Combinatorics, Probability and Computing (2018) 27, 310–333.
doi:10.1017/S0963548317000566

On Zeros of a Polynomial in a Finite Grid

ANURAG BISHNOI[1], PETE L. CLARK[2], ADITYA POTUKUCHI[3] and JOHN R. SCHMITT[4]

[1] Freie Universität Berlin, Institut für Mathematik, Arnimallee 3, 14195 Berlin, Germany.
(e-mail: anurag.2357@gmail.com)
[2]Department of Mathematics, University of Georgia, Athens, GA 30605, USA
(e-mail: plclark@gmail.com)
[3]Department of Computer Science, Rutgers University, Piscataway, NJ 08854, USA
(e-mail: apotu.57@gmail.com)
[4]Department of Mathematics, Middlebury College, Middlebury, VT 05753, USA
(e-mail: jschmitt@middlebury.edu)

Received 3 August 2016; revised 30 March 2017; first published online 15 February 2018

A 1993 result of Alon and Füredi gives a sharp upper bound on the number of zeros of a multivariate polynomial over an integral domain in a finite grid, in terms of the degree of the polynomial. This result was recently generalized to polynomials over an arbitrary commutative ring, assuming a certain 'Condition (D)' on the grid which holds vacuously when the ring is a domain. In the first half of this paper we give a further generalized Alon–Füredi theorem which provides a sharp upper bound when the degrees of the polynomial in each variable are also taken into account. This yields in particular a new proof of Alon–Füredi. We then discuss the relationship between Alon–Füredi and results of DeMillo–Lipton, Schwartz and Zippel. A direct coding theoretic interpretation of Alon–Füredi theorem and its generalization in terms of Reed–Muller-type affine variety codes is shown, which gives us the minimum Hamming distance of these codes. Then we apply the Alon–Füredi theorem to quickly recover – and sometimes strengthen – old and new results in finite geometry, including the Jamison–Brouwer–Schrijver bound on affine blocking sets. We end with a discussion of multiplicity enhancements.

2010 *Mathematics subject classification*: Primary 05E40
Secondary 11T06, 11T71, 51E20, 51E21

1. Introduction

1.1. Notation

We denote the positive integers by $\mathbb{Z}^+$ and the non-negative integers by $\mathbb{N}$. For $n \in \mathbb{Z}^+$, we put $[n] = \{1, 2, \ldots, n\}$.

For us, rings are commutative with multiplicative identity. Throughout this paper R denotes a ring and F denotes a field, each arbitrary unless otherwise specified.

Following [12] and [32], a non-empty subset $S \subset R$ is said to satisfy *Condition (D)* if, for all $x \neq y \in S$, the element $x - y \in R$ is not a zero divisor. A *finite grid* is a subset $A = \prod_{i=1}^{n} A_i$ of R^n (for some $n \in \mathbb{Z}^+$) with each A_i a finite, non-empty subset of R. We say that A satisfies Condition (D) if each A_i does.

For $A \subset R^n$ and $f \in R[\underline{t}] = R[t_1, \ldots, t_n]$, we put

$$Z_A(f) = \{x \in A \mid f(x) = 0\} \quad \text{and} \quad \mathcal{U}_A(f) = \{x \in A \mid f(x) \neq 0\}.$$

1.2. The Alon–Füredi theorem

In [1] Alon and Füredi solved a problem posed by Bárány (based on a result of Komjath) of finding the minimum number of hyperplanes required to cover all points of the hypercube $\{0,1\}^n \subseteq F^n$ *except one*. One such covering is given by n hyperplanes defined by the zeros of the polynomials $t_1 - 1$, $t_2 - 1$, …, $t_n - 1$. Alon and Füredi proved that n is in fact the minimum number. They then generalized this result to all finite grids $A = \prod_{i=1}^{n} A_i \subset F^n$, showing that the minimum number of hyperplanes required to cover all points of A except one is $\sum_{i=1}^{n} (\#A_i - 1)$.

There is also a quantitative refinement: as we vary over families of d hyperplanes which do not cover all points of A, what is the minimum number of points which are missed? To answer this, Alon and Füredi proved the following result.

Theorem 1.1 (Alon–Füredi theorem [1, Theorem 5]). *Let F be a field, let $A = \prod_{i=1}^{n} A_i \subset F^n$ be a finite grid, and let $f \in F[\underline{t}] = F[t_1, \ldots, t_n]$ be a polynomial which does not vanish on all points of A. Then $f(x) \neq 0$ for at least $\min \prod y_i$ elements $x \in A$, where the minimum is taken over all positive integers $y_i \leqslant \#A_i$ with $\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} \#A_i - \deg f$. More concisely (see Section 2.1),*

$$\#\mathcal{U}_A(f) \geqslant \mathfrak{m}\left(\#A_1, \ldots, \#A_n; \sum_{i=1}^{n} \#A_i - \deg f\right).$$

The minimum referred to in Theorem 1.1 is known in all cases – see Lemma 2.1(a) – leading to an explicit form of the bound.

Several proofs of Theorem 1.1 have been given. The original argument in [1] involves the construction of auxiliary polynomial functions of low degree via linear algebra. A second proof was given by Ball and Serra as an application of their punctured combinatorial Nullstellensatz [3, 4]. Recently, López, Renterá-Márquez and Villarreal gave a proof of Alon–Füredi [27], in its coding theoretic formulation (see Section 5). Geil had noticed that the minimum distance of generalized Reed–Muller codes can be determined easily using the Gröbner basis theory [19, Theorem 2]. This technique was then used by Carvalho to give another proof of Theorem 1.1 when F is a finite field [9, Proposition 2.3], which is in fact a special case of an earlier result by Geil and Thomsen [20, Proposition 5] (take all weights equal to 1).

In [13], Clark generalized the Alon–Füredi theorem by replacing the field F with an arbitrary ring R, under the assumption that the finite grid A satisfies Condition (D). This is a modest generalization in that Condition (D) is exactly what is needed for polynomial functions on A to behave as they do in the case of a field, and the proof adapts that of Ball and Serra.

Clark, Forrow and Schmitt [14] used Alon–Füredi to obtain a restricted variable generalization of a theorem of Warning [34] giving a lower bound on the number of zeros of a system of

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

polynomials over a finite field. (Alon–Füredi gives a lower bound on *non-zeros*, but over a finite field $\mathbb{F}_q$, we have Chevalley's trick: $f(x) = 0 \iff 1 - f(x)^{q-1} \neq 0$.) This work also gave several combinatorial applications of this lower bound on restricted variable zero sets.

One of the main goals of this paper is to revisit the Alon–Füredi theorem and give direct combinatorial applications (*i.e.* not of Chevalley–Warning type). We begin by giving the following generalization of the Alon–Füredi theorem.

Theorem 1.2 (generalized Alon–Füredi theorem). *Let R be a ring and let $A_1, \dots, A_n$ be non-empty finite subsets of R that satisfy Condition (D). For $i \in [n]$, let b_i be an integer such that $1 \leqslant b_i \leqslant \#A_i$. Let $f \in R[t_1, \dots, t_n]$ be a non-zero polynomial such that $\deg_{t_i} f \leqslant \#A_i - b_i$ for all $i \in [n]$. Let $\mathcal{U}_A(f) = \{x \in A \mid f(x) \neq 0\}$ where $A = A_1 \times \cdots \times A_n \subset R^n$. Then we have (see Section 2.1)*

$$\#\mathcal{U}_A(f) \geqslant \mathfrak{m}\left(\#A_1, \dots, \#A_n; b_1, \dots, b_n; \sum_{i=1}^{n} \#A_i - \deg f\right).$$

Moreover, this bound is sharp in all cases.

As we shall explain in Section 2.3, one recovers Theorem 1.1 from Theorem 1.2 by taking $b_1 = \cdots = b_n = 1$. Our argument specializes to give a new proof of Alon–Füredi.

In Section 4 we relate the generalized Alon–Füredi theorem to work of DeMillo–Lipton, Schwartz and Zippel. We find in particular that Alon–Füredi implies the result which has become known as the 'Schwartz–Zippel lemma'. In fact, the original result of Zippel (and earlier, DeMillo–Lipton) is a bit different and not implied by Alon–Füredi (see Example 4.7). However, it is implied by generalized Alon–Füredi, and this was one of our motivations for strengthening Alon–Füredi as we have.

The Alon–Füredi theorem has a natural coding theoretic interpretation (see Section 5) as it computes the minimum Hamming distance of the affine grid code $\mathrm{AGC}_d(A)$, an F-linear code of length $\#A$. In this way Alon–Füredi turns out to be the restricted variable generalization of a much older result in the case $A_i = F = \mathbb{F}_q$, the Kasami–Lin–Peterson theorem, which computes the minimum Hamming distance of generalized Reed–Muller codes. We will show that the generalized Alon–Füredi theorem is equivalent to computing the minimum Hamming distance of a more general class of R-linear codes. These generalized affine grid codes have larger distance (though also smaller dimension) than the standard ones, so they may turn out to be useful.

In Section 6, we pursue applications to finite geometry. We begin by revisiting and slightly sharpening the original result of Alon–Füredi on hyperplane coverings. This naturally leads us to partial covers and blocking sets in affine and projective geometries over $\mathbb{F}_q$. Applying Alon–Füredi and projective duality we get a new upper bound, Theorem 6.6, on the number of hyperplanes which do not meet a k-element subset of $\mathrm{AG}(n,q)$. From this result the classical theorems of Jamison–Brouwer–Schrijver on affine blocking sets and Blokhuis–Brouwer on essential points of projective blocking sets follow as corollaries. We are also able to strengthen a recent result of Dodunekov, Storme and Van de Voorde.

Finally, in Section 7 we discuss multiplicity enhancements in the sense of [18]. The material here is most closely related to that of Section 4, but we have placed it at the end because it has a somewhat more technical character than the rest of the paper.

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

2. Preliminaries

2.1. Balls in prefilled bins

Let $a_1,\dots,a_n \in \mathbb{Z}^+$. Consider n bins $A_1,\dots,A_n$ such that A_i can hold up to a_i balls. For $N \in \mathbb{Z}^+$ with $n \leqslant N \leqslant \sum_{i=1}^n a_i$, we define a distribution of N balls in these n bins to be an n-tuple $y = (y_1,\dots,y_n) \in (\mathbb{Z}^+)^n$ with $y_i \leqslant a_i$ for all $i \in [n]$ and $\sum_{i=1}^n y_i = N$. For a distribution y of N balls in n bins, we put $P(y) = \prod_{i=1}^n y_i$. For $n \leqslant N \leqslant \sum_{i=1}^n a_i$ we define $\mathfrak{m}(a_1,\dots,a_n;N)$ to be the minimum value of $P(y)$ as y ranges over all such distributions of N balls in n bins. For $N < n$ we define $\mathfrak{m}(a_1,\dots,a_n;N) = 1$.

Without loss of generality we may – and shall – assume $a_1 \geqslant \cdots \geqslant a_n$. We define the *greedy distribution* $y_G = (y_1,\dots,y_n)$ as follows: first place one ball in each bin; then place the remaining balls into bins from left to right, filling each bin completely before moving on to the next bin, until we run out of balls.

Lemma 2.1. *Let $n \in \mathbb{Z}^+$, and let $a_1 \geqslant \cdots \geqslant a_n$ be positive integers. Let $N \in \mathbb{Z}$ with $n \leqslant N \leqslant a_1 + \cdots + a_n$.*

(a) *We have*

$$\mathfrak{m}(a_1,\dots,a_n;N) = P(y_G) = y_1 \cdots y_n.$$

(b) *Suppose $a_1 = \cdots = a_n = a \geqslant 2$. Then*

$$\mathfrak{m}(a,\dots,a;N) = (r+1)a^{\lfloor (N-n)/(a-1)\rfloor},$$

where $r \equiv N-n \pmod{a-1}$ and $0 \leqslant r < a-1$.

(c) *For all non-negative integers k, we have*

$$\mathfrak{m}(2,\dots,2;2n-k) = 2^{n-k}.$$

(d) *Let $n, a_1,\dots,a_n \in \mathbb{Z}^+$ with $a_1 \geqslant \cdots \geqslant a_n$. Let $N \in \mathbb{Z}$ be such that $N-n = \sum_{i=1}^j (a_i - 1) + r$ for some $j \in \{0,\dots,n\}$ and some r satisfying $0 \leqslant r < a_{j+1}$. Then $\mathfrak{m}(a_1,\dots,a_n;N) = (r+1)\prod_{i=1}^j a_i$.*

Proof. Parts (a)–(c) are [14, Lemma 2.2]. (d) After placing one ball in each bin we are left with $N-n$ balls. The greedy distribution is achieved by filling the first j bins entirely and then putting r balls in bin $j+1$. □

In every distribution $y = (y_1,\dots,y_n)$ we need $y_i \geqslant 1$ for all $i \in [n]$; that is, we must place at least one ball in each bin. So it is reasonable to think of the bins coming *prefilled* with one ball each, and then our task is to distribute the $N-n$ remaining balls so as to minimize $P(y)$. The concept of prefilled bins makes sense more generally: given any $b_1,\dots,b_n \in \mathbb{Z}$ with $1 \leqslant b_i \leqslant a_i$, we may consider the scenario in which the ith bin comes prefilled with b_i balls. If $\sum_{i=1}^n b_i \leqslant N \leqslant \sum_{i=1}^n a_i$, we may restrict to distributions $y = (y_1,\dots,y_n)$ of N balls into bins of sizes $a_1,\dots,a_n$ such that $b_i \leqslant y_i \leqslant a_i$ for all $i \in [n]$, and put

$$\mathfrak{m}(a_1,\dots,a_n;b_1,\dots,b_n;N) = \min P(y),$$

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

where the minimum ranges over this restricted set of distributions. For $N < \sum_{i=1}^n b_i$ we define $\mathfrak{m}(a_1,\dots,a_n;b_1,\dots,b_n;N) := \prod_{i=1}^n b_i$.

Lemma 2.2. *We have*

$$\mathfrak{m}(a_1,\dots,a_n;b_1,\dots,b_n;N) = \prod_{i=1}^n b_i \iff N \leqslant \sum_{i=1}^n b_i.$$

Proof. If $N \leqslant \sum_{i=1}^n b_i$ then

$$\mathfrak{m}(a_1,\dots,a_n;b_1,\dots,b_n;N) = \prod_{i=1}^n b_i$$

by definition unless $N = \sum_{i=1}^n b_i$, and this case is immediate: we have exactly enough balls to perform the prefilling. If $N > \sum_{i=1}^n b_i$, then $\mathfrak{m}(a_1,\dots,a_n;b_1,\dots,b_n;N)$ is the minimum over a set of integers each of which is strictly greater than $\prod_{i=1}^n b_i$. □

In this prefilled context, the greedy distribution y_G is defined by starting with the bins prefilled with $b_1,\dots,b_n$ balls and then distributing the remaining balls from left to right, filling each bin completely before moving on to the next bin. One sees – for example by adapting the argument of [14, Lemma 2.2] – that

$$\mathfrak{m}(a_1,\dots,a_n;b_1,\dots,b_n;N) = P(y_G),$$

when we also have $b_1 \geqslant \cdots \geqslant b_n$. But this may not hold in general, as the following example shows.

Example 2.3. *Let*

$$n = 2, \quad a_1 = 4, \quad a_2 = 3, \quad b_1 = 1, \quad b_2 = 2, \quad N = 4.$$

Then $P(y_G) = 4$ *but* $\mathfrak{m}(4,3;1,2;4) = 3$ *is achieved by the distribution* $(1,3)$.

In general we do not know a simple description of $\mathfrak{m}(a_1,\dots,a_n;b_1,\dots,b_n;N)$. In practice, it can be computed using dynamic programming.

Lemma 2.4. *Let* $a_1,\dots,a_n,b_1,\dots,b_n \in \mathbb{Z}^+$ *with* $1 \leqslant b_i \leqslant a_i$ *for all* $i \in [n]$. *Let* $k \in \mathbb{Z}$ *such that* $b_n \leqslant k \leqslant a_n$. *If*

$$b_1 + \cdots + b_{n-1} \leqslant N - k \leqslant a_1 + \cdots + a_{n-1}$$

for some $N \in \mathbb{Z}$, *then*

$$k \cdot \mathfrak{m}(a_1,\dots,a_{n-1};b_1,\dots,b_{n-1};N-k) \geqslant \mathfrak{m}(a_1,\dots,a_n;b_1,\dots,b_n;N).$$

Proof. Let $y' = (y_1,\dots,y_{n-1})$ be a distribution of $N-k$ balls in the first $n-1$ bins. Then $y = (y_1,\dots,y_{n-1},k)$ is a distribution of N balls in n bins with the last bin getting k balls. Therefore,

$$\mathfrak{m}(a_1,\dots,a_n;b_1,\dots,b_n;N) \leqslant P(y) = k \cdot P(y').$$

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

Since this holds for all such distributions y', we get

$$\mathfrak{m}(a_1,\dots,a_n;b_1,\dots,b_n;N) \leqslant k\cdot\mathfrak{m}(a_1,\dots,a_{n-1};b_1,\dots,b_{n-1};N-k). \qquad \square$$

2.2. Grid reduction and Condition (D)

For any finite grid $A \subset R^n$, evaluation of a polynomial $f \in R[\underline{t}] = R[t_1,\dots,t_n]$ at elements of A gives a ring homomorphism,

$$E_A : R[\underline{t}] \to R^A,\ f \mapsto (x \in A \mapsto f(x)).$$

Let $I(A)$ be the kernel of E_A, that is, the set of polynomials which vanish identically on A. There are some 'obvious' elements of $I(A)$, namely

$$\varphi_i = \prod_{x_i \in A_i} (t_i - x_i), \quad \text{for all } i \in [n].$$

Let $\Phi = \langle \varphi_1,\dots,\varphi_n\rangle$ be the ideal generated by these elements. Then $\Phi \subset I(A)$.

We say a polynomial $f \in R[\underline{t}]$ is *A-reduced* if $\deg_{t_i}(f) < \#A_i$ for all $i \in [n]$. The A-reduced polynomials form an R-submodule $\mathcal{R}_A$ of $R[\underline{t}]$ which is free of rank $\#A$, and the composite map

$$\mathcal{R}_A \hookrightarrow R[\underline{t}] \to R[\underline{t}]/\Phi$$

is an R-module isomorphism [12, Proposition 10], that is, every polynomial $f \in R[\underline{t}]$ differs from a unique A-reduced polynomial $r_A(f)$ by an element of Φ, and we have $E_A(f) = E_A(r_A(f))$. The polynomial $r_A(f)$ can be computed from f by dividing by φ_1, then dividing the remainder by φ_2, and so on. It follows that $\deg r_A(f) \leqslant \deg f$ and $\deg_{t_i} r_A(f) \leqslant \deg_{t_i} f$ for all $i \in [n]$.

Theorem 2.5 (CATS† lemma [12, Theorem 12]). *The following are equivalent.*

(i) *The finite grid A satisfies Condition (D).*
(ii) *If $f \in \mathcal{R}_A$ and $f(x) = 0$ for all $x \in A$, then $f = 0$.*
(iii) *We have $\Phi = I(A)$.*

Remark. These results can be directly used to solve the main problem studied by Alon and Füredi. Let f be a polynomial that vanishes on all points of A except the point $x = (x_1,\dots,x_n)$. Since the polynomial $\prod_{i=1}^n \prod_{\lambda \in A_i \setminus \{x_i\}} (t_i - \lambda)$ is A-reduced and it vanishes everywhere on A except at x, it must be equal to $r_A(f)$. Thus,

$$\deg f \geqslant \deg r_A(f) = \sum_{i=1}^n (\#A_i - 1).$$

Now associate the set of hyperplanes that cover all points of A except one by the product of their corresponding linear polynomials.

† CATS = Chevalley–Alon–Tarsi–Schauz [2, 10, 32].

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

2.3. Generalized Alon–Füredi implies Alon–Füredi

Let $A = \prod_{i=1}^n A_i \subset R^n$ be a finite grid satisfying Condition (D), and for $i \in [n]$ put $a_i = \#A_i$. Suppose $f \in R[t]$ does not vanish identically on A. Let

$$\mathcal{U}_A(f) = \{x \in A \mid f(x) \neq 0\}.$$

Then the Alon–Füredi theorem is the assertion that

$$\#\mathcal{U}_A(f) \geqslant \mathfrak{m}\left(a_1, \ldots, a_n; \sum_{i=1}^n a_i - \deg f\right).$$

The non-vanishing hypothesis on f is equivalent to $r_A(f) \neq 0$. Then $r_A(f)$ satisfies the hypotheses of Theorem 1.2 with $b_1 = \cdots = b_n = 1$. Since $E_A(f) = E_A(r_A(f))$, we have $\mathcal{U}_A(f) = \mathcal{U}_A(r_A(f))$, and thus

$$\begin{aligned}\#\mathcal{U}_A(f) &\geqslant \mathfrak{m}\left(a_1, \ldots, a_n; 1, \ldots, 1; \sum_{i=1}^n a_i - \deg r_A(f)\right)\\ &= \mathfrak{m}\left(a_1, \ldots, a_n; \sum_{i=1}^n a_i - \deg r_A(f)\right) \geqslant \mathfrak{m}\left(a_1, \ldots, a_n; \sum_{i=1}^n a_i - \deg f\right).\end{aligned}$$

3. Proof of the generalized Alon–Füredi theorem

3.1. A preliminary remark

If f satisfies the hypotheses of the generalized Alon–Füredi theorem, then

$$\deg f \leqslant \sum_{i=1}^n \deg_{t_i} f \leqslant \sum_{i=1}^n (a_i - b_i),$$

so

$$\sum_{i=1}^n b_i \leqslant \sum_{i=1}^n a_i - \deg f \leqslant \sum_{i=1}^n a_i.$$

Thus, whereas the conventional Alon–Füredi setup allows the case in which we have too few balls to fill the bins (in which case the result gives the trivial (but sharp!) bound $\#\mathcal{U}_A(f) \geqslant 1$), in our setup we do not need to consider this case.

3.2. Proof of the generalized Alon–Füredi bound

For $i \in [n]$, put $a_i = \#A_i$. We go by induction on n.

Base case. Let $f \in R[t_1]$ be a non-zero polynomial. Suppose f vanishes precisely at the distinct elements $x_1, \ldots, x_k$ of A_1. Dividing f by $t_1 - x_1$ shows $f = (t_1 - x_1)f_1(t_1)$, and (since A_1 satisfies Condition (D)) $f_1(x_i) = 0$ for $2 \leqslant i \leqslant k$. Continuing in this way we get $f = \prod_{i=1}^k (t_1 - x_i) f_k(t_1)$, and thus $\deg f \geqslant k$. So

$$\#\mathcal{U}_A(f) = a_1 - k \geqslant a_1 - \deg f,$$

which is the conclusion of the generalized Alon–Füredi theorem in this case.

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

Induction step. Suppose $n \geqslant 2$ and the result holds for $n-1$. Write

$$f(t_1,\dots,t_n) = \sum_{i=0}^{d_n} f_i(t_1,\dots,t_{n-1})t_n^i,$$

so that $d_n = \deg_{t_n} f$ is the largest index i such that $f_i \neq 0$. Moreover, we have $\deg f_{d_n} \leqslant \deg f - d_n$, and for all $i \in [n-1], \deg_{t_i} f_{d_n} \leqslant \deg_{t_i} f \leqslant a_i - b_i$.

Put $A' = \prod_{i=1}^{n-1} A_i$. By the induction hypothesis, we have

$$\begin{aligned}\#\mathcal{U}_{A'}(f_{d_n}) &\geqslant \mathfrak{m}\left(a_1,\dots,a_{n-1};b_1,\dots,b_{n-1};\sum_{i=1}^{n-1}a_i - \deg f_{d_n}\right)\\ &\geqslant \mathfrak{m}\left(a_1,\dots,a_{n-1};b_1,\dots,b_{n-1};\sum_{i=1}^{n-1}a_i - \deg f + d_n\right).\end{aligned}$$

Let $x' = (x_1,\dots,x_{n-1}) \in \mathcal{U}_{A'}(f_{d_n})$. Then $f(x',t_n) = \sum_{i=0}^{d_n} f_i(x')t_n^i \in R[t_n]$ has degree $d_n \geqslant 0$ since its leading term $f_{d_n}(x')t^{d_n}$ is non-zero. Since A_n satisfies Condition (D), $f(x',t_n)$ vanishes at no more than d_n points of A_n, so there are at least $a_n - d_n$ elements $x_n \in A_n$ such that $(x',x_n) \in \mathcal{U}_A(f)$. Thus

$$\#\mathcal{U}_A(f) \geqslant (a_n - d_n)\mathfrak{m}\left(a_1,\dots,a_{n-1};b_1,\dots,b_{n-1};\sum_{i=1}^{n-1}a_i - \deg f + d_n\right).$$

Since

$$\deg f \leqslant \sum_{i=1}^{n}\deg_{t_i} f = \sum_{i=1}^{n-1}\deg_{t_i} f + d_n$$

and thus

$$\sum_{i=1}^{n-1} b_i \leqslant \sum_{i=1}^{n-1}(a_i - \deg_{t_i} f) \leqslant \sum_{i=1}^{n-1} a_i - \deg f + d_n \leqslant \sum_{i=1}^{n-1} a_i,$$

we may apply Lemma 2.4 with $N = \sum_{i=1}^{n} a_i - \deg f$ and $k = a_n - d_n$, getting

$$\begin{aligned}&(a_n - d_n)\mathfrak{m}\left(a_1,\dots,a_{n-1};b_1,\dots,b_{n-1};\sum_{i=1}^{n-1}a_i - \deg f + d_n\right)\\ &\quad\geqslant \mathfrak{m}\left(a_1,\dots,a_n;b_1,\dots,b_n;\sum_{i=1}^{n}a_i - \deg f\right).\end{aligned}$$

We deduce that

$$\#\mathcal{U}_A(f) \geqslant \mathfrak{m}\left(a_1,\dots,a_n;b_1,\dots,b_n;\sum_{i=1}^{n}a_i - \deg f\right).$$

3.3. Sharpness of the generalized Alon–Füredi bound

For $i \in [n]$, put $a_i = \#A_i$, and let d be an integer such that $0 \leqslant d \leqslant \sum_{i=1}^{n}(a_i - b_i)$ (see Section 3.1). For any distribution $y = (y_1,\dots,y_n)$ of $\sum_{i=1}^{n} a_i - d$ balls in n bins with $b_i \leqslant y_i \leqslant a_i$, for all $i \in [n]$

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

choose a subset $S_i \subset A_i$ of cardinality $a_i - y_i$, and put[‡]

$$f(t) = \prod_{i=1}^{n} \prod_{x_i \in S_i} (t_i - x_i).$$

Then

$$\deg f = \sum_{i=1}^{n} (a_i - y_i) = d,$$

$$\text{for all } i \in [n], \quad \deg_{t_i} f = a_i - y_i \leqslant a_i - b_i$$

and

$$\#\mathcal{U}_A(f) = P(y) = \prod_{i=1}^{n} y_i.$$

Thus, for all finite grids $A = \prod_{i=1}^{n} A_i$ satisfying Condition (D) and all permissible values of $\deg_{t_1} f, \ldots, \deg_{t_n} f$ and $\deg f$, there are instances of equality in the generalized Alon–Füredi bound. The case $b_1 = \cdots = b_n = 1$ yields the (known) sharpness of the Alon–Füredi bound.

3.4. An equivalent formulation

Let us say that a polynomial $f \in R[\underline{t}]$ is *polylinear* (resp. *simple polylinear*) if it is a product of factors (resp. distinct factors) of the form $t_i - x$ for $x \in R$. Petrov has observed (personal communication) that the generalized Alon–Füredi theorem is equivalent to the statement that for any non-zero A-reduced polynomial $f \in R[\underline{t}]$, there is a simple polylinear polynomial $g \in R[\underline{t}]$ with $\deg_{t_i} f = \deg_{t_i} g$ for all $i \in [n]$, $\deg f = \deg g$ and such that $\#Z_A(f) \leqslant \#Z_A(g)$. Thus it is possible to formulate the result without reference to balls in prefilled bins. However, as we will see, having the result in this form is useful for applications.

4. Connections with the Schwartz–Zippel lemma

4.1. Schwartz–Zippel lemma

The material in this section is motivated by a blog post of Lipton [26] which discusses the history of the 'Schwartz–Zippel lemma'. We will further weigh in on the history of this circle of results, discuss various improvements and give the connection to the Alon–Füredi theorem.

Theorem 4.1 (Schwartz–Zippel lemma). *Let R be a domain and let $S \subset R$ be finite and non-empty. Let $f \in R[\underline{t}] = R[t_1, \ldots, t_n]$ be a non-zero polynomial. Then*

$$\#Z_{S^n}(f) \leqslant (\deg f)(\#S)^{n-1}. \tag{4.1}$$

Proof. Let $s = \#S$. The statement is equivalent to

$$\#\mathcal{U}_{S^n}(f) \geqslant s^{n-1}(s - \deg f).$$

[‡] An empty product is understood to take the value 1.

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

If $\deg f \geqslant s$, then (4.1) asserts that f has no more zeros in S^n than the size of S^n: true. So the non-trivial case is $\deg f < s$. Then f is S^n-reduced, so

$$\#\mathcal{U}_{S^n}(f) \geqslant \mathfrak{m}(s,\dots,s;ns-\deg f) = s^{n-1}(s-\deg f)$$

by Alon–Füredi and because the greedy distribution is $(s,\dots,s,s-\deg f)$. □

The case of the Schwartz–Zippel lemma in which $R = S = \mathbb{F}_q$ is due to Ore [30]. Thus the Schwartz–Zippel lemma may be viewed as a 'restricted variable Ore theorem', although it is not the most general result along those lines. In fact, the same argument establishes the following.

Theorem 4.2 (generalized Schwartz–Zippel lemma). *Let $A = \prod_{i=1}^n A_i \subset R^n$ be a finite grid satisfying Condition (D), and suppose $\#A_1 \geqslant \cdots \geqslant \#A_n$. Let $f \in R[\underline{t}] = R[t_1,\dots,t_n]$ be a non-zero polynomial. Then*

$$\#Z_A(f) \leqslant (\deg f)\prod_{i=1}^{n-1}\#A_i.$$

Indeed, the non-trivial case of this result is precisely the case $\#A_1 \geqslant \cdots \geqslant \#A_n > \deg f$ of Alon–Füredi, and the greedy distribution is $(\#A_1,\dots,\#A_{n-1},\#A_n-\deg f)$.

4.2. Schwartz's theorem

The Schwartz–Zippel lemma appears in Schwartz's 1980 paper as a corollary of a more general upper bound on zeros of a polynomial over a domain [33, Corollary 1]. We give a version over an arbitrary ring.

Theorem 4.3 (Schwartz theorem [33, Lemma 1]). *Let $f = f_n \in R[t_1,\dots,t_n]$ be a non-zero polynomial and let $d_n = \deg_{t_n} f_n$. Let $f_{n-1} \in R[t_1,\dots,t_{n-1}]$ be the coefficient of $t_n^{d_n}$ in f_n. Let $d_{n-1} = \deg_{t_{n-1}} f_{n-1}$, and let $f_{n-2} \in R[t_1,\dots,t_{n-2}]$ be the coefficient of $t_{n-1}^{d_{n-1}}$ in f_{n-1}. Continuing in this manner we define for all $1 \leqslant i \leqslant n$ a polynomial $f_i \in R[t_1,\dots,t_i]$ with $\deg_{t_i} f_i = d_i$. Let $A = \prod_{i=1}^n A_i$ be a finite grid satisfying Condition (D). Then*

$$\#Z_A(f) \leqslant \#A\sum_{i=1}^n \frac{d_i}{\#A_i}.$$

Proof. For $i \in [n]$, put $a_i = \#A_i$. We go by induction on n. The base case is the same as that of Theorem 1.2: essentially the root-factor phenomenon of high school algebra, used with some care because R need not be a domain. Inductively we suppose the result holds for polynomials in $n-1$ variables and in particular for $f_{n-1} \in R[t_1,\dots,t_{n-1}]$ and $A' = \prod_{i=1}^{n-1} A_i$. Let $x' = (x_1,\dots,x_{n-1}) \in A'$. If $f_{n-1}(x') = 0$, it may be the case that $f_n(x',x_n) = 0$ for all $x_n \in A_n$. But if not, then $f_n(x',t_n) \in R[t_n]$ has at most d_n zeros in A_n. Thus the number of zeros of $f = f_n$ in A is at most

$$\#A_n \cdot \#A'\left(\sum_{i=1}^{n-1}\frac{d_i}{a_i}\right) + d_n\#A' = \#A\sum_{i=1}^n \frac{d_i}{a_i}.$$

□

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

Proposition 4.4. *Theorem 4.3 implies Theorem 4.2.*

Proof. The coefficient of $t_1^{d_1}\cdots t_n^{d_n}$ in f is non-zero, so $\sum_{i=1}^n d_i \leqslant \deg f$, and thus

$$\#Z_A(f) \leqslant \#A \sum_{i=1}^n \frac{d_i}{\#A_i} \leqslant (\#A_1 \cdots \#A_{n-1}) \sum_{i=1}^n d_i \leqslant (\deg f) \prod_{i=1}^{n-1} \#A_i. \qquad \square$$

4.3. DeMillo–Lipton and Zippel

The following result was proved by DeMillo and Lipton in [16] and then independently by Zippel [35].

Theorem 4.5 (DeMillo–Lipton–Zippel theorem). *Let R be a domain, let*

$$f \in R[\underline{t}] = R[t_1, \ldots, t_n]$$

be a non-zero polynomial, and let $d \in \mathbb{Z}^+$ be such that $\deg_{t_i} f \leqslant d$ for all $i \in [n]$. Let $S \subset R$ be a non-empty set with more than d elements. Then

$$\#Z_{S^n}(f) \leqslant (\#S)^n - (\#S - d)^n.$$

Proof. Put $s = \#S$. We go by induction on n, and the $n = 1$ case is by now familiar. Assume the result for $n-1$. Since $\deg_{t_n} f \leqslant d$, we have

$$\#\{x_n \in S \mid f(t_1, \ldots, t_{n-1}, x_n) = 0\} \leqslant d,$$

so there are at least $s - d$ values of x_n such that $g = f(t_1, \ldots, t_{n-1}, x_n)$ is a non-zero polynomial. By induction, g has at most $s^{n-1} - (s-d)^{n-1}$ zeros on S^{n-1}. So

$$\begin{aligned} \#Z_{S^n}(f) &\leqslant ds^{n-1} + (s-d)(s^{n-1} - (s-d)^{n-1}) \\ &= ds^{n-1} + s^n - ds^{n-1} - (s-d)^n = s^n - (s-d)^n. \end{aligned} \qquad \square$$

Just like the Schwartz–Zippel lemma, a stronger version of the DeMillo–Lipton–Zippel theorem can be proved with essentially the same argument. We leave the proof – or rather this proof – to the reader.

Theorem 4.6 (generalized DeMillo–Lipton–Zippel theorem). *Let R be a ring, let*

$$f \in R[t_1, \ldots, t_n]$$

be a non-zero polynomial, and for $i \in [n]$ put $d_i = \deg_{t_i} f$. Let $A = \prod_{i=1}^n A_i$ be a finite grid satisfying Condition (D). We suppose that $1 \leqslant d_i < a_i$ for all $i \in [n]$. Then

$$\#\mathcal{U}_A(f) \geqslant \prod_{i=1}^n (\#A_i - d_i).$$

Now for a somewhat unsettling remark: the DeMillo–Lipton–Zippel theorem does not imply the Schwartz–Zippel lemma nor is it implied by any of Schwartz's results!

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

Example 4.7. *Let S be a finite subset of R containing 0, satisfying Condition (D), and of size $s \geqslant 3$. Let $f = t_1 t_2 \in R[t_1, t_2]$. Then we have*

$$\#Z_{S^2}(f) = 2s - 1.$$

DeMillo–Lipton–Zippel gives

$$\#Z_{S^2}(f) \leqslant s^2 - (s-1)^2 = 2s - 1.$$

Schwartz's theorem gives

$$\#Z_{S^2}(f) \leqslant s^2 \left(\frac{1}{s} + \frac{1}{s} \right) = 2s.$$

The Alon–Füredi theorem gives

$$\#Z_{S^2}(f) \leqslant s^2 - \mathfrak{m}(s, s; 2s-2) = s^2 - s(s-2) = 2s.$$

Thus neither Alon–Füredi nor Schwartz implies DeMillo–Lipton–Zippel. For the other direction, take $f = t_1 + t_2$. DeMillo–Lipton–Zippel gives $\#Z_{S^2}(f) \leqslant s^2 - (s-1)^2 = 2s - 1$, while the other results give $\#Z_{S^2}(f) \leqslant s$.

But we can still relate Schwartz–Zippel and DeMillo–Lipton–Zippel as follows.

Proposition. *Generalized Alon–Füredi implies generalized DeMillo–Lipton–Zippel.*

Proof. For $i \in [n]$, put $a_i = \#A_i$ and $b_i = a_i - d_i$, so $1 \leqslant b_i \leqslant a_i$ for all $i \in [n]$. Moreover, $\deg f \leqslant \sum_{i=1}^n d_i$, so the generalized Alon–Füredi theorem gives

$$\begin{aligned} \#\mathcal{U}_A &\geqslant \mathfrak{m}\left(a_1, \ldots, a_n; b_1, \ldots, b_n; \sum_{i=1}^n a_i - \deg f \right) \geqslant \mathfrak{m}\left(a_1 \ldots, a_n; b_1, \ldots, b_n; \sum_{i=1}^n (a_i - d_i) \right) \\ &= \mathfrak{m}\left(a_1, \ldots, a_n; b_1, \ldots, b_n; \sum_{i=1}^n b_i \right) = \prod_{i=1}^n b_i = \prod_{i=1}^n (a_i - d_i). \end{aligned}$$

□

Generalized DeMillo–Lipton–Zippel is equivalent to the case $\deg f = \sum_{i=1}^n \deg_{t_i} f$ of generalized Alon–Füredi. In particular, the bound is sharp in every case.

5. Connections with coding theory

In this section we make use of some terminology (only) from coding theory. Definitions can be found in [25, Chapter 3], for example.

Consider the polynomial ring $\mathbb{F}_q[\underline{t}] = \mathbb{F}_q[t_1, \ldots, t_n]$. Since $\mathbb{F}_q$ is finite, we can take $\mathbb{F}_q^n$ itself as a finite grid, and in fact many aspects of the theory presented here were worked out in this case in the early part of the twentieth century. In particular, we say a polynomial is *reduced* if it is $\mathbb{F}_q^n$-reduced, and this notion was introduced by Chevalley in his seminal work [10] on polynomial systems of low degree. We denote the the set of reduced polynomials by $\mathcal{P}(n, q)$; it is an $\mathbb{F}_q$-vector space of dimension q^n. The evaluation map gives an $\mathbb{F}_q$-linear isomorphism

$$E : \mathcal{P}(n, q) \to \mathbb{F}_q^{\mathbb{F}_q^n}, \; f \mapsto (x \in \mathbb{F}_q^n \mapsto f(x)).$$

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

Fixing an ordering $\alpha_1, \ldots, \alpha_{q^n}$ of $\mathbb{F}_q^n$, this isomorphism allows us to identify each $f \in \mathcal{P}(n,q)$ with its *value table* $(f(\alpha_1), \ldots, f(\alpha_{q^n}))$. For $d \in \mathbb{N}$ we let $\mathcal{P}_d(n,q)$ denote the set of all reduced polynomials of degree at most d.

Definition. The set of all value tables of all polynomials in $\mathcal{P}_d(n,q)$ is called the dth-order generalized Reed–Muller code of length q^n, denoted by $\mathrm{GRM}_d(n,q)$.

For $q = 2$, these codes were introduced and studied by Muller [29] and Reed [31]. An explicit formula for minimum distance of the generalized Reed–Muller codes was given by Kasami, Lin and Peterson [24], which we will recover using Alon–Füredi. A systematic study of these codes in terms of the polynomial formulation was conducted by Delsarte, Goethals and MacWilliams in [15], where they also classified all the minimum weight codewords.

Theorem 5.1 (Kasami, Lin and Peterson). *The minimum weight of the dth-order generalized Reed–Muller code $\mathrm{GRM}_d(n,q)$ is equal to $(q-b)q^{n-a-1}$, where $d = a(q-1)+b$ with $0 < b \leqslant q-1$.*

Proof. The minimum weight of $\mathrm{GRM}_d(n,q)$ is equal to the least number of non-zero values taken by a non-zero reduced polynomial of degree at most d, which by Alon–Füredi is $\mathfrak{m}(q, \ldots, q; nq-d)$. Moreover, we have

$$(nq-d)-n = n(q-1)-a(q-1)-b = (n-a-1)(q-1)+q-1-b,$$

and

$$0 \leqslant q-1-b < q-1,$$

so by Lemma 2.1 we have

$$\mathfrak{m}(q, \ldots, q; nq-d) = (q-b)q^{n-a-1}. \qquad \square$$

The generalized Alon–Füredi theorem can also be stated in terms of coding theory. Let $A = \prod_{i=1}^n A_i$ be a finite grid in a ring R^n satisfying Condition (D), with $a_i = \#A_i$ for $i \in [n]$. Given positive integers $b_i \leqslant a_i$ for all $i \in [n]$, and a natural number $d \leqslant \sum_{i=1}^n (a_i - b_i)$, we define the *generalized affine grid code* $\mathrm{GAGC}_d(A; b_1, \ldots, b_n)$ as the set of value tables of all polynomials $f \in R[\underline{t}]$ with $\deg_{t_i} f \leqslant a_i - b_i$ for all $i \in [n]$ and $\deg f \leqslant d$ evaluated on A. We put

$$\mathrm{AGC}_d(A) = \mathrm{GAGC}_d(A; 1, \ldots, 1)$$

and speak of *affine grid codes*.

Theorem 5.2. *The minimum weight of* $\mathrm{GAGC}_d(A; b_1, \ldots, b_n)$ *is*

$$\mathfrak{m}\left(a_1, \ldots, a_n; b_1, \ldots, b_n; \sum_{i=1}^n a_i - d\right).$$

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

Affine grid codes were studied in [27] (under the name of affine Cartesian codes) where they proved the following result.

Theorem 5.3 ([27, Theorem 3.8]). *Let F be a field and $A = \prod_{i=1}^{n} A_i \subset F^n$ a finite grid with $\#A_1 \geqslant \cdots \geqslant \#A_n \geqslant 1$. Then the minimum weight of* $\mathrm{AGC}_d(A)$ *is*

$$\begin{cases} \#A_1 \cdots \#A_{k-1}(\#A_k - \ell) & \text{if } d \leqslant \sum_{i=1}^{n}(\#A_i - 1) - 1, \\ 1 & \text{if } d \geqslant \sum_{i=1}^{n}(\#A_i - 1), \end{cases}$$

where $k, \ell \in \mathbb{Z}$ are such that $d = \sum_{i=k+1}^{n}(\#A_i - 1) + \ell$, $k \in [n]$ and $\ell \in [\#A_k - 1]$.

Proof. The minimum weight of $\mathrm{AGC}_d(A)$ is

$$\mathfrak{m}\left(\#A_1, \ldots, \#A_n; \sum_{i=1}^{n} \#A_i - d\right).$$

So the result follows from Lemma 2.1, as the greedy distribution of

$$\sum_{i=1}^{n} \#A_i - d = \sum_{i=1}^{k-1}(\#A_i - 1) + (\#A_k - 1 - \ell) + n$$

balls is $(\#A_1, \ldots, \#A_{k-1}, \#A_k - \ell, 1, \ldots, 1)$. □

Remark.

(a) The paper [27] makes no mention of Alon–Füredi. Their proof of Theorem 5.3 is self-contained and thus gives a proof of Alon–Füredi with the balls in bins constant replaced by its explicit value $P(y_G)$. On the other hand it is longer than the other proofs of Alon–Füredi appearing in the literature.

(b) Our proof of Theorem 5.3 works for a grid $A \subset R^n$ satisfying Condition (D).

(c) When $b_1 \geqslant \cdots \geqslant b_n$, the greedy algorithm computes $\mathfrak{m}(a_1, \ldots, a_n; b_1, \ldots, b_n; N)$, and we could give a similarly explicit description of $\mathrm{GAGC}_d(A_1; b_1, \ldots, b_n)$.

(d) While (binary) Reed–Muller codes are mentioned in [1] under Corollary 1, the connection between the Alon–Füredi theorem and generalized Reed–Muller codes is not explored.

6. Applications to finite geometry

6.1. Partial coverings of grids by hyperplanes

By a *hyperplane* in R^n we mean a polynomial $H = c_1 t_1 + \cdots + c_n t_n + r \in R[\underline{t}]$ for which at least one c_i is not a zero-divisor. (Referring to the polynomial itself rather than its zero locus in R^n will make the discussion cleaner.) A family $\mathcal{H} = \{H_i\}_{i=1}^{d}$ *covers* $x \in R^n$ if $H_i(x) = 0$ for some $i \in [n]$; $\mathcal{H}$ *covers* a subset $S \subset R^n$ if it covers every point of S, and $\mathcal{H}$ *partially covers* S otherwise. For a family $\mathcal{H} = \{H_i\}_{i=1}^{d}$ of hyperplane in R^n, put

$$f_{\mathcal{H}} = \prod_{i=1}^{d} H_i.$$

Thus $f_{\mathcal{H}}$ is a polynomial of degree d. If $\mathcal{H}$ covers A, then f vanishes identically on A. If R is a domain the converse holds, and thus $\mathcal{H}$ covers A if and only if $f_{\mathcal{H}} \in \langle \varphi_1, \ldots, \varphi_n \rangle$. We now revisit

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

the original combinatorial problem studied by Alon and Füredi, which is part (c) of the following theorem. However, our proof is via Theorem 1.1 instead of the approach used in [1].

Theorem 6.1. *Let R be a domain, let $A = \prod_{i=1}^n A_i \subset R^n$ be a finite grid, and let $\mathcal{H} = \{H_i\}_{i=1}^d$ be family of hyperplanes in R^n.*

(a) *If $\mathcal{H}$ partially covers A, then $\mathcal{H}$ fails to cover at least $\mathfrak{m}\left(\#A_1, \ldots, \#A_n; \sum_{i=1}^n \#A_i - d\right)$ points of A.*
(b) *For all $d \in \mathbb{Z}^+$, there is a family of hyperplanes $\{H_i = t_{j_i} - x_i\}_{i=1}^d$ with $j_i \in [n]$ and $x_i \in A_i$ which covers all but exactly $\mathfrak{m}\left(\#A_1, \ldots, \#A_n; \sum_{i=1}^n \#A_i - d\right)$ points of A.*
(c) *If $\mathcal{H}$ covers all but exactly one point of A, then $d \geqslant \sum_{i=1}^n (\#A_i - 1)$.*

Proof.

(a) As above, $\mathcal{H}$ covers $x \in R^n$ if and only if $f_{\mathcal{H}}(x) = 0$. Apply Alon–Füredi.
(b) The sharpness construction of Section 3.3 is precisely of this form.
(c) If $\mathcal{H}$ covers all points of A except one, then

$$1 = \#\mathcal{U}_A(H_1 \cdots H_d) \geqslant \mathfrak{m}\left(\#A_1, \ldots, \#A_n; \sum_{i=1}^n \#A_i - d\right),$$

so Lemma 2.2 gives $\sum_{i=1}^n \#A_i - d \leqslant n$, that is, $d \geqslant \sum_{i=1}^n (\#A_i - 1)$. □

We complement Theorem 6.1 by computing the minimum cardinality of a hyperplane covering of a finite grid (not necessarily specifying Condition (D)) over a ring R.

Theorem 6.2. *Let $A = \prod_{i=1}^n A_i \subset R^n$ be a finite grid, and let $\mathcal{H} = \{H_i\}_{i=1}^d$ be a hyperplane covering of A. Then $d \geqslant \min \#A_i$.*

Proof. First we observe that if A satisfies Condition (D) then the result is almost immediate: going by contraposition, if $d \leqslant \#A_i - 1$ for all $i \in [n]$ then $f_{\mathcal{H}}$ is non-zero and A-reduced, so it cannot vanish identically on A.

Now we give a non-polynomial method proof in the general case. Without loss of generality assume $\#A_1 \geqslant \cdots \geqslant \cdots \geqslant \#A_n$. We claim that any hyperplane $H = \sum_{i=1}^n c_i t_i + g$ covers at most $\prod_{i=1}^{n-1} \#A_i$ points of A: this suffices, for then $d \geqslant \#A_n$.

Proof of claim. Fix $i \in [n]$ such that c_i is not a zero-divisor in R. Let $\pi : R^n \to R^{n-1}$ be the projection $(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$. Then

$$A = \coprod_{x' = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in \pi(A)} \{x_1\} \times \cdots \times \{x_{i-1}\} \times A_i \times \{x_{i+1}\} \times \cdots \times \{x_n\}$$

is a partition of A into $\#\pi(A) = \prod_{j \neq i} \#A_j$ non-empty subsets, each one of which meets H in at most one point. So*********

$$\#(Z(H) \cap A) \leqslant \prod_{j \neq i} \#A_j \leqslant \prod_{i=1}^{n-1} \#A_i.$$ □

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

Conjecture 6.3. *Let R be a ring, and let $A_1,\dots,A_n \subset R$ be non-empty (but possibly infinite). Let $\mathcal{H} = \{H_j\}_{j\in J}$ be a covering of the grid $A = \prod_{i=1}^n A_i$ by hyperplanes. Then $\#J \geqslant \min_{i=1}^n \#A_i$.*

Remark.

(a) For $i \in [n]$, let $B_i \subset A_i \subset R$. Then we need at least as many hyperplanes to cover $\prod_{i=1}^n A_i$ as we do to cover $\prod_{i=1}^n B_i$. Together with Theorem 6.2 it follows that in the setting of Conjecture 6.3 we need at least $\min(\#A_i, \aleph_0)$ hyperplanes. Thus Conjecture 6.3 holds when R is countable.

(b) When R is a field and $A = R^n$, Conjecture 6.3 is a case of [11, Theorem 3].

6.2. Partial covers and blocking sets in finite geometries

The same ideas can be used to prove old and new results about Desarguesian projective and affine spaces over finite fields.

Let $\mathrm{PG}(n,q)$ denote the n-dimensional projective space over $\mathbb{F}_q$ (the same object would in some other circles be denoted by $\mathbb{P}^n(\mathbb{F}_q)$) and let $\mathrm{AG}(n,q)$ denote the n-dimensional affine space over $\mathbb{F}_q$ (resp. $\mathbb{A}^n(\mathbb{F}_q)$). The set $\mathrm{AG}(n,q)$ comes equipped with a sharply transitive action of the additive group of $\mathbb{F}_q^n$ and thus a choice of a point $x \in \mathrm{AG}(n,q)$ induces an isomorphism $\mathrm{AG}(n,q) \cong \mathbb{F}_q^n$. We will make such identifications without further comment.

A *partial cover* of $\mathrm{PG}(n,q)$ is a set of hyperplanes that do not cover all the points. The points missed by a partial cover are called *holes*.

Theorem 6.4. *Let $\mathcal{H}$ be a partial cover of $\mathrm{PG}(n,q)$ of size $k \in \mathbb{Z}^+$. Then $\mathcal{H}$ has at least $\mathfrak{m}(q,\dots,q;nq-k+1)$ holes.*

Proof. Let $H \in \mathcal{H}$. Then $\mathrm{PG}(n,q) \setminus H \cong \mathrm{AG}(n,q)$ so $\mathcal{H} \setminus H$ is a partial cover of $\mathbb{F}_q^n$ by $k-1$ hyperplanes. As above, there are at least $\mathfrak{m}(q,\dots,q;nq-(k-1))$ points not covered by $\mathcal{H}$. □

Corollary 6.5. *If $0 \leqslant a < q$, a partial cover of $\mathrm{PG}(n,q)$ of size $q+a$ has at least $q^{n-1} - aq^{n-2}$ holes.*

Proof. By Theorem 6.4 there are at least $\mathfrak{m}(q,\dots,q;(n-1)q-a+1)$ holes. Since $0 \leqslant a < q$, the greedy distribution is $(q,\dots,q,q-a,1)$, and the result follows. □

Dodunekov, Storme and Van de Voorde have shown that a partial cover of $\mathrm{PG}(n,q)$ of size $q+a$ has at least $q^{n-1} - aq^{n-2}$ holes if $0 \leqslant a < (q-2)/3$ [17, Theorem 17]. Corollary 6.5 gives an improvement in that the restriction on a is relaxed. They also show that if $a < (q-2)/3$ and the number of holes are at most q^{n-1}, then they are all contained in a single hyperplane. We cannot make any such conclusions from our arguments.

Projective duality yields a dual form of Theorem 6.4: k points in $\mathrm{PG}(n,q)$ which do not meet all hyperplanes must miss at least $\mathfrak{m}(q,\dots,q;nq-k+1)$ of them.

Theorem 6.6. *Let S be a set of k points in $\mathrm{AG}(n,q)$. Then there are at least $\mathfrak{m}(q,\dots,q;nq-k+1)-1$ hyperplanes of $\mathrm{AG}(n,q)$ which do not meet S.*

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

Proof. Add a hyperplane at infinity to get to the setting of $\mathrm{PG}(n,q)$ and then apply the dual form of Theorem 6.4. □

The general problem of the number of linear subspaces missed by a given set of points in $\mathrm{PG}(n,q)$ is studied by Metsch in [28]. We wish to note that Theorem 6.6 gives the same bounds as part (a) of [28, Theorem 1.2] for the specific case when the linear subspaces are hyperplanes.

A *blocking set* in $\mathrm{AG}(n,q)$ or $\mathrm{PG}(n,q)$ is a set of points that meets every hyperplane. The union of the coordinate axes in $\mathbb{F}_q^n$ yields a blocking set in $\mathrm{AG}(n,q)$ of size $n(q-1)+1$. Doyen conjectured in a 1976 Oberwolfach lecture that $n(q-1)+1$ is the least possible size of a blocking set in $\mathrm{AG}(n,q)$. A year later two independent proofs appeared, by Jamison [23], and then a (simpler) proof by Brouwer and Schrijver [8]. We are in a position to give another proof.

Corollary 6.7 (Jamison–Brouwer–Schrijver). *The minimum size of a blocking set in* $\mathrm{AG}(n,q)$ *is* $n(q-1)+1$.

Proof. Let $B \subset \mathrm{AG}(n,q)$ be a blocking set of cardinality at most $n(q-1)$. By Theorem 6.6 and Lemma 2.2 there are at least

$$\mathfrak{m}(q,\dots,q;nq-n(q-1)+1)-1 = \mathfrak{m}(q,\dots,q;n+1)-1 \geqslant 1$$

hyperplanes which do not meet B. □

Turning now to $\mathrm{PG}(n,q)$, every line is a blocking set. But classifying blocking sets that do not contain any line is one of the major open problems in finite geometry. For a survey on blocking sets in finite projective spaces, see [7, Chapter 3].

If $B \subset \mathrm{PG}(n,q)$, $x \in B$ and H is a hyperplane in $\mathrm{PG}(n,q)$, then H *is a tangent to* B *through* x if $H \cap B = \{x\}$. An *essential point* of a blocking set B in $\mathrm{PG}(n,q)$ is a point x such that $B \setminus \{x\}$ is not a blocking set. A point x of B is essential if and only if there is a tangent hyperplane to B through x.

Theorem 6.8. *Let* B *be a blocking set in* $\mathrm{PG}(n,q)$ *and let* x *be an essential point of* B. *There are at least* $\mathfrak{m}(q,\dots,q;nq-\#B+2)$ *tangent hyperplanes to* B *through* x.

Proof. Let H be a tangent hyperplane to B through x. Then $B' = B \setminus \{x\} \subset \mathrm{PG}(n,q) \setminus H \cong \mathrm{AG}(n,q)$. By Theorem 6.6 there are at least $\mathfrak{m}(q,\dots,q;nq-\#B+2)-1$ hyperplanes in $\mathrm{AG}(n,q)$ that do not meet B'. Since B is a blocking set, all of these hyperplanes, when seen in $\mathrm{PG}(n,q)$, must meet x. Thus there are at least $\mathfrak{m}(q,\dots,q;nq-\#B+2)$ tangent hyperplanes to B through x. □

Corollary 6.9 (Blokhuis–Brouwer [6]). *Let* B *be a blocking set in* $\mathrm{PG}(2,q)$ *of size* $2q-s$. *There are at least* $s+1$ *tangent lines through each essential point of* B.

Proof. By Theorem 6.8, each essential point of B has at least

$$\mathfrak{m}(q,q;2q-(2q-s)+2) = \mathfrak{m}(q,q;s+2)$$

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

tangent lines. Since $\#B = 2q - s < q^2 + q + 1 = \#\mathrm{PG}(2,q)$, there exists $x \in \mathrm{PG}(2,q) \setminus B$. There are $q+1$ lines through x, so $2q - s = \#B \geqslant q+1$. Thus $s+1 \leqslant q$, so the greedy distribution is $(s+1,1)$ and $\mathfrak{m}(q,q;s+2) = s+1$. □

Corollary 6.10 (Theorem 7 of [17]). *If $0 \leqslant a < q$, there are at least $q^{n-1} - aq^{n-2}$ tangent hyperplanes through each essential point of a blocking set of size $q+a+1$ in $\mathrm{PG}(n,q)$.*

Proof. By Theorem 6.8 and the proof of Corollary 6.5, each essential point of B has at least $\mathfrak{m}(q,\ldots,q;nq-(q+a+1)+2) = q^{n-1} - aq^{n-2}$ tangent hyperplanes. □

7. Multiplicity enhancements

That one can assign to a zero of a polynomial a positive integer called *multiplicity* is a familiar concept in the univariate case. The definition of the multiplicity $m(f,x)$ of a multivariate polynomial $f \in R[\underline{t}]$ at a point $x \in R^n$ (see Section 7.2) may be less familiar, but the concept is no less useful. All of the main results considered thus far are upper bounds on $\#Z_A(f)$, the number of zeros of a polynomial f on a finite grid. By a *multiplicity enhancement* we mean the replacement of $\#Z_A(f)$ by $\sum_{x\in A} m(f,x)$ in such an upper bound. The prototypical example: for a non-zero univariate polynomial f over a field F we have $\sum_{x\in F} m(f,x) \leqslant \deg f$.

Recently, multiplicity enhancements have become part of the polynomial method toolkit. In [18] Dvir, Kopparty, Saraf and Sudan gave a multiplicity enhancement of the Schwartz–Zippel lemma. This was a true breakthrough with important applications in both combinatorics and theoretical computer science. In Section 4 we saw that the original work of Schwartz, DeMillo–Lipton and Zippel consists of more than the Schwartz–Zippel lemma, and gave some extensions of this work, in particular working over an arbitrary ring. So it is natural to consider multiplicity enhancements of these results. We do so here, giving a multiplicity enhancement of Theorem 4.3 and thus also of Theorem 4.2. On the other hand the Alon–Füredi theorem does not allow for a multiplicity enhancement (at least not in the precise sense described above), as we will see in Example 7.11.

This is a situation where working over a ring under Condition (D) makes things a bit harder. Lemma 7.7 pushes through the single variable root-factor phenomenon under Condition (D).

In places our treatment closely follows that of [18]. We need to set things up over a ring, whereas they work over a field. Nevertheless, their work carries over verbatim much of the time, and when this is the case we state the result in the form we need it, cite the analogous result in [18] and omit the proof.

7.1. Hasse derivatives

Let $R[\underline{t}] = R[t_1,\ldots,t_n]$. For $I = (i_1,\ldots,i_n) \in \mathbb{N}^n$, put

$$\underline{t}^I = t_1^{i_1} \cdots t_n^{i_n}$$

and $|I| = \sum_{j=1}^n i_j = \deg \underline{t}^I$. Thus, $\{\underline{t}^I\}_{I\in\mathbb{N}^n}$ is an R-basis for $R[\underline{t}]$. We put

$$\binom{I}{J} = \prod_{k=1}^n \binom{i_k}{j_k},$$

taking $\binom{i}{j} = 0$ if $j > i$.

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

For $J \in \mathbb{N}^n$, let $D^J : R[\underline{t}] \to R[\underline{t}]$ be the unique R-linear map such that

$$D^J(\underline{t}^I) = \binom{I}{J}\underline{t}^{I-J}.$$

We have $D^J(\underline{t}^I) = 0$, unless $J \leqslant I$. Repeated application of the identity

$$t^n = (t-x+x)^n = \sum_{j=0}^{n}\binom{n}{j}x^{n-j}(t-x)^j$$

leads to the Taylor expansion: for $f \in R[\underline{t}]$ and $x \in R^n$,

$$f(\underline{t}) = \sum_J D^J(f)(x)(\underline{t}-x)^J. \tag{7.1}$$

Applying the automorphism $\underline{t} \mapsto \underline{t}+x$ gives the alternate form

$$f(\underline{t}+x) = \sum_J D^J(f)(x)\underline{t}^J.$$

These $D^J(f)$ were defined in [22] and are now called *Hasse derivatives*.

Proposition 7.1 ([18, Proposition 2.3]). *Let $f \in R[\underline{t}]$, and let $I, J \in \mathbb{N}^n$.*

(a) *If f is homogeneous of degree d and $D^I(f)$ is non-zero, then $D^I(f)$ is homogeneous of degree $d-|I|$.*

(b) *We have*

$$D^J(D^I(f)) = \binom{I+J}{I}D^{I+J}(f).$$

Lemma 7.2 (Leibniz rule). *Let $n = 1$, $i \in \mathbb{N}$ and $g, h \in R[t]$. Then we have*

$$D^i(gh) = \sum_{j=0}^{i} D^j(g)D^{i-j}(h). \tag{7.2}$$

Proof. *Step 1.* Recall Vandermonde's identity: for $m, n, r \in \mathbb{N}$, we have

$$\binom{m+n}{i} = \sum_{j=0}^{i}\binom{m}{j}\binom{n}{i-j}.$$

(If we have a set consisting of m red balls and n blue balls, then for $0 \leqslant j \leqslant i$, the number of i element subsets containing exactly j red balls is $\binom{m}{j}\binom{n}{i-j}$, so $\sum_{j=0}^{i}\binom{m}{j}\binom{n}{i-j}$ is the total number of i element subsets of an $m+n$ element set.)

Step 2. Using Vandermonde's identity, we get

$$D^i(t^m t^n) = \binom{m+n}{i}t^{m+n-i} = \sum_{j=0}^{i}\binom{m}{j}\binom{n}{i-j}t^{m+n-i} = \sum_{j=0}^{i} D^j(t^m)D^{i-j}(t^{m+n-i}).$$

By R-linearity of the Hasse derivatives, this establishes (7.2). □

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

7.2. Multiplicities

Let $f \in R[\underline{t}]$ be non-zero and $x \in R^n$. The *multiplicity of f at x*, denoted $m(f,x)$, is the natural number m such that $D^J(f)(x) = 0$ for all J with $|J| < m$ and $D^J(f)(x) \neq 0$ for some J with $|J| = m$. We put $m(0,x) = \infty$ for all $x \in R^n$.

Lemma 7.3 ([18, Lemma 2.4]). *For $f \in R[\underline{t}]$, $x \in R^n$ and $I \in \mathbb{N}^n$, we have*

$$m(D^I(f),x) \geqslant m(f,x) - |I|.$$

Given a vector $\underline{f} = (f_1,\ldots,f_k) \in R[\underline{t}]^k$, we put $m(\underline{f},x) = \min_{1\leqslant j\leqslant k} m(f_j,a)$.

Proposition 7.4 ([18, Proposition 2.5]). *Let $X_1,\ldots,X_n,Y_1,\ldots,Y_\ell$ be independent indeterminates. Let $\underline{f} = (f_1,\ldots,f_k) \in R[X_1,\ldots,X_n]^k$ and let $\underline{g} = (g_1,\ldots,g_n) \in R[Y_1,\ldots,Y_\ell]^n$. We define $\underline{f}\circ\underline{g} \in R[Y_1,\ldots,Y_\ell]^k$ to be $\underline{f}(g_1,\ldots,g_n)$.*

(a) *For any $a \in R^\ell$ we have*

$$m(\underline{f}\circ\underline{g},a) \geqslant m(\underline{f},\underline{g}(a))m(\underline{g}-\underline{g}(a),a).$$

(b) *We have*

$$m(\underline{f}\circ\underline{g},a) \geqslant m(\underline{f},\underline{g}(a)).$$

Corollary 7.5 ([18, Corollary 2.6]). *Let $f \in R[\underline{t}]$ and let $a,b \in R^n$. Then for all $c \in R$ we have*

$$m(f(a+tb),c) \geqslant m(f,a+cb).$$

Lemma 7.6. *Let $f,g,h \in R[t]$ be non-zero polynomials such that $f = gh$, and let $x \in R$. If $g(x)$ is not a zero divisor, then $m(f,x) = m(h,x)$.*

Proof. Let $m := m(f,x) \geqslant 1$ (for $m = 0$ the result is easily proved). By Lemma 7.2, for all $i \in \mathbb{N}$ we have

$$D^i(f) = D^0(g)D^i(h) + \cdots + D^i(g)D^0(h).$$

Thus $D^i(f)(x) = 0$ for all $i < m(h,x)$ and $m(f,x) \leqslant m(h,x)$.

We now show that $D^i(h)(x) = 0$ for all $i < m$, which will give $m(h,x) \geqslant m(f,x)$ and complete the proof. For $i = 0$, we have $0 = f(x) = g(x)h(x)$, and since $g(x)$ is not a zero divisor, we must have $h(x) = D^0(h)(x) = 0$. Now

$$D^1(f)(x) = D^0(g)(x)D^1(h)(x) + D^1(g)(x)D^0(h)(x) = g(x)D^1(h)(x).$$

Again, since $g(x)$ is not a zero divisor, $D^1(f)(x) = 0$ if and only if $D^1(h)(x) = 0$. Continuing in this way, at the ith step we have $D^j(h)(x) = 0$ for all $j < i$ and thus $D^i(f)(x) = g(x)D^i(h)(x)$. Since $g(x)$ is not a zero divisor and $i < m$, we see that $D^i(h)(x) = 0$. □

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

Lemma 7.7. *Let R be a ring, and let $f \in R[t]$ be a polynomial of degree $d \geqslant 1$. Let $A = \{x_1, \ldots, x_n\} \subseteq R$ be a finite set satisfying Condition (D). Then*

$$\sum_{x \in A} m(f,x) \leqslant d.$$

Proof. We will prove this by induction on n. For $n = 1$, we have that $(t - x_1)^{m(f,x_1)}$ divides f and hence $\deg f \geqslant m(f,x_1)$. So suppose $n \geqslant 2$. Write $f = (t - x_n)^{m(f,x_n)} g$. Since A satisfies Condition (D), the element $(x_i - x_n)^{m(f,x_i)}$ of R is not a zero divisor for any $i \in [n-1]$. Therefore, by Lemma 7.6 we have $m(f,x_i) = m(g,x_i)$ for all $i \in [n-1]$. By the induction hypothesis we get

$$\sum_{i=1}^{n-1} m(f,x_i) = \sum_{i=1}^{n-1} m(g,x_i) \leqslant \deg g = \deg f - m(f,x_n),$$

from which the result follows. □

Remark. An earlier draft of this work contained a different proof of Lemma 7.7. In place of Lemma 7.6, we used the following: if $f = \prod_{i=1}^{m}(t - x_i)$, $g = \prod_{j=1}^{m}(t - y_j) \in R[t]$ and $x_i - y_j \in R^\times$ for all i and j, then f and g generate the unit ideal of $R[t]$. This was proved using some commutative algebra (localization and Nakayama's lemma) and was thus a bit of a departure from the rest of the paper. The interested reader can find the details in the arXiv version of this paper [5].

Lemma 7.8 (DKSS lemma). *Let $A = \prod_{i=1}^{n} A_i \subset R^n$ be a finite subset satisfying Condition (D). Let $f \in R[t]$, and write*

$$f = \sum_{j=0}^{d_n} f_j(t_1, \ldots, t_{n-1}) t_n^j$$

with $f_{d_n} \neq 0$. Put $A' = \prod_{i=1}^{n-1} A_i$. For all $x' = (x_1, \ldots, x_{n-1}) \in A'$, we have

$$\sum_{x \in A_n} m(f,(x',x)) \leqslant (\#A_n) m(f_{d_n}, x') + d_n.$$

Proof. Choose $I' \in \mathbb{N}^{n-1}$ such that $|I'| = m(f_{d_n}, x')$ and $D^{I'}(f_{d_n})(x') \neq 0$. Put $I = I' \times \{0\} \in \mathbb{N}^n$. Then

$$D^I(f) = \sum_{j=0}^{d_n} D^{I'}(f_j) t_n^j,$$

so $D^I f \neq 0$. By Lemma 7.3, we have

$$m(f,(x',x)) \leqslant |I| + m(D^I(f),(x',x)) = m(f_{d_n},x') + m(D^I(f),(x',x)).$$

Apply Corollary 7.5 to $D^I(f)$ with $a = (x',0)$, $b = (0,1)$ and $c = x$: we get

$$m(D^I(f),(x',x)) \leqslant m(D^I(f)(x',t_n),x).$$

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

Summing over $x \in A_n$ gives

$$\sum_{x\in A_n} m(f,(x',x)) \leqslant (\#A_n)m(f_{d_n},x') + \sum_{x\in A_n} m(D^I(f)(x',t_n),x).$$

Since $I = I' \times \{0\}$, $D^I(f)(x',t_n)$ has degree d_n and thus Lemma 7.7 gives

$$\sum_{x\in A_n} m(D^I(f)(x',t_n),x) \leqslant d_n.$$

The result follows. □

Remark. The case of Lemma 7.8 in which R is a field and $A_1 = \cdots = A_n$ is due to Dvir, Kopparty, Saraf and Sudan [18, pp. 8–9]. Our proof follows theirs very closely, but uses Lemma 7.7 in place of the root-factor phenomenon.

7.3. Multiplicity enhanced Schwartz theorem

Theorem 7.9 (multiplicity enhanced Schwartz theorem). *Let R be a ring, let $A = \prod_{i=1}^n A_i \subset R^n$ be finite, non-empty and satisfy Condition (D), and let $f = f_n \in F[t_1,\dots,t_n]$ be a non-zero polynomial. Let $d_n = \deg_{t_n} f$, and let $f_{n-1} \in R[t_1,\dots,t_{n-1}]$ be the coefficient of $t_n^{d_n}$ in f_n. Let $d_{n-1} = \deg_{t_{n-1}} f_{n-1}$, and let $f_{n-2} \in R[t_1,\dots,t_{n-2}]$ be the coefficient of $t_{n-2}^{d_{n-2}}$ in f_{n-2}. Continuing in this manner we define for all $1 \leqslant i \leqslant n$ a polynomial $f_i \in R[t_i,\dots,t_n]$ with $\deg_{t_i} f_i = d_i$. Then*

$$\sum_{x\in A} m(f,x) \leqslant \#A \sum_{i=1}^n \frac{d_i}{\#A_i}.$$

Proof. We use induction on n. The case $n = 1$ is Lemma 7.7. Suppose the result holds for polynomials in $n-1$ variables. Let $A' = \prod_{i=1}^{n-1} A_i$. Applying Lemma 7.8 and then the induction hypothesis, we get

$$\begin{aligned}\sum_{x\in A} m(f,x) &= \sum_{x'\in A'} \sum_{x\in A_n} m(f,(x',x)) \leqslant \#A_n \sum_{x'\in A'} m(f_{n-1},x') + \#A' d_n \\ &\leqslant \#A_n \#A' \sum_{i=1}^{n-1} \frac{d_i}{\#A_i} + \#A \frac{d_n}{\#A_n} = \#A \sum_{i=1}^n \frac{d_i}{\#A_i}.\end{aligned}$$

□

Theorem 7.10 (multiplicity enhanced generalized Schwartz–Zippel). *Let $A = \prod_{i=1}^n A_i \subset R^n$ be a finite grid satisfying Condition (D), and suppose $\#A_1 \geqslant \cdots \geqslant \#A_n$. Let $f \in R[\underline{t}] = R[t_1,\dots,t_n]$ be a non-zero polynomial. Then*

$$\sum_{x\in A} m(f,x) \leqslant \deg f \prod_{i=1}^{n-1} \#A_i.$$

Proof. This follows from Theorem 7.9 as Theorem 4.2 does from Theorem 4.3. □

Remark.

(a) When R is a field, Theorem 7.9 was proved by Geil and Thomsen [20, Theorem 5]. They also build closely on [18].

https://doi.org/10.1017/S0963548317000566 Published online by Cambridge University Press

(b) Recent work of Geil and Thomsen [21, Prop. 17] shows that equality holds in Theorem 7.11 when R is a field and f is polylinear (see Section 3.4).

7.4. A counterexample

It is natural to ask whether Alon–Füredi holds in multiplicity enhanced form, that is, whether the bound

$$\#Z_A(f) \leqslant \#A - \mathfrak{m}\left(\#A_1, \ldots, \#A_n; \sum_{i=1}^{n} \#A_i - \deg f\right)$$

could be improved to

$$\sum_{x \in A} m(f,x) \leqslant \#A - \mathfrak{m}\left(\#A_1, \ldots, \#A_n; \sum_{i=1}^{n} \#A_i - \deg f\right).$$

The following example shows that such an improvement does not always hold.

Example 7.11. *Let $n = 2$ and $R = A_1 = A_2 = \mathbb{F}_q$. Let $d_1, d_2 \in \mathbb{Z}^+$ be such that $d_1, d_2 < q \leqslant d_1 + d_2 + 1$. Then $f = t_1^{d_1} t_2^{d_2}$ is A-reduced, and we have*

$$\sum_{x \in A} m(f,x) = qd_1 + qd_2 > q^2 - 2q + d_1 + d_2 + 1 = q^2 - \mathfrak{m}(q,q;2q - d_1 - d_2).$$

Notice that the polynomial $f = t_1^{d_1} t_2^{d_2}$ is polylinear (see Section 3.4). As far as we know it may be true that $\sum_{x \in A} m(f,x)$ is maximized among all polynomials of fixed degree when f is a polylinear polynomial. We leave this as an open problem.

Acknowledgements

We wish to thank Olav Geil and Fedor Petrov for helpful conversations and useful pointers to the literature.

References

[1] Alon, N. and Füredi, Z. (1993) Covering the cube by affine hyperplanes. *European J. Combin.* **14** 79–83.
[2] Alon, N. and Tarsi, M. (1992) Colorings and orientations of graphs. *Combinatorica* **12** 125–134.
[3] Ball, S. and Serra, O. (2009) Punctured combinatorial Nullstellensätze. *Combinatorica* **29** 511–522.
[4] Ball, S. and Serra, O. (2011) Erratum: Punctured combinatorial Nullstellensätze. *Combinatorica* **31** 377–378.
[5] Bishnoi, A., Clark, P. L., Potukuchi, A. and Schmitt, J. R. (2017) On zeros of a polynomial in a finite grid. arXiv:1508.06020v2
[6] Blokhuis, A. and Brouwer, A. E. (1986) Blocking sets in Desarguesian projective planes. *Bull. London Math. Soc.* **18** 132–134.
[7] Blokhuis, A., Sziklai, P. and Szőnyi, T. (2011) Blocking sets in projective spaces. In *Current Research Topics in Galois Geometry* (J. De Beule and L. Storme, eds), Nova Academic, pp. 61–84.
[8] Brouwer, A. E. and Schrijver, A. (1978) The blocking number of an affine space. *J. Combin. Theory Ser. A* **24** 251–253.

[9] Carvalho, C. (2013) On the second Hamming weight of some Reed–Muller type codes. *Finite Fields Appl.* **24** 88–94.
[10] Chevalley, C. (1935) Démonstration d'une hypothèse de M. Artin. *Abh. Math. Sem. Univ. Hamburg* **11** 73–75.
[11] Clark, P. L. (2012) Covering numbers in linear algebra. *Amer. Math. Monthly* **119** 65–67.
[12] Clark, P. L. (2014) The combinatorial Nullstellensätze revisited. *Electron. J. Combin.* **21** #P4.15.
[13] Clark, P. L. Fattening up Warning's second theorem. arXiv:1506.06743
[14] Clark, P. L., Forrow, A. and Schmitt, J. R. (2017) Warning's second theorem with restricted variables. *Combinatorica* **37** 397–417.
[15] Delsarte, P., Goethals, J.-M. and MacWilliams, F. J. (1970) On generalized Reed–Muller codes and their relatives. *Inform. Control* **16** 403–442.
[16] DeMillo, R. A. and Lipton, R. (1978) A probabilistic remark on algebraic program testing. *Inform. Process. Lett.* **7** 193–195.
[17] Dodunekov, S., Storme, L. and Van de Voorde, G. (2010) Partial covers of $\mathrm{PG}(n,q)$. *European J. Combin.* **31** 1611–1616.
[18] Dvir, Z., Kopparty, S., Saraf, S. and Sudan, M. (2013) Extensions to the method of multiplicities, with applications to Kakeya sets and mergers. *SIAM J. Comput.* **42** 2305–2328.
[19] Geil, O. (2008) On the second weight of generalized Reed–Muller codes. *Des. Codes Cryptogr.* **48** 323–330.
[20] Geil, O. and Thomsen, C. (2013) Weighted Reed–Muller codes revisited. *Des. Codes Cryptogr.* **66** 195–220.
[21] Geil, O. and Thomsen, C. (2017) More results on the number of zeros of multiplicity at least r, *Discrete Mathematics*, **79**, 384–410.
[22] Hasse, H. (1936) Theorie der höheren Differentiale in einem algebraischen Funktionenkörper mit vollkommenem Konstantenkörper bei beliebiger Charakteristik. *J. Reine Angew. Math.* **175** 50–54.
[23] Jamison, R. E. (1977) Covering finite fields with cosets of subspaces. *J. Combin. Theory Ser. A* **22** 253–266.
[24] Kasami, T., Lin, S. and Peterson, W. W. (1968) Generalized Reed–Muller codes. *Electron. Commun. Japan* **51** 96–104.
[25] van Lint, J. H. (1999) *Introduction to Coding Theory*, third edition, Vol. 86 of Graduate Texts in Mathematics, Springer.
[26] Lipton, R. The curious history of the Schwartz–Zippel lemma. https://rjlipton.wordpress.com/2009/11/30
[27] López, H. H., Renterá-Márquez, C. and Villarreal, R. H. (2014) Affine Cartesian codes. *Des. Codes Cryptogr.* **71** 5–19.
[28] Metsch, K. (2006) How many s-subspaces must miss a point set in $\mathrm{PG}(d,q)$. *J. Geom.* **86** 154–164.
[29] Muller, D. (1954) Application of Boolean algebra to switching circuit design and to error detection. *IRE Trans. Electronic Computers* **EC-3** (3) 6–12.
[30] Ore, Ö. (1922) Über höhere Kongruenzen. *Norsk Mat. Forenings Skrifter Ser. I* #7.
[31] Reed, I. S. (1954) A class of multiple-error-correcting codes and the decoding scheme. *IRE Trans. Information Theory* **PGIT-4** 38–49.
[32] Schauz, U. (2008) Algebraically solvable problems: Describing polynomials as equivalent to explicit solutions. *Electron. J. Combin.* **15** #R10.
[33] Schwartz, J. T. (1980) Fast probabilistic algorithms for verification of polynomial identities. *J. Assoc. Comput. Mach.* **27** 701–717.
[34] Warning, E. (1935) Bemerkung zur vorstehenden Arbeit von Herrn Chevalley. *Abh. Math. Sem. Hamburg* **11** 76–83.
[35] Zippel, R. (1979) Probabilistic algorithms for sparse polynomials. In *Proc. EUROSAM 79*, Vol. 72 of Lecture Notes in Computer Science, Springer, pp. 216–226.