_{j \in I'} Fy_j$ have a nonzero intersection. For example, every minimal set gives an irreducible family of vectors. Under the conditions of Theorem 2.2, the family $y_0, \ldots, y_m$ is irreducible. It can be shown that for every irreducible family $z_0, \ldots, z_m$ of vectors there are nonzero $c_j \in F$ such that $y_j = c_j z_j$ satisfy the conditions of Theorem 2.2. So the conditions of the next two theorems are not really more general than those of Theorem 2.2 (but the conclusions are stronger).

Theorem 4.1. *Let $m \geq 2$, $y_j \in F[t]$, $\gcd(y_1, \ldots, y_m) = 1$, and $y_0, y_1, \ldots, y_m$ be an irreducible family. Let d be the dimension of $Fy_1 + \cdots + Fy_m$ over F and let $d' \leq d$ be as above. Then:*

(a) $\deg(y_0) \leq (\nu(y_0 \ldots y_m) - 1)(d'-1)d/2.$

Furthermore, if $\deg(y_1) = \cdots = \deg(y_m)$, *then:*

(b) $\deg(y_0) \leq (\nu(y_0 \ldots y_{\mathrm{m}}) - 1)(d'-1)d/2.$

Proof. First we consider the special case when $d = m$. Since the family $y_0, y_1, \ldots, y_m$ is irreducible, $d' = d = m$ and $y_0 = \sum_{j=1}^{m} c_j y_j$ with $0 \neq c_j \in F$. Applying Theorem 3.1, we obtain that

$$\deg(y_0) \le \deg(y_m) \le -(m-1)m + \sum\nolimits_{\alpha 1(\bar{F})} \Delta_\alpha(c_1y_1, \ldots, c_my_m).$$

Now we use that $\Delta_\alpha(c_1y_1, \ldots, c_my_m) \le (m-1)m/2$ for all $\alpha \in P^1(\bar{F})$ and that in the case $\hat{\ }(y_1) = \cdots = \deg(y_m)$ we have

$$\begin{aligned}\Delta_\infty(c_1y_1, \ldots, c_my_m) &= \min(\deg(y_m) - \deg(y_0), m-1) \\ &\le \deg(y_m) - \deg(y_0)\end{aligned}$$

to obtain the theorem.

Now we consider the general case. Note that $y_j \ne 0$ for all js because of the irreducibility condition. We proceed by induction on m. If $m = 2$, then either $d = 2$ and we are done by the special case above or $d = 1$ and all y_j are constants, so both our conclusions say that $0 \le 0$ which is true. Assume now that $d < m \ge 3$.

By the irreducibility condition, we can write y_0 as a linear combination of $y_1, \ldots, y_m$. Dropping js with zero coefficients, we obtain a minimal $I_0 \subset \{0, 1, \ldots, m\}$ containing 0 with $2 \le \mathrm{card}(I_0) \le d' + 1 \le d + 1 \le m$. The corresponding family $\{y_j: j \in I_0\}$ is an irreducible family. Let $\{y_j : j \in I_1\}$ be an irreducible family containing I_0 and with maximal possible $\mathrm{card}(I_1) \le m$. Set $m' = \mathrm{card}(I_1) - 1$, $V_1 = \sum_{j \in I_1} Fy_j$, $d_1 = \dim(V_1)$, and $z_1 = \gcd(V1)$.

By the induction hypothesis, AQ3

$$\begin{aligned}\deg(y_0/z_1) &\le \left(\nu(\prod_{j \in I_1}(y_j/z_1) - e\right) d_1(d'-1)/2 \\ &\le (\nu(y_0y_1 \ldots y_m) - e)d_1(d'-1)/2 \end{aligned} \tag{4.2}$$

where $e = 2$ when $\deg(y_1) = \cdots = \deg(y_m)$, and $e = 1$ otherwise.

When $z_1 = 1$, e.g., $d' = d$, we are done. In general, we need an upper bound for $\deg(z_1)$ to finish the proof. Let I_2 be the complement of I_1 in $\{0, 1, \ldots, m\}$ and $V_2 = \sum_{j \in I_2} Fy_j$. By the irreducibility condition, $V_1 \cap V_2 \ne 0$.

Let $0 \ne z = \sum_{j \in I'} c_jy_j \in V_1 \cap V_2$ be a vector with the least number of nonzero coefficients when written as a linear combination of $\{y_j: j \in I_2\}$ with minimal possible number $\mathrm{card}(I')$ of nonzero coefficients, where $I' \subset I_2$ and $c_j \in F$. This relation is minimal, hence $I_1 \cup I'$ is irreducible. By the maximality of I_1, we have $I' = I_2$. Thus, $m_2 := \mathrm{card}(I_2) = \dim(I_2) = d - d_1 = m - m_1 \le d'$. Set $z_2 = \gcd(V_2)$. Note that $\gcd(z_1, z_2) = 1$.

Applying Theorem 3.1, we obtain that

$$\begin{aligned}\deg(z_1) &\le \deg(z/z_2) \le \max(\deg(y_j/z_2)\colon j \in I_2) \\ &\le -(m_2-1)m_2 + \sum_{\alpha \in P^1(\bar{F})} \Delta_\alpha(c_j y_j/z_2\colon j \in I_2).\end{aligned} \tag{4.3}$$

Now we use different upper bounds for

$$\Delta_\alpha(c_j y_j/z_2\colon j \in I_2) = \Delta_\alpha(c_j y_j\colon j \in I_2)$$

in (4.3) for different $\alpha \in P^1(\bar{F})$. When α is not a zero of $(z/z_2)y_0 \dots y_m$ nor $\alpha = \infty$, we have $\Delta_\alpha(c_j y_j\colon j \in I_2) = 0$.

When α is a zero of $y_0 \dots y_m$, we use that

$$\Delta_\alpha(c_j y_j : j \in I_2) \le (m_2-1)m_2/2 \le (d-d_1)(d'-1)/2.$$

We use the same bound when $\alpha = \infty$, unless $\deg(y_1) = \cdots = \deg(y_m)$ in which case we use the following more precise bound:

$$\begin{aligned}\Delta_\infty(c_j y_j\colon j \in I_2) &= \min(m_2-1, \max(\deg(y_j)\colon j \in I_2) - \deg(z/z_2)) \\ &\le \max(\deg(y_j)\colon j \in I_2) - \deg(z_1).\end{aligned}$$

Finally, when α is a zero of z/z_2 but not a zero of $y_0 \dots y_m$, we use that

$$\Delta_\alpha(c_j y_j\colon j \in I_2) \le \mathrm{ord}\alpha(z/z_2) = \mathrm{ord}_\alpha;(z) - \mathrm{ord}\alpha(z_1).$$

Substituting these bounds to (4.3), we obtain that

$$\deg(z_1) \le \deg(z/z_2) \le (\nu(y_0 \dots y_m) - e)(d-d_1)(d'-1)/2.$$

Adding this to (4.2), we obtain the conclusions of the theorem.

Remark. Our Theorem 4.1(a) improves the main result of Zannier (1993, see (4) on page 88) with $g = 0$. The number μ in Zannier (1993) is our d, and we replaced $\mu(\mu-1)/2$ in the upper bound of Zannier (1993) by $d'(d-1)/2 \le d(d-1)/2$. The number S in Zannier (1993) is our $\nu(y_0 \dots y_m) + 1$ when $S \ne 0$; when $S = 0$ all terms are constant and $\nu(y_0 \dots y_m) = 0$. Our Theorem 4.1(b) shows that S in Zannier (1993) can be replaced by $S - 1$ in the case when all degrees are the same. Our proof followed those of Brownwell and Masser (1986) and Zannier (1993).

We neither stated nor proved extensions of our results to the case of arbitrary genus g although there are no difficulties in doing this. The number S in Brownwell and Masser (1986), Zannier (1993) can be replaced by the number $S' \leq S$ of places where not all $\mathrm{ord}_\alpha(y_j)$ are the same.

Our next result is similar to Theorem B of Brownwell and Masser (1986) (in the case $g=0$). For any nonzero polynomials $y_0, \ldots, y_m \in F[t]$ and any $\alpha \in P^1(\bar{F})$ let $\mu(\alpha)+1$ be the number of js with minimal value of $\mathrm{ord}_\alpha(y_j)$ (over $0 \leq j \leq m$). Note that $1 \leq \mu(\alpha) \leq \mathrm{m}$. When m 1 and either $y_0 = y_1 + \cdots + y_m$ or a family $y_0, \ldots, y_m$ is irreducible, we have $1 \leq \mu(\alpha)$ for all α. When $\gcd(y_0, \ldots, y_m) = 1$, the minimal value is 0 for all $\alpha \in \bar{F}$. for $\alpha = \infty$, $\mu(\infty)$ is the number of js with maximal $\deg(y_j)$.

Theorem 4.4. *Let $m \geq 2$, $y_j \in F[t]$, $\gcd(y_1, \ldots, y_m) = 1$, and $y_0, y_1, \ldots, y_m$ be an irreducible family. Let d be the dimension of $Fy_1 + \cdots + Fym$ over F and let $d' \leq d$ be as above (i.e., $d'+1$ is the maximal cardinality of the minimal sets). Then*

$$\deg(y_0) \leq (d'-1)\sum_{j=0}^{m} \nu(y_j).$$

If either $d=m$ or the index 0 belongs to a minimal set I_0 with $\mathrm{card}(I_0) = d'$, then $\deg(y_0) \leq -d'(d'-1)/2 + (d'-1)\sum_{j=0}^{m} \nu(y_j)$.

Proof. We follow closely the proof of Theorem 4.1.

Case 1. $d=m$. In this case $d' = d = m$ and $y_0 = \sum_{j=1}^{m} c_j y_j$ with $0 \neq c_j \in F$. So our statement follows from Theorem 3.1 because

$$\Delta_\alpha(c_1 y_1, \ldots, c_m y_m) \leq m(m-1)/2 \text{ for } \alpha = \infty \quad \text{and}$$
$$\Delta_\alpha(c_1 y_1, \ldots, c_m y_m) \leq (m-1)\sum_{j=0}^{m} \mathrm{sign}\,(\mathrm{ord})^\alpha(y_j))$$

for all $\alpha \in \bar{F}$.

Case 2. The index 0 belongs to a minimal set I_0 with $\mathrm{card}(I_0) = d'$.

We proceed by induction on m. If $m=2$, then either $d=2$ and we are done by Case 1 or $d=1$ and all y_j are constants, so our conclusion is that $0 \leq 0$. Assume now that $d < m \geq 3$.

Let I_0, I_1, I_2, d_1, m_1, z, z_1, z_2, m_2 be as in the proof of Theorem 4.1 with the following additional condition: we choose $I_0 \ni 0$ with $\mathrm{card}(I_0) = d'$, so $\mathrm{card}(I_1) \geq d'$.

By the induction hypothesis, AQ3

$$\deg(y_0/z_1) \leq -(d'-1)d'/2 + (d'-1)\sum_{j\in I_1} \nu(y_j/z_1). \tag{4.5}$$

If $z_1 = 1$, we are done because the right hand side in (4.5) is less than the right hand side in the conclusion of Theorem 4.4. Let now $z_1 \neq 1$. Then $d' \geq d_2 \geq 2$. We set $I_3 = I_2 \cup \{m+1\}$ and $y_{m+1} = z$. Applying Theorem 3.1, we obtain that

$$\begin{aligned} \deg(z_1) \leq \deg(z/z_2) &\leq d_2 \sum_{j\in I_3} \nu(y_j/z_2) \\ &\leq (d'-1)\nu(z/z_2) + (d'-1)\sum_{j\in I_2} \nu(y_j). \end{aligned} \tag{4.6}$$

Now it remains to add (4.5) and (4.6) and to observe that $\nu(z/z_2) \leq \nu(z_1)$ and $\sum_{j\in I_1} \nu(y_j/z_1) + \nu(z_1) \leq \sum_{j\in I_1} \nu(y_j)$.

General case. Set $M = \max(\deg(y_1), \ldots, \deg(y_m))$. We pick a $\beta \in \bar{F}$ which is not a zero of $y_0 \ldots y_m$ and replace $y_j = y_j(t) \in F[t]$ by $z_j = t^M y(1/t + \beta)$. Note that the family $z_0, z_1, \ldots, z_m \in F[t]$ is irreducible, $\gcd(z_1, \ldots, z_m) = 1$, $\deg(z_j) = M$ for all j, $\nu(y_j) = \nu(z_j)$ for at least two js, and $\nu(y_j) \leq \nu(z_j) \leq \nu(y_j) + 1$ for all j.

Now we can use Case 2 with y_j replaced by z_j and obtain that

$$\begin{aligned} \deg(y_0) \leq M = \deg(z_0) &\leq m - 2 - d'(d'-1)/2 + (d'-1) \\ &\sum_{j=0}^{m} \nu(z_j) = m - 2 - d'(d'-1)/2 + (d'-1)\sum_{j=0}^{m} \nu(y_j). \end{aligned}$$

If $m - 2 - d'(d'-1)/2 \leq 0$, we are done. Otherwise we replace $z_j = z_j(t)$ by $z_j(t^N)$ with a large integer N and obtain that

$$MN \leq m - 2 - d'(d'-1)/2 + M(d'-1)\sum_{j=0}^{m} \nu(y_j).$$

Dividing both sides by N and sending N to ∞, we obtain the conclusion of the theorem.

Corollary 4.7. *Under the conditions of Theorem* 4.1*,*

$$\deg(y_0) \le (d-1)\sum_{j=0}^{m} \nu(y_j).$$

If not all y_j are constant, then

$$\deg(y_0) < (m-1)\sum_{j=0}^{m} \nu(y_j).$$

If not all y_j are constant and they are pairwise coprime, then

$$\deg(y_0) < (d'-1)\sum_{j=0}^{m} \nu(y_j).$$

Proof. This follows from Theorem 4.1 using the following observations. If all y_j are constant, i.e., $d=1$, then the conclusion is that $0 \le 0$. When y_j are constant and they are pairwise coprime we can apply Theorem 3.1 or Theorem 4.1 to a minimal set containing the index 0.

Our next result is similar to Theorem B of Brownwell and Masser (1986) (in the case $g=0$). For any nonzero polynomials $y_0, \ldots, y_m \in F[t]$ and any $\alpha \in P^1(\bar{F})$ let $\mu(\alpha)+1$ be the number of js with minimal value of $\mathrm{ord}_\alpha(y_j)$ (over $0 \le j \le m$). Note that $1 \le \mu(\alpha) \le m$. When $m \ge 1$ and either $y_0 = y_1 + \cdots + y_m$ or a family $y_0, \ldots, y_m$ is irreducible, we have $1 \le \mu(\alpha)$ for all α. When $\gcd(y_0, \ldots, y_m) = 1$, the minimal value is 0 for all $\alpha \in \bar{F}$. For $\alpha = \infty$, $\mu(\infty)$ is the number of js with maximal $\deg(y_j)$.

Theorem 4.8. *Under the conditions of Theorem* 4.1*,*

$$\deg(y_0) \le 2d' - d'^2 - d + \sum_{\alpha \in P^1(\bar{F})} (d(d-1)/2 - \mu'(\alpha)(\mu'(\alpha)-1)/2),$$

where $\mu'(\alpha) = max(d-d',\ d-m+\mu(\alpha))$.

Proof. The proof is similar to those of Theorems 4.1 and 4.4 and we leave it to the reader.

5. NONZERO CHARACTERISTIC CASE

Now let $p = \mathrm{char}(F) \neq 0$. It was shown in Paley (1933, p. 54) that the equation

$$x_1^n + x_2^n = x_3^n + x_4^n \tag{5.1}$$

has nontrivial solutions for every n. When n is odd or F is algebraically closed, (5.1) is essentially the same equation as (1.1) with $m = 3$.

In Vaserstein (1991) it is shown that for every n, every polynomial in $F[t]$ is the sum of at most $n_{(p)}$ n-th powers provided that F is an algebraically closed field with $\mathrm{char}(F) = p \neq 0$ and n is coprime with p. Here $n_{(p)} = \prod(a_i + 1) - 1 \leq n$, where the a_i are the digits of n in base p. Moreover Vaserstein (1991) gives solutions with pairwise coprime x_i. It follows that (1.1) has nontrivial solutions (with an arbitrary x_0) when $m \geq S(n, p)$, where $S(n, p)$ is the minimum of $(nk)_{(p)}$ over all natural-numbers k. Obviously, $S(n, p)$ could be small for very large n. For example, $S(n, p) = 3$ if $n > 2$ divides $p^k + 1$ for some k. So for all such n and any x_0, the equation (1.1) with $m = 3$ has nontrivial solutions in $\bar{F}[t]$, where $\bar{F}$ is an algebraic closure of F.

Thus, when $p \neq 0$, no upper bound on m prevents the existence of indecomposable primitive nonconstant solutions of (1.1) in $F[t]$, and one cannot place any bound of the form $m < f(n)$ with $f(n) \to \infty$ as $n \to \infty$ to prevent pairwise coprime nonconstant solutions. One can hope to obtain a bound of the form $n < f(S(n, p))$ (under additional conditions on solutions).

The main trouble is that the determinant D in Sec. 3 may vanish even when $y_1, \ldots, y_m$ are linearly independent over F. In fact (Kaplansky, 1957) it vanishes if and only if $y_1, \ldots, y_m$ are linearly dependent over $F(t^p)$. If $D \neq 0$, then it is easy to obtain much sharper conclusions (cf. the case $k = 1$ of the next proposition).

For any integer $k \geq 1$, we set $F_k = F(t^{p^k})$. For any natural number $n = \sum s_i p_i$ written in base p, we set $\mathrm{mod}(m, p^k) = \sum_{i=0}^{k-1} s_i p^i$. For any nonzero polynomial $z = z(t) = c\prod(t - \alpha)^{n(\alpha)} \in F[t]$, where $0 \neq c \in F$, $\alpha \in \bar{F}$, $n(\alpha) = \mathrm{ord}_\alpha(z)$, and any integer $k \geq 1$, we define $\deg_k(z) = \sum_\alpha \mathrm{mod}(n(\alpha), p^k)$. It is a nondecreasing function of k and $\deg_k(z) = \deg(z)$ when $p^k \geq \deg(z)$. We denote by $\nu_k(z)$ the number of exponents $n(\alpha)$ which are not divisible by p^k. In particular, $\nu(z) = \nu_1(z)$.

Proposition 5.2. *Let $m \geq 2$ and $y_0 = y_1 + \cdots + y_m$ with $y_j \in F[t]$. Suppose that $\gcd(y_0, \ldots, y_m) = 1$. For some integer $k \geq 1$ assume that $y_1, \ldots, y_m$ are linearly independent over F_k. Then*

$$\begin{aligned}\deg(y_0) &\leq -m(m-1)/2 + \sum_{j=0}^{m} \deg_k(y_j)\\ &\leq -m(m-1)/2 + (p^k+1)\sum_{j=0}^{m} \nu_k(y_j).\end{aligned}$$

Moreover, in the case $k = 1$,

$$\deg(y_0) \leq -m(m-1)/2 + (m-1)\sum_{j=0}^{m} \nu(y_j).$$

Proof. The condition of linear independence implies that $p^k \geq m$. We write $y_j = y_j(t) = \sum_{s=0}^{p^k-1} y_{j,s}(t)^{p^k} t^s$ with $y_{j,s}(t) = y_{j,s} \in F[t]$. Such a representation is unique which is easy to see by induction on k. Also by induction on k it is easy to see that

$$\operatorname{ord}_\alpha(y_{j,s}) \geq (\operatorname{ord}_\alpha(y_j) - \operatorname{mod}(\operatorname{ord}_\alpha(y_j), p^k))/p^k$$

for all α, j, s. Now we consider the matrix $(y_{j,s})_{1 \leq j \leq m, 0 \leq s \leq p^k - 1}$. Its rank is m. We choose a square m by m submatrix with nonzero determinant $D_0 \in F[t]$. Since $p^k \deg(y_{j,s}) + s \leq \deg(y_j)$, we have $\deg(y_{j,s}) \leq (\deg(y_j) - s)/p^k$ and

$$\deg(D_0) \leq \Big(-m(m-1)/2 + \sum_{j=1}^{m} \deg(y_j)\Big)/p^k.$$

On the other hand,

$$\operatorname{ord}_\alpha(D_0) \geq \sum_{j=0}^{m} (\operatorname{ord}_\alpha(y_j) - \operatorname{mod}(\operatorname{ord}_\alpha(y_j), p^k))/p^k$$

for every $\alpha \in \bar{F}$ which is not a zero of y_0. Moreover this inequality also holds for every $\alpha \in \bar{F}$ which is a zero of y_0, because $D_0 = D_j$, where D_j is the determinant of the matrix obtained by replacing $y_{j,s}$ by $y_{0,s}$, and $0 \neq c_j \in F[t^{p^l}] \subset F[t^{p^k}]$ for all j.

Thus,

$$\sum_{j=0}^{m}(\deg(y_j) - \deg_k(y_j))/p^k$$
$$= \sum_{j=0}^{m}\sum_{\alpha\in\bar{F}}(\mathrm{ord}_\alpha(y_j) - \mathrm{mod}\,(\mathrm{ord}_\alpha(y_j), p^k))/p^k \leq \sum_{\alpha\in\bar{F}}\mathrm{ord}_\alpha D_0$$
$$= \deg(D_0) \leq (-m(m-1)/2 + \sum_{j=1}^{m}\deg\,(y_j))/p^k,$$

hence we obtain the first conclusion.

Now we consider the case $k=1$. Note that the linear independence implies that $m \leq p$ and that the Wronskian $D_0 = \det(y_j^{(i)})_{0\leq j\leq m, 1\leq i\leq m-1} \neq 0$. It is clear that the row j is divisible by $(t-\alpha)^n$with

$$\begin{aligned} n &= \mathrm{ord}_\alpha(y_j) - (\text{the last digit of } \mathrm{ord}_\alpha(y_j) \text{ in base } p) \\ &\leq \mathrm{ord}_\alpha(y_j) - p + 1 \leq \mathrm{ord}_\alpha(y_j) - m + 1. \end{aligned}$$

Since $c_0 D_0 = c_j D_j$, where D_j is the determinant obtained when we replace y_j by y_0, it follows that

$$\deg(D_0) \geq -(m-1)\nu(y_j) + \textstyle\sum_{j=0}^{m}\deg(y_j).$$

On the other hand,

$$\deg(D_0) \leq -m(m-1)/2 + \textstyle\sum_{j=1}^{m}\deg(y_j).$$

Comparing this upper bound for $\deg(D_0)$ with the above lower bound and taking into account that $\deg(y_j) = \deg_l(y_j)$, we obtain the conclusion.

Remarks. We do not pursue in this paper improvements of the proposition similar to Theorem 3.1.

We could prove the proposition using divided derivatives (cf. García and Voloch, 1987 and Wang, 1996).

If $y_1, \ldots, y_m$ are linearly independent over F then they are linearly independent over F_k for all sufficiently large k. However the proposition gives a meaningful upper bound on $\deg(y_0)$ only when k is small.

The proposition can be easily generalized to the situation when every linearly independent over F subsequence in $y_1, \ldots, y_m$ stays linearly independent over F_k. This is always true for sufficiently large k, but the larger the k the weaker the conclusion.

REFERENCES

Beukers, F. (1998). The Diophantine equation $Ax^p + By^q = Cz^r$. *Duke Math. J.* 91:61–88.

Beukers, F., Schlickewei, H.P. (1966). The equation $x + y = 1$ in finitely generated & groups. *Acta Arith.* 78:189–199.

Brownwell, D., Masser, D. (1986). Vanishing sums in function fields. *Math. Proc. Cambridge Philos. Soc.* 100:427–434.

Browkin, J., Brzeziński, J. (1994). Some remarks on the *abc*-conjecture. *Math. Comp.* 62:931–939.

Darmon, H., Granville, A. (1995). On the equations $z^m = F(x,y)$ and $Ax^p + By^q = Cz^r$. *Bull. London Math. Soc.* 27:513–543.

Elkies, N.D. (1988). On $A^4 + B^4 + C^4 = D^4$. *Math. Computation* 51:825–836.

García, A., Voloch, J.F. (1987). Wronskians and linear independence in fields of prime characteristic. *Manuscripta Math.* 59:457–469.

Kaplansky, I. (1957). *An Introduction to Differential Algebra.* Paris: Hermann.

Lander, L.J., Parkin, T.R. (1996). Counterexamples to Euler's conjecture on sums of like powers. *Bull. Amer. Math. Soc.* 72:1079.

Lang, S. (1999). *Math Talks for Undergraduates.* New York: Springer-Verlag.

Mason, R.C. (1986). Norm form equations. *I. J. of Number Theory* 22:190–207.

Mueller, J. (1993). The abc-Inequality and the generalized Fermat equation in function & fields. *Acta Arithmetica* LXIV.1:7–18.

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

Newman, D.J., Slater, M. (1979). Waring's problem for the ring of polynomials. *J. Number Theory* 11:477–487.

Paley, R.E.A.C. (1933). Theorems on polynomials in a Galois field. *Quar. J. Math.* 4:52–63.

Stepanov, S.A. (1982). Diophantine equations over function fields (in Russian). *Mat. Zametki* 32:753–764, 909.

Vaserstein, L. (1987). Waring's problems for algebras over fields. *J. Number Theory* 26:286–298.

Vaserstein, L. (1992). Ramsey's theorem and Waring's problem for algebras over fields. In: Proceedings of the Workshop at The Ohio State University, June, 17–26, 1991, Berlin, New York: Walter de Gruyter, pp. 435–441.

Voloch, J.F. (1985). Diagonal equations over function fields. *Bol. Soc. Brasil. Mat.* 16:29–39.

Voloch, J.F. (1998). The equation $ax + by = 1$ in characteristic p. *J. Number Theory* 73:195–200.

Wang, J.T.-Y. (1996). The truncated second main theorem of function fields. *J. Number Theory* 58:139–157.

Wooley, T.D. (1992). Large improvement in Waring's problem. *Ann. Math.* 135(1):131–164.

Zannier, U. (1993). Some remarks on the S-unit equation in function fields. *Acta Arith.* 64:87–98.

Received April 16, 2001