RESTRICTED VARIABLE CHEVALLEY-WARNING THEOREMS

ANURAG BISHNOI AND PETE L. CLARK

Abstract. We pursue various restricted variable generalizations of the Chevalley-Warning theorem for low degree polynomial systems over a finite field. Our first such result involves variables restricted to Cartesian products of the Vandermonde subsets of $\mathbb{F}_q$ defined by Gács-Weiner and Sziklai-Takáts. We then define an invariant $\underline{\omega}(X)$ of a nonempty subset of $\mathbb{F}_q^n$. Our second result involves X-restricted variables when the degrees of the polynomials are small compared to $\underline{\omega}(X)$. We end by exploring various classes of subsets for which $\underline{\omega}(X)$ can be bounded from below.

1. Introduction

We denote the set of non-negative integers by $\mathbb{N}$ and the set of positive integers by $\mathbb{Z}^+$. Throughout, $\mathbb{F}_q$ denotes a finite field of order q and characteristic p. For a finite, nonempty subset X of a field F, we put

$$\varphi_X(t) \coloneqq \prod_{x \in X} (t - x),$$

so φ_X generates the ideal of all polynomials $f \in F[t]$ that vanish identically on X.

Our point of departure is the following results of Chevalley and Warning, published in consecutive articles in the same journal in 1935 [Ch35], [Wa35].[1]

Theorem 1.1. *Let $f_1, \ldots, f_r \in \mathbb{F}_q[t_1, \ldots, t_n]$ be polynomials of degrees $d_1, \ldots, d_r \in \mathbb{Z}^+$, and suppose that $d \coloneqq \sum_{j=1}^r d_j < n$. Put*

$$Z = Z(f_1, \ldots, f_r) \coloneqq \{x = (x_1, \ldots, x_n) \in \mathbb{F}_q^n \mid f_1(x) = \ldots = f_r(x) = 0\}$$

be the solution set of the polynomial system.

a) *(Chevalley) We* cannot *have $\#Z = 1$.*
b) *(Chevalley-Warning) We have $p \mid \#Z$.*
c) *(Warning) If Z is nonempty, then $\#Z \geq q^{n-d}$.*

The second author has taken the perspective (e.g. in [Cl14]) that Theorem 1.1a) is a precursor of the following celebrated result.

Theorem 1.2 (Combinatorial Nullstellensatz II [Al99])**.** *Let F be a field, let $n \in \mathbb{Z}^+$, let $d_1, \ldots, d_n \in \mathbb{N}$ and let $f \in F[t_1, \ldots, t_n]$ be a polynomial. We suppose:*

(i) *We have $\deg(f) = d_1 + \ldots + d_n$, and*
(ii) *The coefficient of $t_1^{d_1} \cdots t_n^{d_n}$ in f is nonzero.*

Then, for any subsets $X_1, \ldots, X_n$ of F with $\#X_i = d_i + 1$ for $1 \leq i \leq n$, there exists an $x = (x_1, \ldots, x_n) \in X \coloneqq \prod_{i=1}^n X_i$ such that $f(x) \neq 0$.

[1] Some further information about the history of these results can be found in [CGS21, §1].

To see the connection, we introduce the **Chevalley polynomial**

$$\chi = \chi(f_1, \ldots, f_r) := \prod_{j=1}^{r} (1 - f_j^{q-1}).$$

For all $x \in \mathbb{F}_q^n$, we have

$$\chi(x) = \begin{cases} 1 & \text{if } x \in Z(f_1, \ldots, f_r) \\ 0 & \text{otherwise} \end{cases}.$$

Because the hypotheses and the conclusion of Theorem 1.1a) are stable under translation of variables, to prove Theorem 1.1a), seeking a contradiction we may assume that $Z = \{0\}$, and then the polynomial

$$P := \chi(t_1, \ldots, t_n) - \prod_{i=1}^{n} (1 - t_i^{q-1})$$

evaluates to 0 for all $x \in \mathbb{F}_q^n$. But our hypothesis $\sum_{j=1}^{r} d_j < n$ implies that $\deg(\chi) < (q-1)n$, so the coefficient of $t_1^{q_1} \cdots t_n^{q-1}$ in P is $(-1)^{n+1} \neq 0$. This contradicts Theorem 1.2 with $X_1 = \ldots = X_n = \mathbb{F}_q$.

In turn Theorem 1.2 motivates us to consider **restricted variable generalizations** of Theorem 1.1: instead of considering solutions to our polynomial system $f_1 = \ldots = f_r = 0$ on all of $\mathbb{F}_q^n$, we choose a subset $X \subset \mathbb{F}_q^n$ and look only at

$$Z_X := \{x = (x_1, \ldots, x_n) \in X \mid f_1(x) = \ldots = f_r(x) = 0\}.$$

A similar argument gives the following result of Schauz [Sc08] and Brink [Br11].

Theorem 1.3. *Let F be a field $\chi \in F[t_1, \ldots, t_n]$, let $X_1, \ldots, X_n$ be nonempty finite subsets of F, and put $X := \prod_{i=1}^{n} X_i$.*

a) *If $\deg(\chi) < \sum_{i=1}^{n} (\#X_i - 1)$, then $\#\{x \in \prod_{i=1}^{n} X_i \mid \chi(x) \neq 0\} \neq 1$.*
b) *(Restricted Variable Chevalley Theorem) Let $f_1, \ldots, f_r \in \mathbb{F}_q[t_1, \ldots, t_n]$ have positive degrees. If*

$$(q-1) \sum_{j=1}^{r} \deg(f_j) < \sum_{i=1}^{n} (\#X_i - 1), \tag{1}$$

then we have $\#Z_X \neq 1$.

Just as parts b) and c) of Theorem 1.1 give two different generalizations of Theorem 1.1a), it is natural to ask for restricted variable generalizations of Theorem 1.1b) and of Theorem 1.1c), each generalizing Theorem 1.3b).

The latter has been attained: in [CFS17, Thm. 1.6], Clark-Forrow-Schmitt give a **Restricted Variable Warning Theorem**:[2] the conclusion of Theorem 1.3b) is strengthened to: either $Z_X = \varnothing$ or $\#Z_X$ is at least a certain function of $n, d_1, \ldots, d_r, \#X_1, \ldots, \#X_n$ that is at least 2 when (1) holds.

The former case is addressed by the following result of [Cl14].

[2]In fact [CFS17, Thm. 1.6] is a further ring theoretic generalization motivated by work of Brink [Br11], but in the present paper we will only consider polynomials over a field.

Theorem 1.4. [Cl14, Thm. 19] *Let $f_1, \ldots, f_r \in \mathbb{F}_q[t]$ be polynomials of positive degrees. For $1 \leq i \leq n$ let X_i be a nonempty subset of $\mathbb{F}_q$. Put $X := \prod_{i=1}^n X_i$ and*

$$\varphi_i(t_i) := \varphi_{X_i}(t_i) = \prod_{x_i \in X_i} (t_i - x_i).$$

Let

$$Z_X := \{x = (x_1, \ldots, x_n) \in X \mid P_1(x) = \ldots = P_r(x) = 0\}.$$

Suppose

$$(q-1)\sum_{j=1}^r \deg(f_j) < \sum_{i=1}^n (\#X_i - 1). \tag{2}$$

Then as elements of $\mathbb{F}_q$ we have

$$\sum_{x \in Z_X} \frac{1}{\prod_{i=1}^n \varphi_i'(x_i)} = 0. \tag{3}$$

Theorem 1.4 is proved using the Coefficient Formula [Cl14, Thm. 3], a refinement of Theorem 1.2 due to Schauz [Sc08], Lason [La10] and Karasev-Petrov [KP12].

In the case that $X_1 = \ldots = X_n = \mathbb{F}_q$, we have $\varphi_i(t_i) = t_i^q - t_i$, so $\varphi_i'(t_i) = -1$. In this case (3) becomes the assertion that $(-1)^n \#Z = 0$ in $\mathbb{F}_q$, so $p \mid \#Z$, and we recover Theorem 1.1b). Moreover, (3) certainly implies that $\#Z_X \neq 1$, so we recover Theorem 1.3b). Thus Theorem 1.4 is a simultaneous generalization of the Chevalley-Warning Theorem and the Restricted Variable Chevalley Theorem, so in [Cl14] this result is called the "Restricted Variable Chevalley-Warning Theorem."

In this note we wish to reopen the question of what should, or could, constitute a Restricted Variable Chevalley-Warning Theorem. A distinguishing feature of Theorem 1.1b) is that it gives a "p-adic inequality on $\#Z$" – i.e., a p-divisibility on $\#Z$ – a feature that does not seem to be present in Theorem 1.4. Clearly the condition (2) of Theorem 1.4 is not in general sufficient to deduce $p \mid \#Z_X$: for instance fix $1 \leq I \leq n$, suppose $0 \in X_I$ and take $r = 1$ and $f_1 = t_I$: then $Z_X = \{(x_1, \ldots, x_n) \in \prod_{i=1}^n X_i \mid x_I = 0\}$, so $\#Z_X = \prod_{i \neq I} \#X_i$. This and similar examples show that it is not reasonable to expect Z_X to be divisible by p unless $p \mid \#X_i$ for all i. Moreover, beyond any condition on the sizes of the X_i's, we want somehow to take their structure into account.

Here is an example of how to do this, a generalization of Theorem 1.1b) due to Aichinger-Moosbauer [AM21, p. 62].

Proposition 1.5. *In the setting of Theorem 1.4, suppose moreover that each X_i is a coset: that is, there is a subgroup $G_i \subseteq (\mathbb{F}_q, +)$ and $a_i \in \mathbb{F}_q$ such that $X_i = a_i + G_i$ for all $1 \leq i \leq n$. If (2) holds, then $p \mid \#Z_X$.*

Proof. If for all $1 \leq i \leq n$ there is $c_i \in \mathbb{F}_q^\times$ such that for all $x_i \in X_i$ we have $\varphi_i'(x_i) = c_i$, then we have

$$\sum_{x \in Z_X} \frac{1}{\prod_{i=1}^n \varphi_i'(x_i)} = \sum_{x \in Z_X} \frac{1}{c_1 \cdots c_n} = \frac{\#Z_X}{c_1 \cdots c_n},$$

so when we apply Theorem 1.4 the conclusion is $0 = \frac{\#Z_X}{c_1 \cdots c_n}$ and thus $p \mid \#Z_X$.

For any finite nonempty subset X of a field F, if $\varphi_X(t) := \prod_{x \in X}(t - x)$, then

for all $x \in X$ we have

$$\varphi'_X(x) = \prod_{y \in X \setminus \{x\}} (x - y).$$

Taking $X = X_i = a_i + G_i$, for all $x_i \in a_i + G_i$ we get

$$\varphi'_{X_i}(x_i) = \prod_{y \in (a_i + G_i) \setminus \{x_i\}} (x_i - y) = \prod_{x \in G_i \setminus \{0\}} x =: c_i;$$

that is to say, the value of $\varphi'_{X_i}(x_i)$ does not depend upon the choice of $x_i \in X_i$. $\square$

The argument of Proposition 1.5 will work whenever $X = \prod_{i=1}^n X_i$ and for all $1 \leq i \leq n$ we have that $\varphi'_{X_i}(x) = \varphi'_{X_i}(y)$ for all $x, y \in X_i$. Our first main result is that, for any field F of characteristic $p > 0$ and any finite subset $X \subset F$ of cardinality at least 2, we have $\varphi'_X(x) = \varphi'_X(y)$ for all $x, y \in X$ if and only if X has size divisible by p and is a **Vandermonde set**[3] in the sense of Gács-Weiner [GW03] and Sziklai-Takáts [ST08]. We deduce the following result.

Theorem 1.6 (Restricted Chevalley-Warning for Vandermonde Sets)**.** *For $1 \leq i \leq n$, let $X_i \subset \mathbb{F}_q$ be a Vandermonde set of size divisible by p. Let $f_1, \ldots, f_r \in \mathbb{F}_q[t_1, \ldots, t_n]$ be polynomials of positive degree such that*

$$(q-1) \sum_{j=1}^r \deg(f_j) < \sum_{i=1}^n (\#X_i - 1).$$

Let

$$Z_X := \{x \in \prod_{i=1}^n X_i \mid P_1(x) = \ldots = P_r(x) = 0\}.$$

Then $p \mid \#Z_X$.

We also give some examples of Vandermonde subsets of $\mathbb{F}_q$ of cardinality divisible by p that are not cosets of additive subgroups.

Let F be a field of characteristic $p > 0$. In §3, for any $n \in \mathbb{Z}^+$ and any finite, nonempty subset $X \subset F^n$, we define an invariant $\underline{\omega}(X) \in \mathbb{N}$ that is (almost) a multivariate generalization of the invariant $\omega(Y)$ of a subset $Y \subset \mathbb{F}_q$ considered by Gács-Weiner and Sziklai-Takáts. We show the following result:

Theorem 1.7. *Let $X \subset \mathbb{F}_q^n$ be a nonempty subset, and let $f_1, \ldots, f_r \in \mathbb{F}_q[t_1, \ldots, t_n]$ be polynomials of positive degree. If*

$$(q-1) \sum_{j=1}^r \deg(f_j) < \underline{\omega}(X), \tag{4}$$

then

$$p \mid \#Z_X = \#\{x \in X \mid f_1(x) = \ldots = f_r(x) = 0\}.$$

Of course this focuses attention on what we know about $\underline{\omega}(X)$. We have $\underline{\omega}(X) \geq 1$ iff $p \mid \#X$. Moreover, if X_i is the projection of X onto its ith factor, then we have

$$\underline{\omega}(X) \leq \sum_{i=1}^n (\#X_i - 1). \tag{5}$$

[3] In §2 we will give a self-contained treatment of Vandermonde sets.

We call a subset $X \subset F^n$ **optimal** if equality occurs in (5). In Lemma 3.3 we show that if $X = \prod_{i=1}^n X_i$ and each X_i is Vandermonde of size divisible by p, then X is otpimal. Thus Theorem 1.7 implies Theorem 1.6. Because of this we suggest that Theorem 1.7 is more deserving of the name "Restricted Variable Chevalley-Warning Theorem" than Theorem 1.4.

Remark 1.8. *The main observation of this paper is that a Restricted Variable Chevalley Theorem for a subset $X \subset \mathbb{F}_q^n$ should take into account the "structure" of X via an invariant that measures the vanishing of certain symmetric functions summed over X. This observation is also the point of departure of a recent preprint of Nica* [Ni21]. *The main difference is that while we are interested in restricted variable theorems whose conclusion is $p \mid \#Z_X$, Nica is interested in improvements of the Combinatorial Nullstellensatz and the Coefficient Formula that take the structure of the "grid" $X = \prod_{i=1}^n X_i$ into account.*

His work and ours were done independently, and despite the similarity of content, have virtually no overlap.[4]

2. Vandermonde sets

Vandermonde subsets of finite fields were defined by Gács-Weiner [GW03] in characteristic 2 and by Sziklai-Takáts [ST08] in all characteristics. In this section we study Vandermonde subsets over any field F of characteristic $p > 0$.

2.1. Vandermonde Sets and their Polynomials.

Let Y be a finite subset of F of cardinality $r \geq 1$. For $k \in \mathbb{Z}^+$ we put

$$\pi_k(Y) := \sum_{y \in Y} y^k.$$

Lemma 2.1. *Let $\{0\} \neq Y \subset F$ be a finite subset of size $r \geq 1$.*

a) *If $p \mid r$ then $\pi_k(Y) \neq 0$ for some $1 \leq k \leq r-1$.*
b) *In general we have $\pi_k(Y) \neq 0$ for some $1 \leq k \leq r$.*

Proof. Write $Y = \{y_1, \dots, y_r\}$.
a) Consider the $r \times r$ Vandermonde matrix $V = V(y_1, \dots, y_r)$ associated to $y_1, \dots, y_r$, where the i-th row of V is $(y_1^{i-1}, \dots, y_r^{i-1})$. Say $\pi_k(Y) = 0$ for all $1 \leq k \leq r-1$. Then V evaluated at the column vector $(1, \dots, 1)^T$ is 0, since the the first entry of the product is $r = 0 \in F$ as $p \mid r$ and rest of the entries are $\pi_k(Y)$ for $1 \leq k \leq r-1$. This contradicts the nonsingularity of V.
b) Suppose first that $0 \notin Y$, and consider the matrix $\tilde{V}$ obtained from V by multiplying the jth column by y_j, so $\det \tilde{V} = y_1 \cdots y_r \det(V) \neq 0$. Again we must have that $\tilde{V}(1, \dots, 1)^T \neq 0$ which means that $\pi_k(Y) \neq 0$ for some $1 \leq k \leq r$. If $0 \in Y$, let $Y^\bullet := Y \setminus \{0\}$. As we have just shown, there is $1 \leq k \leq r-1$ such that $\pi_k(Y^\bullet) \neq 0$, so then $\pi_k(Y) = \pi_k(Y^\bullet) \neq 0$. □

For the rest of this section Y will denote a finite subset of F of size $r \geq 2$.

We denote by $\omega(Y)$ the least $k \in \mathbb{Z}^+$ such that $\pi_k(Y) \neq 0$. Lemma 2.1 gives

$$\omega(Y) \leq \begin{cases} r & \text{always} \\ r-1 & \text{if } p \mid \#Y \end{cases}.$$

[4] In both [Ni21] and the present work, Newton's identities are used to relate vanishing of power sums to vanishing of symmetric functions...as surely many others have done as well.

Following Sziklai-Takáts [ST08] we say that Y is **Vandermonde** if $\omega(Y) = r - 1$ and is **super-Vandermonde** if $\omega(Y) = r$.

Lemma 2.2. *Let $Y \subset F$ be finite of cardinality $r \geq 2$.*

a) *For all $k \in \mathbb{Z}^+$ and $\alpha \in F^\times$ we have $\pi_k(Y) = 0 \iff \pi_k(\alpha Y) = 0$. Thus $\omega(Y) = \omega(\alpha Y)$ and Y is Vandermonde $\iff$ αY is Vandermonde.*
b) *Suppose Y is Vandermonde. Then for $\beta \in F^\times$ the translate $Y + \beta$ is Vandermonde iff $p \mid r$.*

Proof. a) Indeed $\pi_k(\alpha Y) = \alpha^k \pi_k(Y)$, and the rest is clear.
b) If $Y = \{y_1, \ldots, y_r\}$ then for all $1 \leq k \leq r - 2$, expanding out each term in

$$\pi_k(Y + \beta) = \sum_{i=1}^{r} (y_i + \beta)^k$$

gives a linear combination of $\pi_j(Y)$ for $j \leq k$ together with the final term $r\beta^k$, which is zero iff $p \mid r$. □

Proposition 2.3. *Let F be a field of characteristic $p > 0$, let $Y = \{y_1, \ldots, y_r\} \subset F$ be finite of cardinality $r \geq 2$, and write*

$$\varphi_Y(t) = \prod_{i=1}^{r} (t - y_i) = t^r + \sum_{i=0}^{r-1} a_i t^i.$$

a) *The following are equivalent:*
   (i) *The subset Y is Vandermonde or super-Vandermonde.*
   (ii) *For $2 \leq i \leq r - 1$, if $p \nmid i$ then $a_i = 0$.*
b) *The following are equivalent:*
   (i) *The subset Y is super-Vandermonde.*
   (ii) *For $1 \leq i \leq r - 1$, if $p \nmid i$ then $a_i = 0$.*

Proof. For independent indeterminates $x_1, \ldots, x_r$ and $0 \leq k \leq r$ let $s_k(x_1, \ldots, x_r)$ be the kth elementary symmetric function and let $\pi_k(x_1, \ldots, x_r) = \sum_{i=1}^{r} x_i^k$ be the kth power sum. Note that $a_i = (-1)^{r-i} s_{r-i}(x_1, \ldots, x_r)$, for $0 \leq i \leq r - 1$. In the ring $\mathbb{Z}[x_1, \ldots, x_r]$ we have **Newton's identities** (see e.g. [Ze84] for a short proof): for all $1 \leq k \leq r$ we have

$$k s_k(x_1, \ldots, x_r) = \sum_{i=1}^{k} (-1)^{i-1} s_{k-i}(x_1, \ldots, x_r) \pi_k(x_1, \ldots, x_r)$$

and

$$\pi_k(x_1, \ldots, x_r) = (-1)^{k-1} k s_k(x_1, \ldots, x_r) + \sum_{i=1}^{k-1} (-1)^{k-1+i} s_{k-i}(x_1, \ldots, x_r) \pi_i(x_1, \ldots, x_r).$$

From these we see that for elements $y_1, \ldots, y_r$ in a field of characteristic p, for any $1 \leq k \leq r$, the vanishing of the power sums $\pi_i(y_1, \ldots, y_r)$ for $1 \leq i \leq k$ is equivalent to the vanishing of the elementary symmetric functions $s_i(y_1, \ldots, y_r)$ for $1 \leq i \leq k$ and $p \nmid i$. The result follows easily. □

Theorem 2.4. *Let F be a field of characteristic $p > 0$. For a nonempty subset $Y \subset F$ or size $r \geq 2$, the following are equivalent:*

(i) *There is $g \in F[t]$ and $c \in F^\times$ such that $\varphi_Y = g(t^p) + ct$.*
(ii) *The polynomial φ_Y' is a nonzero constant (i.e., has degree zero).*

(iii) *The polynomial* φ_Y' *is constant on* Y.
(iv) *The set* Y *is Vandermonde of size a multiple of* p.

Proof. (i) $\iff$ (iv): It follows from Proposition 2.3a) that Y is Vandermonde of size divisible by p if and only if $\varphi_Y = g(t^p) + ct$ for some $c \in F$. Since $p \mid \#Y$, by Lemma 2.1 the set Y is *not* super-Vandermonde, so $c \neq 0$.
(i) $\implies$ (ii): We have $(g(t^p) + ct)' = c$.
(ii) $\implies$ (iii): This is immediate.
(iii) $\implies$ (i): Let $r = \#Y = \deg \varphi_Y$, and let $c \in F$ be the constant value of φ_Y' on Y. Since the polynomial $\varphi_Y(X)$ is separable, for all $y \in Y$ we have $\varphi_Y'(y) \neq 0$, so $c \neq 0$. Consider $f := \varphi_Y - ct$. Then f' has degree at most $r-1$ and vanishes identically on Y, a set of size r, so f' is the zero polynomial. Thus there is $g \in F[t]$ such that $\varphi_Y - ct = f = g(t^p)$. □

As explained in the introduction, Theorem 2.4 implies Theorem 1.6. In fact we get the following more general result that applies to any field of characteristic p.

Theorem 2.5 (Vandermonde Coefficient Formula)**.** *Let* F *be a field of prime characteristic* p. *For* $1 \leq i \leq n$ *let* $X_i \subset F$ *be a finite nonempty Vandermonde subset of cardinality* $d_i + 1$ *divisible by* p. *For* $1 \leq i \leq n$, *let* $c_i = \varphi_i' \in F^\times$. *Put* $d = (d_1, \ldots, d_n)$ *and* $X = \prod_{i=1}^n X_i$. *Let* $f \in F[t_1, \ldots, t_n]$ *be* d*-topped – this means that the monomial* $t_1^{d_1} \cdots t_n^{d_n}$ *does not divide any other monomial appearing in* f *with a nonzero coefficient, and this holds when* $\deg(f) \leq d$ *– and let* $c_d(f)$ *be the coefficient of* $t_1^{d_1} \cdots t_n^{d_n}$ *in* f. *Then we have*

$$(c_1 \cdots c_n)\, c_d(f) = \sum_{x \in X} f(x).$$

Proof. This follows from a suitable version of the Coefficient Formula: e.g. from [Cl14, Thm. 3.9]. □

2.2. **More on Cosets.** Suppose $Y \subset F$ is a coset of a finite subgroup of $(F, +)$. Then φ_Y' is constant on Y: this was shown in the proof of Proposition 1.5 when F is finite, but the argument holds verbatim. Applying Theorem 2.4, we deduce:

Corollary 2.6. *Let* Y *be a coset of a finite subgroup of* $(F, +)$. *Then* Y *is a Vandermonde set.*

Corollary 2.6 was shown for additive subgroups in $\mathbb{F}_{2^a}$ by Gács-Weiner [GW03, Ex. 2.3(i)] and for cosets of additive subgroups in $\mathbb{F}_q$ by Sziklai-Takáts [ST08, Prop. 1.8(i)]. Our proof is different from both of theirs, which draw on the theory of additive polynomials. With this approach one can deduce more about φ_Y when Y is a coset, as we now explain.

Suppose F contains the finite field $\mathbb{F}_q$. A polynomial $f \in F[t]$ is **functionally $\mathbb{F}_q$-linear** if the induced map $E(f) : F \to F$ given by $x \mapsto f(x)$ is an $\mathbb{F}_q$-vector space endomorphism. In characteristic zero, if the induced map $E(f) : F \to F$ were even a group homomorphism then f would have to be of the form $f(t) = at$ for some $a \in F$; however in positive characteristic the functionally $\mathbb{F}_q$-linear polynomials are precisely those of the form $\sum_{i=0}^n a_i t^{q^i}$ for $a_i \in F$. We say that a polynomial $f \in F[t]$ is **additive** if it is functionally $\mathbb{F}_p$-linear: equivalently, $E(f) : (F, +) \to (F, +)$ is a group homomorphism. A polynomial $f \in F[t]$ is **functionally $\mathbb{F}_q$-affine** if $f - f(0)$ is functionally $\mathbb{F}_q$-linear.

Theorem 2.7. *Let $Y \subset F$ be a finite nonempty subset.*

a) *The following are equivalent:*
   (i) *The subset Y is an $\mathbb{F}_q$-subspace of F.*
   (ii) *The polynomial φ_Y is functionally $\mathbb{F}_q$-linear.*

b) *The following are equivalent:*
   (i) *The subset Y is a coset of an $\mathbb{F}_q$-subspace of F.*
   (ii) *The polynomial φ_Y is functionally $\mathbb{F}_q$-affine.*

Proof. When F is finite, these results are special cases of [LN97, Thms. 3.56 and 3.57]. The proofs given there do not use the finiteness of F. □

In particular, if Y is a coset of a finite subgroup of $(F, +)$ then $\varphi_Y - \varphi_Y(0)$ is an additive polynomial, so there are $a_0, \ldots, a_n, b \in F$ such that

$$\varphi_Y(t) = \sum_{i=0}^{n} a_i t^{p^i} + b.$$

Notice that this is stronger than φ_Y just being of the form $g(t^p) + ct$. If $F \supseteq \mathbb{F}_q$ and Y is a coset of an $\mathbb{F}_q$-subspace, then there are $a_0, \ldots, a_n, b \in F$ such that

$$\varphi_Y(t) = \sum_{i=0}^{n} a_i t^{q^i} + b,$$

a stronger conclusion still.

2.3. **More on Vandermonde Sets.** Theorem 1.6 is a generalization of Proposition 1.5. How much of an improvement is it? This comes down to asking how many more Vandermonde subsets of $\mathbb{F}_q$ of cardinality divisible by p there are than cosets of additive subgroups. In [ST08, Prop. 1.8(ii)], Sziklai and Takáts construct a family of such Vandermonde sets over $\mathbb{F}_{q^2}$ with $p = 2$ that are, in general, not cosets of additive subgroups. In particular, certain geometrical objects in finite projective planes known as hyperovals (see [BCP06] and the references therein for the list of known infinite families of hyperovals), give rise to such Vandermonde sets. There are further examples of Vandermonde sets known for small fields (see for example [AH19, Example 7]), but a full classification is out of reach (simply because a full classification of hyperovals appears to be out of reach [Va19]).

We ask the following question related to the enumeration of Vandermonde sets.

Question 2.8. *For a prime number p and $a \in \mathbb{Z}^+$, let $V(p, n)$ be the number of subsets $X \subset \mathbb{F}_{p^n}$ that are Vandermonde of size divisible by p, and let $C(p, n)$ be the number of subsets $X \subset \mathbb{F}_{p^n}$ that are cosets of subgroups of $(\mathbb{F}_{p^n}, +)$. Straightforward calculation gives*

$$C(p, n) = \sum_{d=0}^{n} p^{(1-d)(n-d)} \frac{\# \operatorname{GL}_n(\mathbb{F}_p)}{\# \operatorname{GL}_d(\mathbb{F}_p) \# \operatorname{GL}_{n-d}(\mathbb{F}_p)},$$

where

$$\# \operatorname{GL}_n(\mathbb{F}_p) = \prod_{i=1}^{n} (p^n - p^{n-i+1}).$$

Is it true that for each fixed p we have $C(p, n) = o(V(p, n))$ as $n \to \infty$?

3. The Invariant $\underline{\omega}(X)$

Again let F be a field of characteristic $p > 0$. For a function $f : F^n \to F$ and a finite subset $X \subset F^n$ we put

$$\int_X f := \sum_{x \in X} f(x) \in F.$$

A polynomial $P \in F[t_1, \ldots, t_n]$ determines a function $E(P) : x \in F^n \mapsto P(x) \in F$. We put $\int_X P := \int_X E(P)$.

For $\underline{k} = (k_1, \ldots, k_n) \in \mathbb{N}^n$, we put $|\underline{k}| := k_1 + \ldots + k_n$. For a finite nonempty subset $X \subseteq F^n$ and $\underline{k} = (k_1, \ldots, k_n) \in \mathbb{N}^n$, let

$$\pi_{\underline{k}}(X) := \int_X t_1^{k_1} \cdots t_n^{k_n} = \sum_{x=(x_1,\ldots,x_n)\in X} x_1^{k_1} \cdots x_n^{k_n}.$$

Thus

$$\pi_{\underline{0}}(X) = \int_X 1 = \#X \in F. \tag{6}$$

We put

$$\underline{\omega}(X) := \inf\{|\underline{k}| \mid \pi_{\underline{k}}(X) \neq 0\}.$$

Thus for $d \in \mathbb{N}$ we have $\underline{\omega}(X) \geq d + 1$ iff $\int_X P = 0$ for every $P \in F[t_1, \ldots, t_n]$ of degree at most d. In particular, by (6) we have $\underline{\omega}(X) \geq 1$ iff $\pi_{\underline{0}}(X) = 0$ iff $p \mid \#X$.

Our definition allows $\underline{\omega}(X)$ to be infinite, which would occur if and only if $\pi_{\underline{k}}(X) = 0$ for all $\underline{k} \in \mathbb{N}^n$. But the following result shows that this is never the case.

Lemma 3.1. *Let $X \subset F^n$ be finite and nonempty. For $1 \leq i \leq n$, let*

$$X_i := \{a \in F \mid \exists\ (x_1, \ldots, x_{i-1}, a, x_{i+1}, \ldots, x_n) \in X\}$$

be the projection of X onto its ith coordinate. Then we have

$$\underline{\omega}(X) \leq \sum_{i=1}^{n} (\#X_i - 1)\,. \tag{7}$$

Proof. Choose $x = (x_1, \ldots, x_n) \in X$, and put

$$\delta_{X,x} := \prod_{i=1}^{n} \prod_{y_i \in X_i \setminus \{x_i\}} \frac{t_i - y_i}{x_i - y_i}.$$

Then $\deg \delta_{X,x} = \sum_{i=1}^{n} (\#X_i - 1)$. For $y \in X$ we have $\delta_{X,x}(y) = \begin{cases} 1 & y = x \\ 0 & y \neq x \end{cases}$, so and $\int_X \delta_{X,x} = 1$. □

To be sure: when $n = 1$ and $Y \subseteq F$ is finite of size $r \geq 1$, we have $\underline{\omega}(Y) = \omega(Y)$ if $p \mid \#Y$, while if $p \nmid \#Y$ we have $\underline{\omega}(Y) = 0$ and $\omega(Y) \geq 1$. (In fact, we did not define $\omega(Y)$ for subset of size 1: extending the definition we gave, we would have $\omega(\{x\}) = 1$ if $x \neq 0$, while $\omega(\{0\})$ would be infinite.)

3.1. **Proof of Theorem 1.7.** Once again, let $\chi := \prod_{i=1}^r (1 - f_i^{q-1})$ be Chevalley's polynomial, so for $x \in X$ we have

$$\chi(x) = \begin{cases} 1 & x \in Z_X \\ 0 & x \notin Z_X \end{cases}.$$

Moreover $\deg \chi = (q-1) \sum_{j=1}^r d_j < \underline{\omega}(X)$, so in $\mathbb{F}_q$ we have

$$\# Z_X = \int_X \chi(x) = 0 \in \mathbb{F}_q.$$

Since $\mathbb{F}_q$ has characteristic p, this yields $p \mid \# Z_X$.

3.2. **Optimal Subsets.** A finite nonempty subset $X \subseteq F^n$ is **optimal** if equality holds in (7):

$$\underline{\omega}(X) = \sum_{i=1}^n (\# X_i - 1).$$

Thus when $n = 1$ we get that an optimal subset of F is precisely a set that either has size 1 or is a Vandermonde set of size divisible by p.

Remark 3.2. *That $\mathbb{F}_q^n$ is optimal is the crux of Ax's "Quick Proof of the Chevalley-Warning Theorem"* [Ax64, §2], [CGS21, Thm. 1.1]. *In some sense the proof of Theorem 1.7 corresponds to* the rest *of Ax's proof...which is why it is so short.*

Lemma 3.3. *For $1 \leq i \leq n$, let $X_i \subseteq F$ be finite nonempty, and put $X := \prod_{i=1}^n X_i$.*

a) *For all $\underline{k} \in \mathbb{N}^n$ we have $\pi_{\underline{k}}(X) = \prod_{i=1}^n \pi_{k_i}(X_i)$.*
b) *Let $J := \{1 \leq i \leq n \mid \# X_i \neq 1\}$. Then we have*

$$\underline{\omega}(X) = \sum_{j \in J} \omega(X_j).$$

c) *If each $X_i \subseteq \mathbb{F}_q$ is Vandermonde of size divisible by p, then*

$$\underline{\omega}(X) = \sum_{i=1}^n (\# X_i - 1),$$

so X is optimal.

Proof. a) We have

$$\pi_{\underline{k}}(X) = \sum_{(x_1, \ldots, x_n) \in \prod_{i=1}^n X_i} x_1^{k_1} \cdots x_n^{k_n} = \prod_{i=1}^n \sum_{x_i \in X_i} x_i^{k_i} = \prod_{i=1}^n \pi_{k_i}(X_i).$$

Part b) follows in the case where $\# X_i \geq 2$ for all i, as does part c). If $j \in \{1, \ldots, n\} \setminus J$ and $\underline{k} \in \mathbb{N}^n$ is such that $\pi_{\underline{k}}(X) \neq 0$, then $\pi_{(k_1, \ldots, k_{j-1}, 0, k_{j+1}, \ldots, k_n)}(X) \neq 0$. If $\underline{k} \in \mathbb{N}^n$ is such that $\underline{k}_j = 0$ for all $j \notin J$, then let $\hat{\underline{k}} \in \mathbb{N}^J$ be the corresponding tuple with the indices outside of J removed. We have

$$\pi_{\underline{k}}(X) = \pi_{\hat{\underline{k}}}(\prod_{j \in J} X_j) = \sum_{j \in J} \omega(X_j). \qquad \square$$

Lemma 3.3 and Theorem 1.7 together imply Theorem 1.6.

3.3. **The invariant $\underline{\omega}(X)$ of a coset X.** Now we look more closely at $\underline{\omega}(X)$ for a coset X of a finite additive subgroup of F. The following result reduces us to studying finite subgroups $G \subset (F^n, +)$.

Lemma 3.4 (Affine Invariance)**.** *Let $\mathrm{AGL}_n(F) = F^n \rtimes \mathrm{GL}_n(F)$ be the group of affine transformations of F^n. For all $\sigma \in \mathrm{AGL}_n(F)$ and all finite nonempty $X \subset F^n$ we have $\underline{\omega}(\sigma(X)) = \underline{\omega}(X)$.*

Proof. The group $\mathrm{AGL}_n(F)$ acts on polynomials: if

$$\sigma \in \mathrm{AGL}_n(F) : (x_1, \dots, x_n) \in \mathbb{F}_q^n \mapsto (a_1 + \sum_{j=1}^n m_{1,j} x_j, \dots, a_n + \sum_{j=1}^n m_{n,j} x_j)$$

and $P \in F[t_1, \dots, t_n]$ then

$$P_\sigma = P(a_1 + \sum_{j=1}^n m_{1,i} t_i, \dots, a_n + \sum_{j=1}^n m_{n,j} t_i).$$

This action preserves the degree. So if $\deg(P) < \underline{\omega}(X)$ then

$$\int_{\sigma(X)} P = \int_X P_\sigma = 0,$$

so $\underline{\omega}(\sigma(X)) \leq \underline{\omega}(X)$. Applying this with σ^{-1} and $\sigma(X)$ in place of σ and X we deduce that $\underline{\omega}(\sigma(X)) = \underline{\omega}(X)$. □

Example 3.5. *Let $X \subset \mathbb{F}_p^2$ be a nontrivial, proper coset, so $\#X = p$. The group $\mathrm{AGL}_2(\mathbb{F}_p)$ acts transitively on all nontrivial proper cosets of $\mathbb{F}_p^2$, so by Lemma 3.4 all such cosets have the same invariant $\underline{\omega}(X)$; taking $X = \mathbb{F}_p \times \{0\}$ and applying Lemma 3.3 we find that this common value is $p-1$. However, for the subgroup*

$$\Delta := \{(a, a) \mid a \in \mathbb{F}_p\}$$

we have $\underline{\omega}(\Delta) = (p-1) < 2(p-1) = (\#\mathbb{F}_p - 1) + (\#\mathbb{F}_p - 1)$. Thus optimality is not *$\mathrm{AGL}_n(F)$-invariant. More precisely optimality is translation-invariant and is not generally $\mathrm{GL}_n(F)$-invariant.*

Proposition 3.6. *Let $G \subset \mathbb{F}_q^n$ be an additive subgroup.*

a) *If G is an $\mathbb{F}_q$-subspace, then $\underline{\omega}(G) = (\dim_{\mathbb{F}_q}(G))(q-1)$.*
b) *If $q = p$, then $\underline{\omega}(G) = (\log_p(\#G))(p-1)$.*

Proof. a) Let $e := \dim_{\mathbb{F}_q}(G)$. Since $\mathrm{GL}_n(\mathbb{F}_q)$ acts transitively on $\mathbb{F}_q$-subspaces of dimension e, we have $\underline{\omega}(G) = \underline{\omega}(\mathbb{F}_q^e \times \{0\}^{n-e}) = e(q-1)$.
b) Every finite subgroup of a field of characteristic p is an $\mathbb{F}_p$-subspace, of dimension $\log_p(\#G)$. So part a) applies. □

Thus we can compute $\omega(G)$ when $n = 1$ or $F = \mathbb{F}_p$. The remaining case seems more interesting.

Example 3.7. *There are subgroups A_1, B_1, B_2 of $(F, +)$ with*

$$A_1 \cong (\mathbb{Z}/p\mathbb{Z})^2, \ B_1 \cong B_2 \cong \mathbb{Z}/p\mathbb{Z}.$$

We put

$$A := A_1 \times \{0\}^{n-1}, \ B = B_1 \times B_2 \times \{0\}^{n-2}.$$

Then

$$\underline{\omega}(A) = \underline{\omega}(A_1) = p^2 - 1 > 2(p-1) = \underline{\omega}(B_1 \times B_2) = \underline{\omega}(B),$$

even though $A \cong (\mathbb{Z}/p\mathbb{Z})^2 \cong B$ and thus there is an $\mathbb{F}_p$-linear automorphism $\sigma : F \to F$ such that $B = \sigma(A)$.

In light of Examples 3.5 and 3.7 it is not clear to us what form the determination of $\underline{\omega}(G)$ for an arbitrary finite subgroup $G \subset (F, +)$ should take. The following is the best possible lower bound that takes only the isomorphism class of G into account:

Proposition 3.8 (Aichinger-Moosbauer)**.** *Let F be a field of characteristic $p > 0$, and let $G \subset (F^n, +)$ be a finite subgroup of order p^e. Then*

$$\underline{\omega}(G) \geq e(p-1). \tag{8}$$

Proof. This follows from [AM21, Lemma 8.2 and Lemma 12.1]. □

Remark 3.9. *Combining Proposition 3.8 and Theorem 1.7, we find that if $X \subset \mathbb{F}_q^n$ is a coset of a subgroup of order p^e and $f_1, \ldots, f_r \in \mathbb{F}_q[t_1, \ldots, t_n]$ have positive degree, then if*

$$(q-1)\left(\sum_{j=1}^{r} \deg(f_j)\right) < d(p-1),$$

then $p \mid \#Z_X$. When $q = p$, this result implies Proposition 1.5. If $q = p^a$ with $a > 1$, then a better bound follows from [AM21, Thm. 12.2]*: we have $p \mid \#Z_X$ if*

$$a\left(\sum_{j=1}^{r} \deg(f_j)\right) < p-1. \tag{9}$$

The low degree condition (9) has a different form than the low degree condition (2) of Proposition 1.5, but using an elementary convexity argument, Aichinger and Moosbauer show that if (2) holds then so does (9).

3.4. **Graphs of Functions.** We end by giving a class of subsets X of $\mathbb{F}_q^n$ for which $\underline{\omega}(X)$ can be bounded from below.

Proposition 3.10. *Let $n \geq 2$. For $f \in \mathbb{F}_q[t_1, \ldots, t_{n-1}]$ of degree $d \geq 1$, let*

$$X_f \coloneqq \{(x, f(x)) \in \mathbb{F}_q^n\}$$

be the graph of the associated function $E(f) : \mathbb{F}_q^{n-1} \to \mathbb{F}_q$. Then:

a) *We have $\omega(X_f) \geq \frac{(n-1)(q-1)}{d}$.*
b) *If $\omega(X_f) = \frac{(n-1)(q-1)}{d}$ then $d \mid (n-1)(q-1)$.*
c) *If $n-1 \mid d \mid (n-1)(q-1)$ and the only monomial of degree d in the support of f is $t_1^{\frac{d}{n-1}} \cdots t_{n-1}^{\frac{d}{n-1}}$, then $\omega(X_f) = \frac{(n-1)(q-1)}{d}$.*
d) *If $n = 2$, then $\omega(X_f) = \frac{q-1}{d}$ iff $d \mid q-1$.*

Proof. a) For $\underline{k} = (k_1, \ldots, k_n) \in \mathbb{N}$, we have $\pi_{\underline{k}}(X_f) =$

$$\sum_{(x_1,\ldots,x_{n-1}) \in \mathbb{F}_q^{n-1}} x_1^{k_1} \cdots x_{n-1}^{k_{n-1}} f(x_1, \ldots, x_{n-1})^{k_n} = \int_{\mathbb{F}_q^{n-1}} t_1^{k_1} \cdots t_{n-1}^{k_{n-1}} f(t_1, \ldots, t_{n-1})^{k_n}.$$

Since $\mu(\mathbb{F}_q^{n-1}) = (n-1)(q-1)$, if $\pi_{\underline{k}}(X) \neq 0$ then

$$k_1 + \ldots + k_{n-1} + dk_n = \deg(t^{k_1} \cdots t^{k_{n-1}} f^{k_n}) \geq (n-1)(q-1),$$

so

$$|\underline{k}| = k_1 + \ldots + k_n \geq \frac{k_1}{d} + \ldots + \frac{k_{n-1}}{d} + k_n \geq \frac{(n-1)(q-1)}{d}.$$

b) Clearly equality can only hold if $d \mid (n-1)(q-1)$.
c) If the hypotheses hold, then the only monomial of degree $(n-1)(q-1)$ in the support of $f^{\frac{(n-1)(q-1)}{d}}$ is $t_1^{q-1} \cdots t_{n-1}^{q-1}$, so

$$\pi_{(0,\ldots,0,\frac{(n-1)(q-1)}{d})} X_f = \int_{\mathbb{F}_q^{n-1}} f^{\frac{(n-1)(q-1)}{d}} \neq 0.$$

d) When $n = 2$ and $d \mid q-1$, the conditions of part c) hold. □

When $d = 1$, the subset $X_f \subset \mathbb{F}_q^n$ is an affine $\mathbb{F}_q$-hyperplane, so it follows from Proposition 3.6a) that $\omega(X_f) = (n-1)(q-1)$, so we get another case in which the bound of Proposition 3.10a) is sharp.

References

[AH19] K. Abdukhalikov and D. Ho, *Vandermonde sets and hyperovals.* `https://arxiv.org/pdf/1911.10798.pdf`

[Al99] N. Alon, *Combinatorial Nullstellensatz.* Recent trends in combinatorics (Mátraháza, 1995). Combin. Probab. Comput. 8 (1999), 7–29.

[AM21] E. Aichinger and J. Moosbauer, *Chevalley-Warning type results on abelian groups.* J. Algebra 569 (2021), 30–66.

[Ax64] J. Ax, *Zeroes of polynomials over finite fields.* Amer. J. Math. 86 (1964), 255–261.

[BCP06] Bayens, L., Cherowitzo, W., Penttila, T. *Groups of hyperovals in Desarguesian planes.* Innov. Incidence Geom. 6/7 (2007/08), 37–51.

[Br11] D. Brink, *Chevalley's theorem with restricted variables.* Combinatorica 31 (2011), 127–130.

[CFS17] P.L. Clark, A. Forrow and J.R. Schmitt, *Warning's Second Theorem With Restricted Variables.* Combinatorica 37 (2017), 397–417.

[CGS21] P.L. Clark T. Genao and F. Saia, *Chevalley-Warning at the Boundary.* To appear, *Expositiones Math.*

[Ch35] C. Chevalley, *Démonstration d'une hypothèse de M. Artin.* Abh. Math. Sem. Univ. Hamburg 11 (1935), 73–75.

[Cl14] P.L. Clark, *The Combinatorial Nullstellensätze revisited.* Electron. J. Combin. 21 (2014), no. 4, Paper 4.15, 17 pp.

[GW03] A. Gács and Z. Weiner, *On $(q+t,t)$-arcs of type $(0,2,t)$.* Proceedings of the Conference on Finite Geometries (Oberwolfach, 2001). Des. Codes Cryptogr. 29 (2003), 131–139.

[KP12] R.N. Karasev and F.V. Petrov, *Partitions of nonzero elements of a finite field into pairs.* Israel J. Math. 192 (2012), 143–156.

[La10] M. Lasoń, *A generalization of combinatorial Nullstellensatz.* Electron. J. Combin. 17 (2010), no. 1, Note 32, 6 pp.

[LN97] R. Lidl and H. Niederreiter, *Finite fields.* With a foreword by P. M. Cohn. Second edition. Encyclopedia of Mathematics and its Applications, 20. Cambridge University Press, Cambridge, 1997.

[Ni21] B. Nica, *Polynomials over structured grids.* `https://arxiv.org/abs/2110.05616`

[Sc08] U. Schauz, *Algebraically solvable problems: describing polynomials as equivalent to explicit solutions.* Electron. J. Combin. 15 (2008), no. 1, Research Paper 10, 35 pp.

[ST08] P. Sziklai and M. Takáts, *Vandermonde sets and super-Vandermonde sets.* Finite Fields Appl. 14 (2008), 1056–1067.

[Va19] P. Vandendriessche, *Classification of the hyperovals in $PG(2,64)$.* Electron. J. Combin. 26 (2019), no. 2, Paper No. 2.35, 12 pp.

[Wa35] E. Warning, *Bemerkung zur vorstehenden Arbeit von Herrn Chevalley.* Abh. Math. Sem. Hamburg 11 (1935), 76–83.

[Ze84] D. Zeilberger, *A combinatorial proof of Newton's identities.* Discrete Math. 49 (1984), 319.