Divisibility of Exponential Sums via Elementary Methods

Francis N. Castro, Hugues Randriam, Ivelisse Rubio
and H. F. Mattson, Jr.

Abstract

We present an elementary method for evaluating the order of p-divisibility of exponential sums over a prime field. This method unifies and sometimes improves previously known results of Ax-Katz, Moreno-Moreno, Adolphson-Sperber, and Cao-Sun.

1 Introduction

E. Artin conjectured, and in 1935 C. Chevalley proved, that if $F(x_1, \ldots, x_n)$ is a homogeneous polynomial of total degree $d < n$ over a finite field, then F has a non-trivial zero. Thereafter several improvements to this result and extensions to the number of solutions of systems of polynomial equations of several variables over finite fields have been proved. These works fall into three categories:

- *Non-elementary*: Here are the extension of Ax's result by Katz in [6] and the Newton polyhedra method of Adolphson-Sperber presented in [1]. These proofs use p-adic theory of zeta functions and completely continuous endomorphisms in infinite dimensional p-adic Banach spaces.

- *Semi-elementary*: Here are Ax's result in [3], the extension of Katz's result proved by D. Wan in [13], the improvement of Ax-Katz presented in [8] by Moreno-Moreno, Moreno *et al.*'s tightness result proved in [11], Adolphson-Sperber's new proof of their result in [2]. These results use p-adic analysis combined with Stickelberger's theorem. Finally, Hou's

proof in [12], based on that of Ax, employs ingenious methods to prove Katz's extension of Ax's result.

- *Elementary*: Here are Chevalley's ([16]) and Warning's proofs ([17]) of the conjecture of Artin, the method of reduction to the ground field of Moreno-Moreno presented in [8], the covering method in characteristic 2 presented by Moreno-Moreno in [9], and Wan's proof of Moreno-Moreno's and Ax-Katz's results for prime fields in [14]. Wilson ([18]) gave another elementary proof of Ax-Katz's theorem for prime fields.

The purpose of this paper is to present elementary proofs of the non- and semi-elementary results mentioned above (with a slight variation in the case of Ax-Katz), and at the same time to extend or improve some of these results.

After preliminaries in Section 2, we present in Section 3 our main result: a new proof, entirely elementary in nature, of the prime field case of a theorem on divisibility properties of exponential sums previously obtained by Moreno *et al.* in [11]. The method of the proof can be seen as a generalization to arbitrary (positive) characteristic of the covering method introduced in [7, 9] for characteristic 2. The proof also includes a criterion for exact divisibility, from which tightness follows as in [11].

In Section 4, we derive some consequences of our main result:

- improvement on Adolphson-Sperber's theorem ([1, 2]) in the prime field case (this improvement is hinted at in [11])
- Wan's theorem on diagonal polynomials ([15]), in the prime field case
- Moreno-Moreno's theorem ([8]), which in many cases improves on Ax-Katz's theorem (both being equivalent in the prime field case).

We stress that our proofs are entirely elementary. Note in particular that in [3], Ax asked whether his theorem could be proved using elementary methods. Wan [14] and Wilson [18] did so for prime fields, although these two proofs (as well as Wan's proof of Moreno-Moreno) do not use the relation between exponential sums and the number of solutions of polynomial equations that Ax used. Thus it could be said that our proof gives an affirmative answer to Ax's question, while staying closer to his original strategy.

In Section 5 we improve bounds on the number of zeros of a certain family of polynomials first considered by Cao and Sun in [4].

2 Preliminaries

A covering method for the prime field of characteristic 2, presented in [9], established divisibility properties of an exponential sum for the number of zeros of a polynomial over $\mathbb{F}_2$. This method was used in [10] to give an elementary proof of the Moreno-Moreno result in [8] on the divisibility of the number of zeros of a set of polynomials, for a finite field of characteristic 2. We now generalize the covering method to characteristic p. From now on $\mathbb{F}_q$ denotes the field with q elements.

Let E be a finite subset of ${\mathbb{Z}_{\geq 0}}^n$. Choose a labeling of the elements of E, so that $E = \{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$, with $\mathbf{e}_j = (e_{1j}, \ldots, e_{nj})$, where each e_{ij} is a non-negative integer. By abuse of notation, we will identify E with the matrix

$$E = \begin{pmatrix} e_{11} & \cdots & e_{1N} \\ e_{21} & \cdots & e_{2N} \\ & \vdots & \\ e_{n1} & \cdots & e_{nN} \end{pmatrix}, \tag{1}$$

where the columns represent the $\mathbf{e}_j$'s; let R_i represent the ith row. We assume no column is repeated and no row is 0.

Let $\boldsymbol{\nu} = (\nu_1, \cdots, \nu_N)$ be an N-tuple of non-negative integers. We define the **zero-rank** of $\boldsymbol{\nu}$ with respect to E, denoted by $r_E(\boldsymbol{\nu})$, as the number of rows R_i of E such that $R_i \cdot \boldsymbol{\nu} = 0$.

Let now introduce the *m-covering problem* associated with E: if m is a positive integer, we say that $\boldsymbol{\nu}$ is an m**-covering** when $E\boldsymbol{\nu}^T = \nu_1\mathbf{e}_1 + \cdots + \nu_N\mathbf{e}_N$ has all its entries nonzero and divisible by m, that is, when there exist positive integers $\lambda_1, \ldots, \lambda_n$ such that, for each i,

$$\nu_1 e_{i1} + \cdots + \nu_N e_{iN} = R_i \cdot \boldsymbol{\nu} = m\lambda_i. \tag{2}$$

(We may also call such $\boldsymbol{\nu}$ a **positive solution to (2)** and a $\boldsymbol{\nu}$ that satisfies (2) with λs allowed to be 0 a **solution to (2)**. Thus a positive solution is a solution that has zero-rank $r_E(\boldsymbol{\nu})$ equal to 0.)

We define $\kappa_m(E)$, the m**-th covering number of** E, as the least cardinality of such an m-covering, *i.e.*, the least value of $\nu_1 + \cdots + \nu_N$ for which (2) holds. Thus a minimal positive solution $\boldsymbol{\nu}$ to (2) has **modulus** $|\boldsymbol{\nu}| = \sum_j \nu_j = \kappa_m(E)$.

The following lemma proves that $\kappa_m(E)$ is well-defined (*i.e.*, that positive solutions to (2) exist) and gives some of its elementary properties.

Lemma 2.1 *Let $E \subset {\mathbb{Z}_{\geq 0}}^n$ be as above. Then:*

(i) One has $\kappa_m(E) \leq mn$.

(ii) If $E \subset E'$ (or as matrices, if E' is constructed from E by adding extra columns), then $\kappa_m(E) \geq \kappa_m(E')$.

(iii) Let $n' \geq n$, and let $E' \subset {\mathbb{Z}_{\geq 0}}^{n'}$ be such that E is the one-to-one image of E' under the projection that forgets the last $n' - n$ coordinates (or as matrices, suppose E' is constructed from E by adding extra rows). Then one has $\kappa_m(E) \leq \kappa_m(E')$.

(iv) Consider a direct sum decomposition $\mathbb{Z}^n = \mathbb{Z}^{n_1} \oplus \cdots \oplus \mathbb{Z}^{n_r}$ with $n = n_1 + \cdots + n_r$, and suppose that relative to this decomposition E can be written as $E = E_1 \cup \cdots \cup E_r$ with each $E_i \subset \mathbb{Z}^{n_i}$ (or as matrices, suppose E is the block diagonal matrix constructed from the E_i). Then $\kappa_m(E) = \kappa_m(E_1) + \cdots + \kappa_m(E_r)$.

(v) If $\boldsymbol{\nu}$ is a (not necessarily positive) solution to (2) and if $r = r_E(\boldsymbol{\nu})$, then one may find an integer $t \leq r$, and indices $j_1 < \cdots < j_t$ such that $\nu_{j_1} = \cdots = \nu_{j_t} = 0$, and such that if one defines $\boldsymbol{\nu}'$ by $\nu'_j = \nu_j$ for all j except $\nu'_{j_1} = \cdots = \nu'_{j_t} = m$, then $\boldsymbol{\nu}'$ is a positive solution to (2).

(vi) If $\boldsymbol{\nu}$ is a (not necessarily positive) solution to (2), it satisfies

$$|\boldsymbol{\nu}| \geq \kappa_m(E) - m.r_E(\boldsymbol{\nu}).$$

Proof: Recall we supposed no row of E is zero. We then construct a positive solution $\boldsymbol{\nu}$ as follows: if $n > N$, we may choose all $\nu_j = m$, thus $\kappa_m(E) \leq mN$; if on the other hand $n \leq N$, then for each i choose an index j_i such that $e_{i,j_i} \neq 0$ (some j_i may be repeated, this will only diminish their number) and put $\nu_{j_i} = m$ for these, and $\nu_j = 0$ elsewhere. Thus $\kappa_m(E) \leq m(\min\{n, N\}) \leq mn$, which proves (i).

To prove (ii), label the elements of E' so that $E = \{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ and $E' = \{\mathbf{e}_1, \ldots, \mathbf{e}_{N'}\}$ with $N \leq N'$. Let $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_N)$ be a minimal m-covering of E, and $\boldsymbol{\nu}' = (\nu_1, \ldots, \nu_N, 0, \ldots, 0)$ with $N' - N$ zeros added. Then $\boldsymbol{\nu}'$ is an m-covering of E', so $\kappa_m(E') \leq |\boldsymbol{\nu}'| = |\boldsymbol{\nu}| = \kappa_m(E)$.

For (iii), remark that if $\boldsymbol{\nu}$ is a minimal m-covering of E', then it is also an m-covering of E, so $\kappa_m(E') = |\boldsymbol{\nu}| \geq \kappa_m(E)$.

To prove (iv), consider first $\boldsymbol{\nu}$ a minimal m-covering of E, and write $\boldsymbol{\nu} = (\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_r)$, its decomposition in the direct sum $\mathbb{Z}^n = \mathbb{Z}^{n_1} \oplus \cdots \oplus \mathbb{Z}^{n_r}$. Then each $\boldsymbol{\nu}_i$ is an m-covering for E_i, so $\kappa_m(E) = |\boldsymbol{\nu}| = |\boldsymbol{\nu}_1| + \cdots + |\boldsymbol{\nu}_r| \geq \kappa_m(E_1) + \cdots + \kappa_m(E_r)$. Conversely, for each i let $\boldsymbol{\nu}_i$ be a minimal m-covering for E_i, and let $\boldsymbol{\nu} = (\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_r)$. Then $\boldsymbol{\nu}$ is an m-covering of E, so $\kappa_m(E) \leq |\boldsymbol{\nu}| = |\boldsymbol{\nu}_1| + \cdots + |\boldsymbol{\nu}_r| = \kappa_m(E_1) + \cdots + \kappa_m(E_r)$.

We prove (v) by induction on r. If $r = 0$ there is nothing to prove. Suppose now $r > 0$ and (iii) proved up to $r - 1$. Then since $r > 0$, there exists an i_1 with $R_{i_1} \cdot \boldsymbol{\nu} = 0$, and there exists a j_1 with $e_{i_1 j_1} \neq 0$, so that necessarily $\nu_{j_1} = 0$. Defining $\hat{\boldsymbol{\nu}}$ by $\hat{\nu}_j = \nu_j$ for all j except $\hat{\nu}_{j_1} = m$, we now have $r_E(\hat{\boldsymbol{\nu}}) \leq r - 1$ and apply the induction hypothesis to this $\hat{\boldsymbol{\nu}}$.

We finally deduce (vi) from (v). Indeed, one then has $|\boldsymbol{\nu}'| = |\boldsymbol{\nu}| + mt \leq |\boldsymbol{\nu}| + mr$, while $|\boldsymbol{\nu}'| \geq \kappa_m(E)$ by definition of the covering number. □

Assertion (vi) in the lemma motivates the following:

Definition 2.2 *A (not necessarily positive) solution to the m-covering problem of E will be called* **optimal** *if it satisfies*

$$|\boldsymbol{\nu}| = \kappa_m(E) - m.r_E(\boldsymbol{\nu}). \tag{3}$$

We may view optimality as a generalization for not necessarily positive solutions of the notion of minimality for positive solutions. In particular, a positive solution is minimal if and only if it is optimal in this sense.

Lemma 2.3 *Let $\boldsymbol{\nu}$ be an optimal (not necessarily positive) solution to the m-covering problem of E. Then for all j, one has $\nu_j \leq m$.*

Proof: Let r, t and $\boldsymbol{\nu}'$ be as in Lemma 2.1 (v). Then $\boldsymbol{\nu}'$ is a positive solution to the covering problem, with $\kappa_m(E) \leq |\boldsymbol{\nu}'| = |\boldsymbol{\nu}| + mt \leq |\boldsymbol{\nu}| + mr = \kappa_m(E)$ (the last equality being because $\boldsymbol{\nu}$ is optimal). Thus these inequalities are equalities, so $t = r$; and more important, $\boldsymbol{\nu}'$ is a minimal covering. Since $\boldsymbol{\nu} \leq \boldsymbol{\nu}'$, it suffices to prove $\nu'_j \leq m$ for all j and for all minimal coverings $\boldsymbol{\nu}'$.

Now suppose there is a j_0 and a minimal positive solution $\boldsymbol{\nu}'$ to the covering problem with $\nu'_{j_0} \geq m+1$. Define another solution $\boldsymbol{\nu}''$ to the covering problem by $\nu''_j = \nu'_j$ for all j, except $\nu''_{j_0} = \nu'_{j_0} - m$ (we don't know yet that $\boldsymbol{\nu}''$ is a *positive* solution, but we will prove it very soon).

Remark that $\nu''_{j_0} \geq 1$, so that $(m+1)\nu''_{j_0} \geq \nu''_{j_0} + m = \nu'_{j_0}$, or $\nu''_{j_0} \geq \frac{1}{m+1}\nu'_{j_0}$; and for all the other j, $\nu''_j = \nu'_j$, so again $\nu''_j \geq \frac{1}{m+1}\nu'_j$. All in all, we find

$$\boldsymbol{\nu}'' \geq \frac{1}{m+1}\boldsymbol{\nu}';$$

thus for all i

$$R_i \cdot \boldsymbol{\nu}'' \geq \frac{1}{m+1} R_i \cdot \boldsymbol{\nu}' > 0.$$

So $\boldsymbol{\nu}''$ is a positive solution to the covering problem, with $|\boldsymbol{\nu}''| = |\boldsymbol{\nu}'| - m = \kappa_m(E) - m < \kappa_m(E)$, a contradiction. □

Let now p be a prime and, for integral $k \geq 0$, let $\sigma(k)$ denote the sum of the digits in the base-p expansion of k. That is, if $k = a_0 + a_1 p + a_2 p^2 + \cdots + a_r p^r$ with $0 \leq a_i < p$, then $\sigma(k) = a_0 + a_1 + \cdots + a_r$. Also, let $v_p(k)$ be the exponent on the highest power of p dividing k. It is known that

$$v_p(k!) = \frac{k - \sigma(k)}{p-1}, \tag{4}$$

a fact we'll use later.

For ease of writing we make these conventions: A relation ρ stated between integral vectors, as $\mathbf{a}\rho\mathbf{b}$, means that the relation holds in each coordinate. If stated between a vector and an integer, it means that each coordinate of the vector is in that relation to the integer. For example, $\mathbf{a} = (a_1, \ldots, b_n) \geq \mathbf{b} = (b_1, \ldots, b_n)$ means $a_i \geq b_i$ for $i = 1, \ldots, n$. And $\mathbf{a} \equiv 0 \pmod{m}$ means $a_i \equiv 0 \pmod{m}$ for all i. An exception: if we write $\mathbf{a} \neq \mathbf{b}$, we understand the usual meaning; we don't mean that $\mathbf{a}$ and $\mathbf{b}$ differ in every coordinate, only in at least one coordinate.

For $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{e} = (e_1, \ldots, e_n)$, and $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_N)$, we will also freely write $\mathbf{x}^{\mathbf{e}} = x_1^{e_1} \cdots x_n^{e_n}$, $\boldsymbol{\nu}! = \nu_1! \cdots \nu_N!$, and so on.

Let $\mathcal{S} = \{0, 1\}$ if $p = 2$, and, for $p \geq 3$, let g be a generator of the (cyclic) group of units of $\mathbb{Z}/p^m\mathbb{Z}$, and let $\mathcal{S} = \{0\} \cup \left\{ g^{ip^{m-1}} \mid 0 \leq i \leq p-2 \right\}$. Then the elements of $\mathcal{S}$ are a complete residue system modulo p. If k is a non-negative integer, then

$$\sum_{s\in\mathcal{S}} s^k \equiv \begin{cases} p \mod p^m & \text{if } k=0 \\ p-1 \mod p^m & \text{if } k \text{ is a nonzero multiple of } p-1 \\ \frac{g^{k(p-1)p^{m-1}}-1}{g^{kp^{m-1}}-1} \equiv 0 \mod p^m & \text{if } k \text{ is not divisible by } p-1. \end{cases} \tag{5}$$

In the next section we take m, which we are free to choose, to be n, the number of variables we shall consider.

3 Divisibility of Exponential Sums

In [11] appear tight bounds on the divisibility of some exponential sums and of the number of zeros of a set of polynomials. In this section we present, for prime fields, an elementary proof of the main theorem of [11].

Recall a few classical facts from algebraic number theory (or from the very beginning of the theory of cyclotomic fields): let ζ be a primitive p-th root of unity over $\mathbb{Q}$ and set $\theta = 1-\zeta$. Then in $\mathbb{Q}(\zeta)$ the (principal fractional) ideal $\langle\theta\rangle$ is prime, and the ideal $\langle p\rangle$ splits as $\langle p\rangle = \langle\theta\rangle^{p-1}$. So if we denote by v_p the extension to $\mathbb{Q}(\zeta)$ of the classical p-adic valuation, and also by v_θ the θ-adic valuation, one has $v_\theta(x) = (p-1)v_p(x)$ for all $x \in \mathbb{Q}(\zeta)$. Let then $A = \{x \in \mathbb{Q}(\zeta) \mid v_p(x) \geq 0\} = \{x \in \mathbb{Q}(\zeta) \mid v_\theta(x) \geq 0\}$ be the ring of p-integers (or equivalently θ-integers) in $\mathbb{Q}(\zeta)$.

Alternatively, remark that every $x \in \mathbb{Q}(\zeta) = \mathbb{Q}(\theta)$ can be written uniquely as $x = \sum_{k=0}^{p-2} \lambda_k \theta^k$ for some $\lambda_k \in \mathbb{Q}$. Write $\lambda_k = a_k/b_k$ as an irreducible fraction, with $a_k \in \mathbb{Z}$ and $b_k \in \mathbb{Z}_{>0}$. The non-zero $\lambda_k\theta^k$ in this sum have valuation $v_\theta(\lambda_k\theta^k) = k+(p-1)v_p(\lambda_k)$, which are pairwise distinct since they are distinct modulo $p-1$. Thus $v_\theta(x) = \min_{0\leq k\leq p-2}\{k+(p-1)v_p(\lambda_k)\}$, and a necessary and sufficient condition for $x \in A$ is that $v_p(\lambda_k) \geq 0$ for each k, or equivalently, that no b_k is multiple of p. This, together with the fact that θ satisfies the equation $\theta^{p-1} = \sum_{k=0}^{p-2}(-1)^{p-k}\binom{p}{k+1}\theta^k$, gives a very explicit description of A. In turn, depending on the reader's tastes and preferences, this explicit description could be taken as the definition of A (instead of the previous "abstract" one) for the rest of the paper.

This stated, one then has $pA = \theta^{p-1}A$, and there is also a natural identification $A/\theta A = \mathbb{F}_p$.

For any polynomial $F \in \mathbb{F}_p[\mathbf{x}]$, where $\mathbf{x} = (x_1, \ldots, x_n)$, we set

$$S(F) = \sum_{\mathbf{x} \in \mathbb{F}_p^n} \zeta^{F(\mathbf{x})} \in A. \tag{6}$$

By abuse of notation we will also write F for the polynomial with integral coefficients obtained by lifting $\mathbb{F}_p$ to $\mathcal{S}$. Since ζ^m depends only on m modulo p, the preceding can also be written as

$$S(F) = \sum_{\mathbf{x} \in \mathcal{S}^n} \zeta^{F(\mathbf{x})}. \tag{7}$$

Let now $E = \{\mathbf{e}_1, \ldots, \mathbf{e}_N\} \subset {\mathbb{Z}_{\geq 0}}^n$ be as in the preceding section, with associated matrix

$$E = \begin{pmatrix} e_{11} & \cdots & e_{1N} \\ e_{21} & \cdots & e_{2N} \\ & \vdots & \\ e_{n1} & \cdots & e_{nN} \end{pmatrix}. \tag{8}$$

Define

$$\mathbb{F}_p[\mathbf{x}]_E = \left\{ \sum_{j=1}^{N} a_j x_1^{e_{1j}} \cdots x_n^{e_{nj}} \mid a_1, \ldots, a_N \in \mathbb{F}_p \right\}.$$

This is a vector subspace of $\mathbb{F}_p[\mathbf{x}]$ of dimension N. The row R_i of E records the exponents of the variable x_i in a generic element $F(\mathbf{x}) \in \mathbb{F}_p[\mathbf{x}]_E$, so the assumption that no row is zero means that every variable does indeed occur. On the other hand, the column $\mathbf{e}_j$ of E records the exponents of the j-th monomial in $F(\mathbf{x})$ (at least when $a_j \neq 0$).

Conversely, for arbitrary $F \in \mathbb{F}_p[\mathbf{x}]$, we may define its exponent set $e(F) \subset {\mathbb{Z}_{\geq 0}}^n$ as the set of exponent n-tuples of the monomials that appear in F with non-zero coefficient, so that $F(\mathbf{x}) = \sum_{\mathbf{e} \in e(F)} a_{\mathbf{e}} \mathbf{x}^{\mathbf{e}}$ with $a_{\mathbf{e}} \in \mathbb{F}_p^{\times}$. One then has $F \in \mathbb{F}_p[\mathbf{x}]_E$ if and only if $e(F) \subset E$.

Theorem 3.1 *(i) With the above notations, every $F \in \mathbb{F}_p[\mathbf{x}]_E$ satisfies*

$$v_p\left(S(F)\right) \geq \frac{\kappa_{p-1}(E)}{p-1}.$$

In particular, if $S(F)$ is a rational integer, it is divisible by $p^{\lceil \frac{\kappa_{p-1}(E)}{(p-1)} \rceil}$.

(ii) Conversely, there exists an $F \in \mathbb{F}_p[\mathbf{x}]_E$ such that the preceding inequality is an equality, that is

$$v_p\left(S(F)\right) = \frac{\kappa_{p-1}(E)}{p-1}.$$

Before we proceed to the proof, we make a few remarks.

Remark 3.2 *Part (i) of the theorem can be easily generalized when rows of E are allowed to be zero, since in the sum defining $S(F)$, summing over a variable that does not occur only factors out a constant p while leaving the rest of the sum unchanged.*

Remark 3.3 *Given an $F \in \mathbb{F}_p[\mathbf{x}]$ in which all variables occur, one can use part (i) of the theorem with $E = e(F)$ to get*

$$v_p\left(S(F)\right) \geq \frac{\kappa_{p-1}(e(F))}{p-1}$$

(it may then be convenient to write $\kappa_{p-1}(F)$ for $\kappa_{p-1}(e(F))$). Because of Lemma 2.1 (ii), choosing a larger E will not give a stronger inequality.

In general, it could be computationally expensive to compute $\kappa_{p-1}(E)$; but in Section 4 we will see cases where this computation, hence estimation of divisibility, is easy.

Remark 3.4 *The polynomial F we obtain in part (ii) of the theorem need not satisfy $e(F) = E$, that is, it may be that $e(F)$ is a strict subset of E, or equivalently, that F (when decomposed in the monomial basis of $\mathbb{F}_p[\mathbf{x}]_E$) has some coefficients equal to 0. However, combining Lemma 2.1 (ii) and the preceding remark, we find that its exponent set must at least satisfy $\kappa_{p-1}(e(F)) = \kappa_{p-1}(E)$.*

We now begin the proof of the theorem, which will use several lemmas. Write

$$S(F) = \sum_{\mathbf{x}\in\mathcal{S}^n} (1-\theta)^{F(\mathbf{x})} = \sum_{\mathbf{x}\in\mathcal{S}^n} \prod_{j=1}^{N} (1-\theta)^{a_j x_1^{e_{1j}} \cdots x_n^{e_{nj}}}. \tag{9}$$

Let $M_j = a_j x_1^{e_{1j}} \cdots x_n^{e_{nj}}$. Then let M be any upper bound on all the M_j's as $\mathbf{x}$ varies, e.g., $M = (p^n - 1)^{d+1}$, where d is the total degree of F (we could also take $M = +\infty$, since all the following sums are finite anyway):

$$
\begin{aligned}
S(F) &= \sum_{\mathbf{x}\in\mathcal{S}^n} \prod_{j=1}^{N} (1-\theta)^{M_j} \\
&= \sum_{\mathbf{x}\in\mathcal{S}^n} \sum_{\nu_1=0}^{M} \binom{M_1}{\nu_1} (-\theta)^{\nu_1} \times \cdots \times \sum_{\nu_N=0}^{M} \binom{M_N}{\nu_N} (-\theta)^{\nu_N} \\
&= \sum_{\nu_1=0}^{M} \cdots \sum_{\nu_N=0}^{M} \sum_{\mathbf{x}\in\mathcal{S}^n} (-\theta)^{\sum \nu_j} \binom{M_1}{\nu_1} \cdots \binom{M_N}{\nu_N},
\end{aligned}
$$

or

$$
S(F) = \sum_{\boldsymbol{\nu}=0}^{M} T_{\boldsymbol{\nu}}(F) \tag{10}
$$

with

$$
T_{\boldsymbol{\nu}}(F) = (-\theta)^{|\boldsymbol{\nu}|} \sum_{\mathbf{x}\in\mathcal{S}^n} \binom{M_1}{\nu_1} \cdots \binom{M_N}{\nu_N}. \tag{11}
$$

We first give a rough estimate of the valuation of such a $T_{\boldsymbol{\nu}}(F)$:

Lemma 3.5 *Let $r = r_E(\boldsymbol{\nu})$ be the number of rows of E orthogonal to $\boldsymbol{\nu}$. Then*

$$
v_\theta(T_{\boldsymbol{\nu}}(F)) \geq |\boldsymbol{\nu}| + (p-1)r.
$$

Proof: The Lemma follows from the fact that the r variables associated with the rows orthogonal to $\boldsymbol{\nu}$ do not appear in the sum defining $T_{\boldsymbol{\nu}}(F)$. Why so? Because an "orthogonal" row occurs only when every non-0 ν_j is met by an e_{ij} that is 0. Thus summing over each of these variables will factor out a constant p, leaving the sum over the other variables untouched. □

Let us define $\boldsymbol{\delta} = (\delta_1, \cdots, \delta_N)$ as follows:

$$
\delta_j = \begin{cases} 1 & \text{if } \nu_j > 0 \\ 0 & \text{if } \nu_j = 0. \end{cases}
$$

Lemma 3.6 *For all $\boldsymbol{\ell} = (l_1, \ldots, l_N)$ satisfying $\boldsymbol{\delta} \leq \boldsymbol{\ell} \leq \boldsymbol{\nu}$, one has*

$$
r_E(\boldsymbol{\ell}) = r_E(\boldsymbol{\nu}).
$$

Proof: One has $\boldsymbol{\nu} \leq M$, so by construction $\frac{1}{M}\boldsymbol{\nu} \leq \boldsymbol{\delta}$, and so

$$\frac{1}{M}\boldsymbol{\nu} \leq \boldsymbol{\ell} \leq \boldsymbol{\nu}.$$

Since each row R_i of E has non-negative entries, it follows that

$$\frac{1}{M} R_i \cdot \boldsymbol{\nu} \leq R_i \cdot \boldsymbol{\ell} \leq R_i \cdot \boldsymbol{\nu},$$

and the only possibility for one of the terms in this inequality to be zero is that they all are zero. □

We may write $\nu_j!\binom{M_j}{\nu_j} = M_j(M_j-1)(M_j-2)\cdots(M_j-(\nu_j-1))$, the latter being a polynomial in M_j of degree ν_j, with constant term 0 unless $\nu_j = 0$. More precisely, $s(\nu_j, l_j) \in \mathbb{Z}$ denoting the corresponding Stirling number of the first kind, one has

$$\nu_j!\binom{M_j}{\nu_j} = \sum_{\delta_j \leq l_j \leq \nu_j} s(\nu_j, l_j) M_j^{l_j}. \tag{12}$$

Thus

$$\nu_1! \cdots \nu_N! \binom{M_1}{\nu_1} \cdots \binom{M_N}{\nu_N} = \prod_{j=1}^{N} \sum_{\delta_j \leq l_j \leq \nu_j} s(\nu_j, l_j) M_j^{l_j} \tag{13}$$

$$= \sum_{\boldsymbol{\delta} \leq \boldsymbol{\ell} \leq \boldsymbol{\nu}} s(\nu_1, l_1) \left(a_1 x_1^{e_{11}} \cdots x_n^{e_{n1}}\right)^{l_1} \cdots s(\nu_N, l_N) \left(a_N x_1^{e_{1N}} \cdots x_n^{e_{nN}}\right)^{l_N} \tag{14}$$

$$= \sum_{\boldsymbol{\delta} \leq \boldsymbol{\ell} \leq \boldsymbol{\nu}} s_{\boldsymbol{\nu},\boldsymbol{\ell}} \mathbf{a}^{\boldsymbol{\ell}} x_1^{R_1 \cdot \boldsymbol{\ell}} \cdots x_n^{R_n \cdot \boldsymbol{\ell}} \tag{15}$$

where $\mathbf{a}^{\boldsymbol{\ell}} = a_1^{l_1} \cdots a_N^{l_N}$ and $s_{\boldsymbol{\nu},\boldsymbol{\ell}} = \prod_j s(\nu_j, l_j)$. Therefore,

$$T_{\boldsymbol{\nu}}(F) = \frac{(-\theta)^{|\boldsymbol{\nu}|}}{\boldsymbol{\nu}!} \sum_{\mathbf{x} \in \mathcal{S}^n} \sum_{\boldsymbol{\delta} \leq \boldsymbol{\ell} \leq \boldsymbol{\nu}} s_{\boldsymbol{\nu},\boldsymbol{\ell}} \mathbf{a}^{\boldsymbol{\ell}} x_1^{R_1 \cdot \boldsymbol{\ell}} \cdots x_n^{R_n \cdot \boldsymbol{\ell}} \tag{16}$$

with $\boldsymbol{\nu}!$ defined as $\nu_1! \cdots \nu_N!$. Now setting

$$L_{\boldsymbol{\nu},\boldsymbol{\ell}} = \frac{(-\theta)^{|\boldsymbol{\nu}|}}{\boldsymbol{\nu}!} \sum_{\mathbf{x} \in \mathcal{S}^n} x_1^{R_1 \cdot \boldsymbol{\ell}} \cdots x_n^{R_n \cdot \boldsymbol{\ell}} \tag{17}$$

we arrive at

$$T_{\boldsymbol{\nu}}(F) = \sum_{\boldsymbol{\delta} \leq \boldsymbol{\ell} \leq \boldsymbol{\nu}} L_{\boldsymbol{\nu},\boldsymbol{\ell}} s_{\boldsymbol{\nu},\boldsymbol{\ell}} \mathbf{a}^{\boldsymbol{\ell}}. \tag{18}$$

Note that, in this equation, neither $s_{\boldsymbol{\nu},\boldsymbol{\ell}}$ nor $L_{\boldsymbol{\nu},\boldsymbol{\ell}}$ depend on $\mathbf{a}$ (the coefficients of F); $s_{\boldsymbol{\nu},\boldsymbol{\ell}}$ is a rational integer, while for the moment $L_{\boldsymbol{\nu},\boldsymbol{\ell}}$ is only known to be an element of $\mathbb{Q}(\theta)$.

We now partition the $\boldsymbol{\nu}$s ($0 \leq \boldsymbol{\nu} \leq M$) into two cells:

- $V_{opt} = \{\boldsymbol{\nu} \mid \boldsymbol{\nu}$ is an optimal (not necessarily positive) solution to the $p-1$-covering system of $E\}$ (optimality being in the sense of Definition 2.2);
- V_{other} is the set of all other $\boldsymbol{\nu}$s. That is, $\boldsymbol{\nu}$ may be a non-optimal solution to the system, or $\boldsymbol{\nu}$ may not be a solution at all.

We may thus write

$$S(F) = \sum_{\boldsymbol{\nu} \in V_{opt}} T_{\boldsymbol{\nu}}(F) \; + \sum_{\boldsymbol{\nu} \in V_{other}} T_{\boldsymbol{\nu}}(F). \tag{19}$$

Lemma 3.7 *If $\boldsymbol{\nu} \in V_{other}$, then*

$$T_{\boldsymbol{\nu}}(F) \in \theta^{\kappa_{p-1}(E)+1} A.$$

Proof: We distinguish two cases among all the $\boldsymbol{\ell}$ satisfying $\boldsymbol{\delta} \leq \boldsymbol{\ell} \leq \boldsymbol{\nu}$: there is (or is not) a not necessarily positive solution $\boldsymbol{\ell}$ to the $(p-1)$-covering system.

Case 1. There exists an $\boldsymbol{\ell}$, solution to the $p-1$-covering system. By Lemma 2.1 (vi), one has $|\boldsymbol{\ell}| \geq \kappa_{p-1}(E) - (p-1) r_E(\boldsymbol{\ell})$, with $r_E(\boldsymbol{\ell}) = r_E(\boldsymbol{\nu})$ by Lemma 3.6. So

$$|\boldsymbol{\ell}| \geq \kappa_{p-1}(E) - (p-1) r_E(\boldsymbol{\nu}). \tag{20}$$

Now there are two options:

- If $\boldsymbol{\ell} = \boldsymbol{\nu}$, then $\boldsymbol{\nu}$ is a solution to the covering system, but then $\boldsymbol{\nu}$ is not optimal since by assumption $\boldsymbol{\nu} \in V_{other}$; so

$$|\boldsymbol{\nu}| > \kappa_{p-1}(E) - (p-1) r_E(\boldsymbol{\nu}).$$

- If $\boldsymbol{\ell} \neq \boldsymbol{\nu}$, then $|\boldsymbol{\nu}| > |\boldsymbol{\ell}|$. This together with (20) gives again

$$|\boldsymbol{\nu}| > \kappa_{p-1}(E) - (p-1) r_E(\boldsymbol{\nu}).$$

So whatever the option chosen, we have $|\boldsymbol{\nu}| > \kappa_{p-1}(E) - (p-1)r_E(\boldsymbol{\nu})$, and the conclusion follows from Lemma 3.5. This finishes the proof in case 1.

Case 2. No $\boldsymbol{\ell}$ is a solution to the $(p-1)$-covering system. This means that for each $\boldsymbol{\ell}$ there is a row $R_{i_{\boldsymbol{\ell}}}$ of E with $R_{i_{\boldsymbol{\ell}}} \cdot \boldsymbol{\ell} \not\equiv 0 \pmod{p-1}$. Here (5) tells us that $p^n | \sum_{x_{i_{\boldsymbol{\ell}}} \in \mathcal{S}} x_{i_{\boldsymbol{\ell}}}^{R_{i_{\boldsymbol{\ell}}} \cdot \boldsymbol{\ell}}$, hence $p^n | \sum_{\mathbf{x} \in \mathcal{S}^n} x_1^{R_1 \cdot \boldsymbol{\ell}} \cdots x_n^{R_n \cdot \boldsymbol{\ell}}$. Thus

$$\begin{aligned} v_\theta(L_{\boldsymbol{\nu},\boldsymbol{\ell}}) &\geq |\boldsymbol{\nu}| - (p-1)v_p(\boldsymbol{\nu}!) + (p-1)n \\ &\geq \sum \sigma(\nu_j) \, + \, \kappa_{p-1}(E) \end{aligned}$$

where we used (4) and Lemma 2.1 (i).

To finish the proof, remark that $\boldsymbol{\nu} \neq 0$ (otherwise $\boldsymbol{\ell} = 0$ would be forced, putting us in case 1), so $\sum \sigma(\nu_j) > 0$, and

$$v_\theta(T_{\boldsymbol{\nu}}(F)) \geq \min_{\boldsymbol{\ell}} v_\theta(L_{\boldsymbol{\nu},\boldsymbol{\ell}}) > \kappa_{p-1}(E).$$

□

From this Lemma and (19), we deduce

$$S(F) \equiv \sum_{\boldsymbol{\nu} \in V_{opt}} T_{\boldsymbol{\nu}}(F) \pmod{\theta^{\kappa_{p-1}(E)+1} A}. \tag{21}$$

Now:

Lemma 3.8 *If $\boldsymbol{\nu} \in V_{opt}$, then for all $\boldsymbol{\ell}$ (with $\boldsymbol{\delta} \leq \boldsymbol{\ell} \leq \boldsymbol{\nu}$) one has*

$$L_{\boldsymbol{\nu},\boldsymbol{\ell}} \in \theta^{\kappa_{p-1}(E)} A.$$

Proof: Let $\boldsymbol{\nu} \in V_{opt}$, and let $r = r_E(\boldsymbol{\nu})$ be the number of rows of E orthogonal to $\boldsymbol{\nu}$; for ease of writing, reorder the variables so that one can write $R_1 \cdot \boldsymbol{\nu} = \cdots = R_r \cdot \boldsymbol{\nu} = 0$. Then, since $\boldsymbol{\delta} \leq \boldsymbol{\ell} \leq \boldsymbol{\nu}$, one has $0 \leq R_i \cdot \boldsymbol{\ell} \leq R_i \cdot \boldsymbol{\nu}$, and thus $R_1 \cdot \boldsymbol{\ell} = \cdots = R_r \cdot \boldsymbol{\ell} = 0$. This means that the variables $x_1, \ldots, x_r$ do not appear in the sum defining $L_{\boldsymbol{\nu},\boldsymbol{\ell}}$, so that

$$L_{\boldsymbol{\nu},\boldsymbol{\ell}} = p^r \frac{(-\theta)^{|\boldsymbol{\nu}|}}{\boldsymbol{\nu}!} \sum_{\mathbf{x}' \in \mathcal{S}^{n-r}} x_{r+1}^{R_{r+1} \cdot \boldsymbol{\ell}} \cdots x_n^{R_n \cdot \boldsymbol{\ell}}, \tag{22}$$

in which $\mathbf{x}'$ stands for $(x_{r+1}, \ldots, x_n)$. Now by Lemma 2.3, one has $\boldsymbol{\nu} \leq p-1$, so that $v_p(\boldsymbol{\nu}!) = 0$, and

$$v_\theta(L_{\boldsymbol{\nu},\boldsymbol{\ell}}) \geq (p-1)r + |\boldsymbol{\nu}| \geq \kappa_{p-1}(E),$$

the last inequality stemming from Lemma 2.1 (vi). This proves Lemma 3.8. □

Part (i) of the theorem now follows from this lemma, along with (18) and (21).

We now prove part (ii) of the theorem. Using (18) again, write

$$\sum_{\boldsymbol{\nu}\in V_{opt}} T_{\boldsymbol{\nu}}(F) = \sum_{\boldsymbol{\nu}\in V_{opt}} \sum_{\boldsymbol{\delta}\le\boldsymbol{\ell}\le\boldsymbol{\nu}} L_{\boldsymbol{\nu},\boldsymbol{\ell}} s_{\boldsymbol{\nu},\boldsymbol{\ell}} \mathbf{a}^{\boldsymbol{\ell}} = \theta^{\kappa_{p-1}(E)} P(\mathbf{a}) \tag{23}$$

with

$$P(\mathbf{a}) = \sum_{\boldsymbol{\nu}\in V_{opt}} \sum_{\boldsymbol{\delta}\le\boldsymbol{\ell}\le\boldsymbol{\nu}} (\theta^{-\kappa_{p-1}(E)} L_{\boldsymbol{\nu},\boldsymbol{\ell}}) s_{\boldsymbol{\nu},\boldsymbol{\ell}} \mathbf{a}^{\boldsymbol{\ell}}.$$

Thanks to the last lemma, we have $P(\mathbf{a}) \in A[\mathbf{a}]$, so we can reduce its coefficients modulo θ to get a polynomial

$$P^\sharp(\mathbf{a}) \in \mathbb{F}_p[\mathbf{a}].$$

Putting equations (21) and (23) together, we immediately get the following criterion for exact divisibility:

Proposition 3.9 *With these notations, if the value of $P^\sharp$ at $\mathbf{a}$ is non-zero (where $\mathbf{a}$ are the coefficients of F), then $S(F)$ is divisible by $\theta^{\kappa_{p-1}(E)}$ but not by $\theta^{\kappa_{p-1}(E)+1}$.*

The tightness part in the theorem now follows from this criterion. Indeed, let $\boldsymbol{\mu}$ be any minimal positive solution to the $(p-1)$-covering problem of E. Thus $\boldsymbol{\mu} \in V_{opt}$ and $|\boldsymbol{\mu}| = \kappa_{p-1}(E)$. If $\boldsymbol{\nu} \in V_{opt}$, then by definition one has $|\boldsymbol{\nu}| = \kappa_{p-1}(E) - (p-1) r_E(\boldsymbol{\nu}) \le \kappa_{p-1}(E) = |\boldsymbol{\mu}|$, so if $\boldsymbol{\nu} \ne \boldsymbol{\mu}$, then for all $\boldsymbol{\ell} \le \boldsymbol{\nu}$ one also has $\boldsymbol{\ell} \ne \boldsymbol{\mu}$. Thus the coefficient $d_{\boldsymbol{\mu}}$ of $\mathbf{a}^{\boldsymbol{\mu}}$ in P is

$$\begin{aligned} d_{\boldsymbol{\mu}} &= (\theta^{-\kappa_{p-1}(E)} L_{\boldsymbol{\mu},\boldsymbol{\mu}}) s_{\boldsymbol{\mu},\boldsymbol{\mu}} \\ &= (-1)^{|\boldsymbol{\mu}|} s_{\boldsymbol{\mu},\boldsymbol{\mu}} \frac{\theta^{|\boldsymbol{\mu}|-\kappa_{p-1}(E)}}{\boldsymbol{\mu}!} \sum_{\mathbf{x}\in\mathcal{S}^n} x_1^{R_1\cdot\boldsymbol{\mu}} \cdots x_n^{R_n\cdot\boldsymbol{\mu}} \end{aligned}$$

where we used (17). We compute the valuation of $d_{\boldsymbol{\mu}}$ as follows:

- $s_{\boldsymbol{\mu},\boldsymbol{\mu}} = \prod_j s(\mu_j, \mu_j) = 1$ has valuation 0;

- $\theta^{|\boldsymbol{\mu}|-\kappa_{p-1}(E)}$ has valuation $|\boldsymbol{\mu}|-\kappa_{p-1}(E) = 0$ since $\boldsymbol{\mu}$ is a minimal positive solution;
- $\boldsymbol{\mu}!$ has valuation 0, since $\boldsymbol{\mu} \leq p-1$ (Lemma 2.3);
- $\sum_{\mathbf{x}\in\mathcal{S}^n} x_1^{R_1\cdot\boldsymbol{\mu}} \cdots x_n^{R_n\cdot\boldsymbol{\mu}}$ has valuation 0, thanks to (5), since each $R_i \cdot \boldsymbol{\mu}$ is a non-zero multiple of $p-1$.

So all in all, one has $v_\theta(d_{\boldsymbol{\mu}}) = 0$.

Thanks to Lemma 2.3, each $\boldsymbol{\nu} \in V_{opt}$ satisfies $\boldsymbol{\nu} \leq p-1$, so P, and thus also $P^\sharp$, has degree less than p in each a_j. Lastly, $v_\theta(d_{\boldsymbol{\mu}}) = 0$ implies $P^\sharp$ is not the zero polynomial.

We now use the following well-known lemma, whose proof we omit:

Lemma 3.10 *Let $P^\sharp$ be a polynomial over $\mathbb{F}_p$ of degree $\leq p-1$ in each of its N variables $a_1, \ldots, a_N$. If $P^\sharp$ is not the zero polynomial, then there exists $(\alpha_1, \ldots, \alpha_N) \in \mathbb{F}_p^N$ such that $P^\sharp(\alpha_1, \ldots, \alpha_N) \neq 0$ in $\mathbb{F}_p$.*

Letting $F(\mathbf{x}) = \sum_j \alpha_j \mathbf{x}^{\mathbf{e}_j}$, we can now apply the criterion in proposition 3.9. This finishes the proof of the theorem.

□

In [5] Castro *et al.* generalized a result of Carlitz by determining the exact divisibility of the exponential sum associated to certain polynomials. To obtain their result they used an argument that relies on a non-elementary result of Stickelberger. Using Proposition 3.9 we could obtain the same result over the prime field without using Stickelberger's result.

Before we proceed to applications, we restate in terms of polynomials some parts of Lemma 2.1 that will be used in the following sections in combination with Theorem 3.1.

Lemma 3.11 *(i) Let $F(\mathbf{x})$ be a polynomial in a certain set of variables $\mathbf{x}$, and let y be a new variable. If $\widehat{F}(y, \mathbf{x}) = yF(\mathbf{x})$, then $\kappa_{p-1}(\widehat{F}) \geq \kappa_{p-1}(F)$.*

(ii) Let $\mathbf{x}_1, \ldots, \mathbf{x}_r$ be disjoint sets of variables, and suppose F can be written as $F(\mathbf{x}) = F_1(\mathbf{x}_1) + \cdots + F_r(\mathbf{x}_r)$. Then $\kappa_{p-1}(F) = \kappa_{p-1}(F_1) + \cdots + \kappa_{p-1}(F_r)$.

Proof: Just looking at exponent sets, (i) arises directly from Lemma 2.1 (iii), and (ii) from Lemma 2.1 (iv). □

4 New Elementary Proofs of Classic Results

Armed with Theorem 3.1, we now revisit some classical results, which we can easily obtain by computing or estimating the $(p-1)$-covering $\kappa_{p-1}(E)$. We will first see that the bound improves the result of Adolphson and Sperber [1].

We need some notation to state their theorem. Suppose $F(x_1,\ldots,x_n) = \sum_{j=1}^{N} a_j x_1^{e_{1j}}\cdots x_n^{e_{nj}}$, $(a_j \neq 0)$. Let $\Delta(F)$ be the Newton polyhedron of F, that is, the convex hull in $\mathbb{R}^n$ of the set $\{\mathbf{e}_j\}\cup\{(0,\ldots,0)\}$. Let $\omega(F)$ be the smallest positive rational number such that $\omega(F)\Delta(F)$ contains a point of $\mathbb{Z}^n_{>0}$. A. Adolphson and S. Sperber [1] proved that if F is not a polynomial in some proper subset of the variables $x_1,\ldots,x_n$, then $v_p(S(F)) \geq \omega(F)$. The following theorem says that the bound presented in Theorem 3.1 could be better than $\omega(F)$.

Theorem 4.1 *Let $E = \{\mathbf{e}_1,\ldots,\mathbf{e}_N\}$ and $\omega(F)$ be defined as above. Then*

$$\frac{\kappa_{p-1}(E)}{p-1} \geq \omega(F).$$

Proof: By definition, $\kappa_{p-1}(E)$ is the least sum $\nu_1+\cdots+\nu_N$ over all vectors $\boldsymbol{\nu}$ satisfying $\nu_1\mathbf{e}_1+\cdots+\nu_N\mathbf{e}_N = (\lambda_1(p-1),\ldots,\lambda_n(p-1))$ with positive $\lambda_1,\ldots,\lambda_n$. Dividing the last equation by $p-1$, we obtain

$$\frac{\nu_1}{p-1}\mathbf{e}_1+\cdots+\frac{\nu_N}{p-1}\mathbf{e}_N = (\lambda_1,\ldots,\lambda_n)\in\mathbb{Z}^n_{>0}.$$

Let $k := \sum_{j=1}^{N}\frac{\nu_j}{p-1}$. Then $\frac{1}{k}\sum_{j=1}^{N}\frac{\nu_j}{p-1} = 1$ and $k\sum_{j=1}^{N}\frac{\nu_j}{k(p-1)}\mathbf{e}_j = (\lambda_1,\ldots,\lambda_n)$, a positive n-tuple in $k\Delta(F)$. Thus $\frac{\kappa_{p-1}(E)}{p-1} = \sum_{j=1}^{N}\frac{\nu_j}{p-1} = k \geq \omega(F)$.

□

The next example shows that there are cases where the bound in Theorem 4.1 is stronger than that of Adolphson and Sperber.

Example 4.2 *Let $d \neq 1$ be relatively prime to $p-1$, and let $F(x,y) = a_1x^{d^2} + a_2y^{d^2} + a_3x^dy^d$ be a polynomial over $\mathbb{F}_p$. It can be verified that $\omega(F) = \frac{2}{d^2}$. To compute $\kappa_{p-1}(E)$ we consider all ν_1,ν_2,ν_3 satisfying*

$$d^2\nu_1 + d\nu_3 = \lambda_1(p-1)$$
$$d^2\nu_2 + d\nu_3 = \lambda_2(p-1),$$

with positive λ_1, λ_2. *Since* $\gcd(d, p-1) = 1$, $d|\lambda_1$ *and* $d|\lambda_2$. *Therefore,*

$$\frac{\nu_1 + \nu_2 + \nu_3}{p-1} \geq \frac{2}{d}.$$

Thus $\frac{\kappa_{p-1}(E)}{p-1} \geq \frac{2}{d} > \frac{2}{d^2} = \omega(F)$.

The relation between an exponential sum $S(F) = \sum_{\mathbf{x}\in\mathbb{F}_p^n} \zeta^{F(\mathbf{x})}$ and the number of zeros of a system of polynomials $P_1(\mathbf{x}), \ldots, P_t(\mathbf{x})$ is given by the following well known Lemma.

Lemma 4.3 *Let* ζ *be as in (6),* $P_1(\mathbf{x}), \ldots, P_t(\mathbf{x}) \in \mathbb{F}_p[x_1, \ldots, x_n]$, *and* N *be the number of common zeros of* $P_1, \ldots, P_t$. *Then*

$$N = p^{-t} \sum_{\mathbf{x}\in\mathbb{F}_p^n, \mathbf{y}\in\mathbb{F}_p^t} \zeta^{y_1P_1(\mathbf{x})+\ldots+y_tP_t(\mathbf{x})}.$$

Now Theorem 3.1 gives an elementary proof of the following theorem, which was proved for a general finite field $\mathbb{F}_{p^f}$ in [11].

Theorem 4.4 *Let* $P_1(\mathbf{x}), \ldots, P_t(\mathbf{x}) \in \mathbb{F}_p[x_1, \ldots, x_n]$, *and* N *be the number of common zeros of* $P_1, \ldots, P_t$. *Introduce* t *extra variables* $y_1, \ldots, y_t$ *and define a new polynomial* F *in* $n+t$ *variables by* $F(\mathbf{x}, \mathbf{y}) = y_1P_1(\mathbf{x}) + \ldots + y_tP_t(\mathbf{x})$. *Then,* $p^{\frac{\kappa_{p-1}(F)}{p-1}-t}$ *divides* N, *and this divisibility is tight.*

Remark 4.5 *Note that, in this case, to compute* $\kappa_{p-1}(F)$, *we need to include equations associated to each of the variables* $y_1, \ldots, y_t$ *(as shown in (24)). When we sum these* t *equations we get an expression for the modulus* $|\boldsymbol{\nu}|$ *of the* $(p-1)$*-covering that is a multiple of* $p-1$. *Hence,* $\frac{\kappa_{p-1}(F)}{p-1}$ *is always an integer.*

In [15], Wan obtained an improvement of Ax's theorem for diagonal equations. Our method gives an elementary proof for prime fields of Wan's result.

Theorem 4.6 (Wan) *Let* $F(x_1, \ldots, x_n) = a_1 x_1^{d_1} + \cdots + a_n x_n^{d_n} + \beta$ *be a polynomial over* $\mathbb{F}_p$ *and let* N *be the number of zeros of* F *over* $\mathbb{F}_p$*, Then* p^μ *divides* N*, where*

$$\mu = \left\lceil \frac{1}{\gcd(p-1, d_1)} + \cdots + \frac{1}{\gcd(p-1, d_n)} \right\rceil - 1.$$

In fact Wan stated this theorem with the condition that all d_i divide $p-1$, in which case this becomes

$$\mu = \left\lceil \frac{1}{d_1} + \cdots + \frac{1}{d_n} \right\rceil - 1.$$

His version might seem to be a special case; however, the two formulations are easily seen to be equivalent, so ours does not add any generality. We stick to this "artificially general" formulation since this does not make the proof more difficult.

Proof: This is a special case of Theorem 5.2 with the polynomial $G = \beta$ and hence $m = 0$. □

Another classic result for which we can give a new elementary proof is Moreno-Moreno's Theorem ([8]). This new proof is easily obtained from an estimate of the $(p-1)$-covering $\kappa_{p-1}(E)$:

Recall that for an integer $k \geq 0$, $\sigma(k)$ denotes the sum of the digits in the base-p expansion of k. We define the p-**weight degree** of a monomial $\mathbf{x}^{\mathbf{e}} = x_1^{e_1} \ldots x_n^{e_n}$ as $w_p(\mathbf{x}^{\mathbf{e}}) = \sigma(e_1) + \ldots + \sigma(e_n)$ and the p-**weight degree of a polynomial** F, $w_p(F)$, as the largest p-weight degree of the monomials in F.

Theorem 4.7 *Let* $P_1(\mathbf{x}), \ldots, P_t(\mathbf{x})$ *be polynomials in* $x_1, \ldots, x_n$ *over* $\mathbb{F}_{p^f}$*. For* $k = 1, \ldots, t$*, let* ℓ_k *be the* p*-weight degree of* P_k*, and define* μ *as the smallest integer satisfying*

$$\mu \geq \frac{f(n - \sum_k \ell_k)}{\max_k \ell_k}.$$

Then p^μ *divides* N*, the number of common zeros of* $P_1, \ldots, P_t$ *in* ${\mathbb{F}_{p^f}}^n$*.*

Proof: Let $f = 1$ and consider $F(\mathbf{x}) = y_1 P_1 + \ldots + y_t P_t \in \mathbb{F}_{p^f}[x_1, \ldots, x_n, y_1, \ldots, y_t]$. We will use Theorem 4.4 to prove that

$$\frac{\kappa_{p-1}(F)}{p-1} \geq t + \frac{n - \sum_k \ell_k}{\max_k \ell_k}.$$

Note that, since we are working over the prime field $\mathbb{F}_p$, we reduce each power of x_i mod $x_i^p - x_i$, yielding a polynomial taking the same values but of degree less than p in each variable. Then the p-weight degree and the degree coincide.

For each k let N_k denote the number of monomials in $P_k(\mathbf{x})$. Let the j^{th} monomial in $P_k(\mathbf{x})$ have (total) degree ℓ_{kj}. We let E_k denote the matrix $e(P_k(\mathbf{x}))$ as in Section 3.

To compute $\kappa_{p-1}(F)$ we employ $\boldsymbol{\nu} \in \mathbb{Z}_{\geq 0}{}^{\sum N_k}$. We take $\boldsymbol{\nu} = (\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_t)$ with $\boldsymbol{\nu}_k = (\nu_{1k}, \ldots, \nu_{N_k k})$.

Without loss of generality we assume ℓ_1 to be the largest degree of the given polynomials and the first monomial of polynomial P_k to be of degree ℓ_k. To compute $\kappa_{p-1}(F)$ we consider $\boldsymbol{\nu}$ satisfying the following matrix equation with $\boldsymbol{\lambda} \geq 1$:

$$\begin{pmatrix} E_1 & E_2 & \cdots & E_t \\ \begin{matrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ & \vdots & & \\ 0 & 0 & \cdots & 0 \end{matrix} & \begin{matrix} 0 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 \\ & \vdots & & \\ 0 & 0 & \cdots & 0 \end{matrix} & \cdots & \begin{matrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ & \vdots & & \\ 1 & 1 & \cdots & 1 \end{matrix} \end{pmatrix} \boldsymbol{\nu}^T = (p-1)\boldsymbol{\lambda}^T. \tag{24}$$

If we multiply (24) on the left by $(1^n 0^t)$, the vector with n 1s followed by t 0s, we sum all the rows of the E_ks, obtaining

$$\sum_{j,k} \ell_{kj} \nu_{jk} \geq n(p-1).$$

But, as we assumed, $\ell_k \geq \ell_{kj}$ for all j, so

$$\ell_1 |\boldsymbol{\nu}_1| + \cdots + \ell_t |\boldsymbol{\nu}_t| \geq n(p-1). \tag{25}$$

Now we use the equations coming from the lower t rows of (24). For $k = 1, \ldots, t$ we multiply row $n+k$ by $\ell_1 - \ell_k$ to get

$$(\ell_1 - \ell_k)|\boldsymbol{\nu}_k| \geq (\ell_1 - \ell_k)(p-1).$$

Summing these over k yields

$$\ell_1 \sum_k |\boldsymbol{\nu}_k| - \sum_k \ell_k |\boldsymbol{\nu}_k| \geq (p-1) \left(t\ell_1 - \sum_k \ell_k \right).$$

When we add (25) to this, we get

$$\ell_1 |\boldsymbol{\nu}| \geq (p-1) \left(n + \sum_{k=1}^{t} (\ell_1 - \ell_k) \right)$$

for any solution $\boldsymbol{\nu}$ to equation (24). This implies that

$$\frac{\kappa_{p-1}(E)}{p-1} \geq t + \frac{n - \sum_k \ell_k}{\ell_1},$$

and, using Theorem 4.4, we prove Moreno-Moreno's result for the prime field.

The case for a general f follows via the same technique of "reduction to the ground field" presented in the second part of the proof of Theorem 1 of [10]. □

We recall how this Moreno-Moreno theorem relates to Ax-Katz's. From Moreno-Moreno the number of solutions in $\mathbb{F}_{p^f}$ is divisible by

$$p^{\left\lceil f \frac{n - \sum_k \ell_k}{\max_k \ell_k} \right\rceil},$$

where the ℓ_i are the p-weight degrees, while Ax-Katz gives divisibility by

$$p^{f \left\lceil \frac{n - \sum_k d_k)}{\max_k d_k} \right\rceil},$$

where the d_i are the ordinary degrees. These bounds coincide when $f = 1$ and are not comparable in general, however in many cases Moreno-Moreno gives an improvement (typically, when the degrees are big compared to the characteristic).

5 A New Result

Recently, in [4], Wei Cao and Qi Sun improved the Chevalley-Warning-Ax-Katz estimates for certain polynomials, over any finite field. In this section, we improve their bound in the prime-field case.

Let

$$F = \sum_{i=1}^{r} a_i x_{i1}^{d_{i1}} x_{i2}^{d_{i2}} \cdots x_{in_i}^{d_{in_i}} + G(y_1, \ldots, y_m)$$

be a polynomial in $n = m + \sum_i n_i$ distinct variables over $\mathbb{F}_p$, set $d_i = \gcd(d_{i1}, \cdots, d_{in_i}, p-1)$, and consider the polynomial $\widetilde{F} = \sum_{i=1}^{r} a_i (x_{i1} x_{i2} \cdots x_{in_i})^{d_i} + G(y_1, \cdots, y_m)$. Their result, specialized to $\mathbb{F}_p$, is:

Theorem 5.1 (Cao-Sun) *Let $N(F)$ be the number of solutions of F. With the above notation,*

$$v_p(N(F)) \geq \lceil \frac{n - \deg(\widetilde{F})}{\deg(\widetilde{F})} \rceil.$$

Note that since

$$\frac{n - \deg(\widetilde{F})}{\deg(\widetilde{F})} = \frac{n_1 + \cdots + n_r + m}{\deg(\widetilde{F})} - 1 \leq \frac{n_1}{\deg(\widetilde{F})} + \cdots + \frac{n_r}{\deg(\widetilde{F})} + \frac{m}{\deg(G)} - 1$$

$$\leq \frac{n_1}{n_1 d_1} + \cdots + \frac{n_r}{n_r d_r} + \frac{m}{\deg(G)} - 1 = \sum_{i=1}^{r} \frac{1}{d_i} + \frac{m}{\deg(G)} - 1,$$

the following theorem improves Cao-Sun's bound in the prime-field case.

Theorem 5.2 *With the above notation,*

$$v_p(N(F)) \geq \lceil \sum_{i=1}^{r} \frac{1}{d_i} + \frac{m}{\deg(G)} \rceil - 1.$$

This result could be obtained from Adolphoson-Sperber's result in [1] using arguments from linear programming. However, our method is more straightforward.

To prove this theorem, we need the following lemma on the exponential sum of a monomial.

Lemma 5.3 *Let $d = \gcd(d_1, \ldots, d_n, p-1)$. Then*

$$v_\theta(S(x_1^{d_1} \cdots x_n^{d_n})) \geq \kappa_{p-1}\left(x_1^{d_1} \cdots x_n^{d_n}\right) \geq \frac{p-1}{d}.$$

Proof: From Theorem 3.1, $v_\theta(S(x_1^{d_1}\cdots x_n^{d_n})) \geq \kappa_{p-1}\left(x_1^{d_1}\cdots x_n^{d_n}\right)$, which is the smallest ν_1 satisfying

$$\begin{aligned} \nu_1 d_1 &= \lambda_1(p-1) \\ &\vdots \\ \nu_1 d_n &= \lambda_n(p-1) \end{aligned} \tag{26}$$

for positive λs.

Since $d = \gcd(d_1, \ldots, d_n, p-1)$, we can find a Bézout-type relation

$$d = \alpha_1 d_1 + ... + \alpha_n d_n + \beta(p-1). \tag{27}$$

Combining (27) and (26) we get

$$\nu_1 d = (\alpha_1\lambda_1 + ... + \alpha_n\lambda_n + \nu_1\beta)(p-1).$$

So $\nu_1 d$ is a non-zero multiple of $(p-1)$, thus $\nu_1 d \geq p-1$, from which the lemma follows. □

In the proof of the theorem we also use the following well known result.

Lemma 5.4 *Let G be a polynomial in m variables. Then*

$$\nu_\theta\left((S(G)\right) \geq \kappa_{p-1}(G) \geq \frac{(p-1)m}{\deg(G)}.$$

Proof: (of Theorem 5.2) Let $\widehat{F} = yF$. Then, by Theorem 3.1 and Lemmas 3.11, 4.3, and 5.4 we see that

$$\begin{aligned} \nu_p\left(N(F)\right) = \nu_p\left(S(\widehat{F})\right) - 1 &\geq \lceil \sum_{i=1}^{r} \frac{\kappa_{p-1}\left(x_{i1}^{d_{i1}}\cdots x_{in_i}^{d_{in_i}}\right)}{p-1} + \frac{\kappa_{p-1}(G)}{p-1} \rceil - 1 \\ &\geq \lceil \sum_{i=1}^{r} \frac{1}{d_i} + \frac{m}{\deg(G)} \rceil - 1. \end{aligned}$$

□

Example 5.5 *Let* $F = x_1^{21} + x_2^{20}x_3^{14} + x_4^{11}x_5^4 + y_1^{10} + y_2^{10} + y_1y_2$ *over* $\mathbb{F}_{31}$. *Cao-Sun's result does not give information about* $N(F)$ *since* $\widetilde{F} = x_1^3 + (x_2x_3)^2 + x_4x_5 + y_1^{10} + y_2^{10} + y_1y_2$. *Applying the above theorem, however, we obtain*

$$v_p\left(N(F)\right) \geq \lceil \frac{1}{3} + \frac{1}{2} + 1 + \frac{2}{10} \rceil - 1 = \lceil \frac{61}{30} \rceil - 1 = 2.$$

Corollary 5.6 *With the above notation, if at least one of the* d_i*'s is equal to 1, then* $p|N(F)$ *whenever* $r > 1$ *or* $G \neq 0$.

References

[1] A. Adolphson and S. Sperber, p-adic Estimates for Exponential Sums and the Theorem of Chevalley-Warning, *Ann. Sci. Ec. Norm. Super.*, 4^e série, vol **20**, pp. 545-556, 1987.

[2] A. Adolphson and S. Sperber, p-adic Estimates for Exponential Sums, *Lectures Notes in Mathematics*, **1454**, pp. 11-22, Springer 1990.

[3] J. Ax, Zeros of Polynomials over Finite Fields, *Am. J. Math*, vol. **86** pp. 255-261, 1964.

[4] W. Cao and Q. Sun, Improvements upon the Chevalley-Warning-Ax-Katz-type Estimates, *J. Number Theory*, **122** pp. 135-141, 2007.

[5] Castro, F. N., Rubio, I. and Vega, J., Divisibility of Exponential Sums and Solvability of Certain Equations over Finite Fields, *The Quart. J. Math.*, doi: 10.1093/qmath/han013, 2008.

[6] N. M. Katz, On a Theorem of Ax, *Am. J. Math.,* **93**, 1971 pp. 485–499.

[7] O. Moreno, C. Cáceres and M. Alonso, An Improved and Simplified Binary Ax Theorem, *Proc. 1994 IEEE International Symposium on Information Theory*, Trondheim, Norway June-July 1994.

[8] O. Moreno and C.J. Moreno, Improvement of the Chevalley-Warning and the Ax-Katz theorems, *Amer. J. Math.* **117:**1 (1995) pp. 241-244.

[9] O. Moreno and C. J. Moreno, The MacWilliams-Sloane Conjecture on the Tightness of the Carlitz-Uchiyama Bound and the Weights of Dual of BCH Codes, *IEEE Trans. Inform. Theory* **40:6** (1994) pp. 1894-1907.

[10] O. Moreno, F.N. Castro and H. F. Mattson, Jr., Correction to "Divisibility Properties for Covering Radius of Certain Cyclic Codes", *IEEE Trans. Inform. Theory* **52:4**, (2006) 1798-1799.

[11] O. Moreno, K. Shum, F.N. Castro and P. J. Kumar, Tight Bounds for Chevalley-Warning-Ax Type Estimates, with Improved Applications, *Proc. London Math. Soc.*, **88**(2004) pp. 545-564.

[12] Xiang-Dong Hou, A note on the proof of a theorem of Katz, *Finite Fields and Their Applications*, **11** (2005) pp. 316-319.

[13] D. Wan, An Elementary Proof of a Theorem of Katz, *Amer. J. Math*, **111** (1989) pp. 1-8.

[14] D. Wan, A Chevalley-Warning approach to p-adic estimates of character sums, *Proc. AMS Math. Soc.*, **3**(1995) pp. 45-54.

[15] D. Wan, Zeros of Diagonal Equations over Finite Fields, *Proc. A.M.S.*, **103**, 4, (1988) pp. 1049-1052.

[16] C. Chevalley, Demonstration d'une hypothese de M. Artin, *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg,* **11**, pp 73–75, 1935.

[17] E. Warning, Bermerkung zur vorstehenden Arbeit von Herrn Chevalley, *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg,* **11**, pp 76–83, 1935.

[18] R. M. Wilson, A Lemma on Polynomials Modulo p^m and Applications to Coding Theory, *Discrete Math.*, (**306**), pp. 3154-3165, 2006.