

## SÁRKÖZY'S THEOREM IN FUNCTION FIELDS

BEN GREEN

ABSTRACT. Sárközy proved that dense sets of integers contain two elements differing by a  $k$ th power. The bounds in quantitative versions of this theorem are rather weak compared to what is expected. We prove a version of Sárközy's theorem for polynomials over  $\mathbb{F}_q$  with polynomial dependencies in the parameters.

More precisely, let  $P_{q,n}$  be the space of polynomials over  $\mathbb{F}_q$  of degree  $< n$  in an indeterminate  $T$ . Let  $k \geq 2$  be an integer and let  $q$  be a prime power. Set  $c(k, q) := (2k^2 D_q(k)^2 \log q)^{-1}$ , where  $D_q(k)$  is the sum of the digits of  $k$  in base  $q$ . If  $A \subset P_{q,n}$  is a set with  $|A| > 2q^{(1-c(k,q))n}$ , then  $A$  contains distinct polynomials  $p(T), p'(T)$  such that  $p(T) - p'(T) = b(T)^k$  for some  $b \in \mathbb{F}_q[T]$ .

## 1. INTRODUCTION

Let  $q$  be a prime power, and  $n$  an integer. Write  $P_{q,n}$  for the  $n$ -dimensional vector space over  $\mathbb{F}_q$  consisting of all polynomials  $c_0 + c_1 T + \cdots + c_{n-1} T^{n-1}$  (where  $T$  is an indeterminate) of degree  $< n$ .

**Theorem 1.1.** *Let  $k \geq 2$  be an integer and let  $q$  be a prime power. Set  $c(k, q) := (2k^2 D_q(k)^2 \log q)^{-1}$ , where  $D_q(k)$  is the sum of the digits of  $k$  in base  $q$ . If  $A \subset P_{q,n}$  is a set with  $|A| > 2q^{(1-c(k,q))n}$ , then  $A$  contains distinct polynomials  $p(T), p'(T)$  such that  $p(T) - p'(T) = b(T)^k$  for some  $b \in \mathbb{F}_q[T]$ .*

In certain cases, a stronger result than Theorem 1.1 is rather trivial. For example, if  $k = q^r$  then

$$(a_0 + a_1 T + \cdots + a_m T^m)^k = a_0 + a_1 T^k + \cdots + a_m T^{mk}.$$

Taking  $m = \lfloor \frac{n-1}{k} \rfloor$ , a bound of the form  $|A| \ll q^{(1-\frac{1}{k})n}$  follows quickly from the pigeonhole principle. In most cases, however, no trivial argument of this type is available.

---

The author is supported by a Simons Investigator Grant and by ERC Starting Grant 274938, *Approximate Algebraic Structure*. This work was entirely carried out while the author was visiting IMS, Singapore, and he thanks the institute for their hospitality.

Our method also proves the following result, which is qualitatively more general than Theorem 1.1 though has a slightly worse exponent.

**Theorem 1.2.** *Let  $F \in \mathbb{F}_q[T]$  be a polynomial of degree  $k$  with zero constant term. Set  $c'(k, q) := (2k^2(q-1)^2(1 + \log_q k)^2 \log q)^{-1}$ . If  $A \subset P_{q,n}$  is a set with  $|A| > 2q^{(1-c'(k,q))n}$ , then  $A$  contains distinct polynomials  $p(T), p'(T)$  such that  $p(T) - p'(T) = F(b(T))$  for some  $b \in \mathbb{F}_q[T]$ .*

These results are natural function field analogues of a well-known result of Sárközy [7], who showed that if  $F \in \mathbb{Z}[X]$  is a polynomial with  $F(0) = 0$ , and if  $A \subset \{1, \dots, N\}$  is a set with  $|A| \geq N(\log N)^{-c_F}$ , then  $A$  contains distinct elements  $a, a'$  such that  $a - a' = F(x)$  for some  $x \in \mathbb{Z}$ . However, the bounds we obtain in the function field setting are considerably stronger than the natural analogue of Sárközy's bound, and correspond to a bound of the form  $|A| \geq N^{1-c_F}$  for the integer version of the problem. No such bound is known, for any polynomial  $F$  of degree greater than 1.

For the best known bounds on the integer version of the problem, which are a little better than Sárközy's bounds, see [1]. For previous results on the function field problem see [4]. Regarding lower bounds, the only reference we know is an unpublished manuscript by C. Link [5], which adapts some constructions of Ruzsa [6] from the integers. In particular examples are given of sets  $A \subset P_{5,2n}$  with  $|A| = 5^{3n/2}$  (thus  $|A| = N^{3/4}$ , where  $N = 5^{2n} = |P_{5,2n}|$ ) such that  $A - A$  does not contain any squares other than 0. We thank J. Wolf for bringing this reference to our attention.

Theorems 1.1 and 1.2 are fairly straightforward consequences of the following statement, which is the “natural” result output by the methods of this paper.

**Theorem 1.3.** *Let  $d \in \mathbb{N}$ , and suppose that  $0 < m \leq n$ . Let  $\Phi = (\phi_1, \dots, \phi_n) : \mathbb{F}_q^m \rightarrow \mathbb{F}_q^n$  be a polynomial map, in which each  $\phi_i$  is a polynomial over  $\mathbb{F}_q$  in  $m$  variables with total degree  $\leq d$ . Suppose that  $|\Phi^{-1}(0)|$  is coprime to  $q$  (and so, in particular,  $0 \in \text{im}(\Phi)$ ). Let  $A \subset \mathbb{F}_q^n$  be a set such that  $A - A$  intersects  $\text{im}(\Phi)$  only at 0. Then  $|A| \leq 2q^n e^{-m^2/2nd^2}$ .*

The main tool in proving Theorem 1.3 will be a simple – but previously overlooked – lemma from the recent breakthrough work of Croot, Lev and Pach [2].

*Acknowledgement.* We are grateful to Will Sawin for comments on an earlier version of this note. His comments led to a streamlined version

of the argument, giving somewhat better dependencies in the exponents in the main theorems by avoiding the appeal to the Chevalley-Warning theorem.

## 2. MULTIVARIATE POLYNOMIALS

Every function  $f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q$  agrees pointwise with a polynomial

$$P_f(x) = \sum_{\alpha \in \{0,1,\dots,q-1\}^n} \tilde{f}(\alpha) x^\alpha, \quad (2.1)$$

where here  $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$  (and  $x = (x_1, \dots, x_n)$ ), and  $\tilde{f}(\alpha)$  takes values in  $\mathbb{F}_q$ . To prove this, note that the space  $V$  of functions  $f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q$  and the space  $W$  of polynomials of the above type are both vector spaces over  $\mathbb{F}_q$  of dimension  $q^n$ . Hence it suffices to show that the natural “evaluation” map  $V \rightarrow W$  is injective, and this may be done by induction on  $n$ , using the fact that if a polynomial in one variable of degree  $\leq q-1$  vanishes on  $\mathbb{F}_q$  then it is the zero polynomial.

Define the degree of a polynomial in  $W$  to be  $\max |\alpha|$ , where  $|\alpha| = \alpha_1 + \cdots + \alpha_n$ , and the maximum is taken over those  $\alpha$  for which  $\tilde{f}(\alpha) \neq 0$ . Note that the degree is always at most  $(q-1)n$ .

The notation  $\tilde{f}(\alpha)$  is meant to suggest that we think of this as a kind of “ $\mathbb{F}_q$ -valued Fourier transform” of  $f$ . We have the following formula for these “Fourier coefficients”  $\tilde{f}(\alpha)$  in terms of  $f$ .

**Lemma 2.1.** *We have*

$$\tilde{f}(\alpha) = \sum_{x \in \mathbb{F}_q^n} f(x) \prod_{i=1}^n \sigma_{\alpha_i}(x_i),$$

where

$$\sigma_a(x) = \begin{cases} -x^{q-1-a} & \text{if } a > 0 \\ 1_{x=0} & \text{if } a = 0. \end{cases}$$

*Proof.* We begin with the right hand side. Starting with (2.1), this may be expanded as

$$\sum_{\beta \in \{0,1,\dots,q-1\}^n} \tilde{f}(\beta) \prod_{i=1}^n S_{\alpha_i, \beta_i}$$

where

$$S_{a,b} = \sum_{x \in \mathbb{F}_q} x^b \sigma_a(x).$$

Thus it suffices to show that when  $a, b \in \{0, 1, \dots, q-1\}$  we have  $S_{a,b} = 1_{a=b}$ . We have  $S_{0,b} = 0^b$  and  $S_{q-1,0} = -\sum_{x \in \mathbb{F}_q} x^0 = 0$ , so

the claim holds in these cases. Also if  $a > 0$  then we have  $S_{a,a} = -\sum_{x \in \mathbb{F}_q} x^{q-1} = -\sum_{x \in \mathbb{F}_q^*} 1 = 1$ . In the remaining cases we have  $a > 0$ ,  $a \neq b$  and  $(a, b) \neq (q-1, 0)$ , so  $b-a$  is not a multiple of  $q-1$ . In all such cases, since  $a > 0$  we have

$$S_{a,b} = \sum_{x \in \mathbb{F}_q} x^{q-1+b-a}.$$

If  $\lambda \in \mathbb{F}_q^*$  then the multiplication-by- $\lambda$  map is a bijection on  $\mathbb{F}_q$ , and so  $S_{a,b} = \lambda^{q-1+b-a} S_{a,b}$ . But since  $q-1+b-a$  is a positive integer and not a multiple of  $q-1$  there is a choice of  $\lambda$  for which  $\lambda^{q-1+b-a} \neq 1$ . It follows that  $S_{a,b} = 0$ . This concludes the proof that  $S_{a,b} = 1_{a=b}$  in all cases, and hence the lemma.  $\square$

We note that when  $q = 2$  this formula takes a particularly simple form, namely

$$\tilde{f}(\alpha) = \sum_{x \in \mathbb{F}_2^n} f(x) \prod_{i: \alpha_i=0} 1_{x_i=0}.$$

### 3. $\text{im}(\Phi)$ AS THE SUPPORT OF A POLYNOMIAL OF LOW DEGREE

Using the results of the last section, we establish the following key lemma.

**Lemma 3.1.** *Suppose that  $\Phi : \mathbb{F}_q^m \rightarrow \mathbb{F}_q^n$  is a polynomial map of degree  $\leq d$ , that is to say  $\Phi = (\phi_1, \dots, \phi_n)$  with each  $\phi_i$  an  $m$ -variable polynomial over  $\mathbb{F}_q$  with total degree  $\leq d$ . Suppose that  $|\Phi^{-1}(0)|$  is coprime to  $q$ . Then there is a polynomial  $P \in \mathbb{F}_q[x_1, \dots, x_n]$ , with no variable having degree  $> (q-1)$ , such that*

- (1)  $\deg P \leq (q-1)(n - \frac{m}{d})$ ;
- (2)  $P$  is supported on  $\text{im}(\Phi)$  (that is,  $P(x) = 0$  when  $x \notin \text{im}(\Phi)$ );
- (3)  $P(0) \neq 0$ .

*Proof.* Let  $f(x) = |\Phi^{-1}(x)|$ . Then we will take  $P = P_f$  to be the polynomial representation of  $f$ , as in (2.1). Note that  $P_f(x) = f(x)$  when evaluated pointwise, and so properties (2) and (3) are immediate. It remains to verify property (1), to which end we must show that  $\tilde{f}(\alpha) = 0$  when  $|\alpha|$  satisfies the bound in (1). For this, we use the formula obtained in Lemma 2.1, which implies that

$$\tilde{f}(\alpha) = \sum_{a \in \mathbb{F}_q^m} \prod_{i=1}^n \sigma_{\alpha_i}(\phi_i(a)).$$

Now observe that  $\sigma_{\alpha_i}$  is a polynomial of degree  $q - 1 - \alpha_i$ , this being immediate from the definition when  $\alpha_i > 0$ , and true also when  $\alpha_i = 0$  since we have the identity  $1_{x=0} = 1 - x^{q-1}$ . It follows that  $\prod_{i=1}^n \sigma_{\alpha_i}(\phi_i(a))$  is a polynomial in  $a = (a_1, \dots, a_m)$  of degree at most  $\sum_{i=1}^n (q - 1 - \alpha_i)d = (n(q - 1) - |\alpha|)d$ . If this is  $< (q - 1)m$  then, when we sum over  $a \in \mathbb{F}_q^m$ , we get zero: indeed the lowest-degree monomial with nonzero sum is  $\prod_{i=1}^m a_i^{q-1}$ . Thus indeed  $\tilde{f}(\alpha) = 0$  so long as  $(n(q - 1) - |\alpha|)d < m(q - 1)$ , which is equivalent to  $|\alpha| > (q - 1)(n - \frac{m}{d})$ . This concludes the proof.  $\square$

#### 4. CROOT–LEV–PACH PRINCIPLE

In this section we complete the proof of Theorem 1.3 by applying the idea of Croot, Lev and Pach [2]. Suppose that  $d, m, \Phi$  are as in the statement of the theorem, and suppose that  $A \subset \mathbb{F}_q^n$  is a set for which  $A - A$  and  $\text{im}(\Phi)$  intersect only at 0. Let  $P$  be the polynomial constructed in Lemma 3.1. Then  $P(a - a') = 0$  when  $a \neq a'$  (since  $P$  is supported on  $\text{im}(\Phi)$ ) but  $P(a - a') \neq 0$  when  $a = a'$  (since  $P(0) \neq 0$ ). It follows that the rank of the  $|A| \times |A|$  matrix  $M$  with  $M_{a,a'} = P(a - a')$  is precisely  $|A|$ . On the other hand, it follows from the ideas in [2] that the rank of  $M$  is at most twice the number of indices  $\alpha \in \{0, 1, \dots, q - 1\}^n$  with  $|\alpha| \leq \frac{1}{2} \deg P$ .

The variant of the proof of this last fact given in [3] is sufficiently short that we can give it here: we have  $M = XYZ$  where

- $X$  is the  $|A| \times q^n$  matrix with  $(a, \alpha)$ -entry  $a^\alpha$  (where  $\alpha \in \{0, 1, \dots, q - 1\}^n$ );
- $Y$  is the  $q^n \times q^n$  matrix with  $(\alpha, \beta)$ -entry  $c_{\alpha, \beta}$ , where  $f(x + y) = \sum_{\alpha, \beta} c_{\alpha, \beta} x^\alpha y^\beta$ ;
- $Z$  is the  $q^n \times |A|$  matrix with  $(\beta, b)$ -entry  $b^\beta$ .

Thus  $\text{rk}(M) \leq \text{rk}(Y)$ . However, we may write

$$f(x + y) = \sum_m m(x) a_m(y) + m(y) b_m(x)$$

where  $m$  ranges over the set of monomials  $m(t) = t^\gamma$  with  $\gamma \in S$ , the set of all indices  $\gamma \in \{0, 1, \dots, q - 1\}^n$  with  $|\gamma| \leq \frac{1}{2} \deg f$ , and so  $Y$  is the sum of two matrices, one with row rank at most  $|S|$  and the other with column rank at most  $|S|$ .

Returning to our main argument, putting these observations together shows that  $|A|$  is at most twice the number of indices  $\alpha \in \{0, 1, \dots, q -$

$1\}^n$  with  $|\alpha| \leq \frac{1}{2}(q-1)(n - \frac{m}{d})$ , or in other words

$$|A| \leq 2q^n \mathbf{P}(X_1 + \cdots + X_n \leq \frac{1}{2}(q-1)(n - \frac{m}{d})),$$

where the  $X_i$  are i.i.d. random variables, each having the uniform distribution on  $\{0, 1, \dots, q-1\}$ . Equivalently,

$$|A| \leq 2q^n \mathbf{P}(Y_1 + \cdots + Y_n \leq -\frac{m}{d}), \quad (4.1)$$

where the  $Y_i$  are i.i.d. random variables, each uniformly distributed on the points  $-1 + j(\frac{2}{q-1})$ ,  $j = 0, 1, \dots, q-1$ , which all lie in  $[-1, 1]$ .

With some calculation, one could presumably obtain a fairly satisfactory asymptotic for the right-hand side of (4.1). However, this would not be particularly straightforward and in any case, in our opinion, the general behaviour of the exponents in Theorem 1.3 is of more interest than any numerical constants. To obtain this we can use the quite general and rather tidy large deviation inequality of Bernstein type due to Hoeffding, namely  $\mathbf{P}(Y_1 + \cdots + Y_n \leq -t) \leq e^{-t^2/2n}$ . This applies to any independent random variables  $Y_i$  taking values in  $[-1, 1]$  and with  $\mathbf{E}Y_i = 0$ . By (4.1), this gives the bound  $|A| \leq 2q^n e^{-m^2/2nd^2}$ , which is the same as the one claimed in Theorem 1.3.

## 5. SÁRKÖZY'S THEOREM IN FUNCTION FIELDS

The remaining task is to deduce Theorems 1.1 and 1.2 from Theorem 1.3.

*Proof of Theorem 1.1.* Identify  $P_{q,n}$  with  $\mathbb{F}_q^n$  via the map  $c_0 + c_1T + \cdots + c_{n-1}T^{n-1} \mapsto (c_0, \dots, c_{n-1})$ . Set  $m := \lfloor \frac{n-1}{k} \rfloor + 1$ . We have

$$\begin{aligned} & (a_0 + a_1T + \cdots + a_{m-1}T^{m-1})^k \\ &= g_0(a_0, \dots, a_m) + g_1(a_0, \dots, a_m)T + \cdots + g_{n-1}(a_0, \dots, a_m)T^{n-1}, \end{aligned}$$

where the  $g_i$  are polynomials with  $g_i(0) = 0$  and  $\deg g_i \leq D_q(k)$ , the sum of the digits of  $k$  in base  $q$ . To see this last point, note that if  $k = b_0 + b_1q + \cdots + b_rq^r$  in base  $q$  then

$$\begin{aligned} & (a_0 + a_1T + \cdots + a_{m-1}T^{m-1})^k \\ &= (a_0 + a_1T + \cdots + a_{m-1}T^{m-1})^{b_0} (a_0 + a_1T^q + \cdots + a_{m-1}T^{qm})^{b_1} \cdots \end{aligned}$$

Define  $\Phi : \mathbb{F}_q^m \rightarrow \mathbb{F}_q^n$  by taking  $\Phi = (\phi_1, \dots, \phi_n)$ , where

$$\phi_i(a_0, \dots, a_{m-1}) = g_{i-1}(a_0, \dots, a_{m-1}).$$

Then, under the identification of  $P_{q,n}$  with  $\mathbb{F}_q^n$ ,  $\text{im}(\Phi)$  is precisely the set of  $k$ th powers of polynomials. Furthermore  $\deg \Phi \leq D_q(k)$ , and

$\Phi^{-1}(0) = \{0\}$ . Thus we may apply Theorem 1.3 with  $m = \lfloor \frac{n-1}{k} \rfloor + 1 \geq \frac{n}{k}$  and  $d = D_q(k)$  and thereby obtain Theorem 1.1.

We remark that Will Sawin has pointed out to us that if  $q = p^r$  then one could, if desired, replace  $D_q(k)$  by  $D_p(k)$ . This may be done by considering the vector space of degree  $< n$  polynomials over  $\mathbb{F}_q$  as a vector space of degree  $nr$  over  $\mathbb{F}_p$ , noting that the  $k$ th power map is a polynomial map of degree  $D_p(k)$ .

*Proof of Theorem 1.2.* This is almost the same as the proof of Theorem 1.1, except that now we can only bound  $\deg \Phi$  by  $\sup_{k' \leq k} D_q(k') \leq (q-1)(1+\log_q k)$ . We must also be a little more careful in checking that  $|\Phi^{-1}(0)|$  is coprime to  $q$ . To see this, note that  $|\Phi^{-1}(0)|$  is the number of polynomials  $p \in \mathbb{F}_q[T]$  with  $F(p(T)) = 0$ . Since  $F$  has zero constant term,  $p(T) = 0$  is such a polynomial. If  $F(T) = a_1T + \cdots + a_mT^m$  with  $a_m \neq 0$  and  $p(T) = b_0 + b_1T + \cdots + b_rT^r$  with  $b_r \neq 0$ , then  $F(p(T)) = a_mb_rT^{rm} + \text{lower order terms} \neq 0$ . Thus in fact  $p(T) \equiv 0$  is the unique such polynomial, whence  $|\Phi^{-1}(0)| = 1$ .

## REFERENCES

- [1] A. Balog, J. Pelikan, J. Pintz and E. Szemerédi, *Difference sets without  $\kappa$ th powers*, Acta Math. Hungar. **65** (1994), no. 2, 165–187.
- [2] E. Croot, V. Lev and P. Pach, *Progression-free sets in  $\mathbf{Z}_4^n$  are exponentially small*, preprint, arXiv:1605.01506.
- [3] J. Ellenberg and D. Gijswijt, *On large subsets of  $\mathbb{F}_q^n$  with no 3-term arithmetic progression*, preprint, arXiv:1605.09223
- [4] Lê Thái Hoang and Y.-R.-Liu, *On sets of polynomials whose difference set contains no squares*, Acta Arith. **161** (2013), 127–143.
- [5] Cordelia Link, *On sets whose difference set contains no squares*, manuscript (2008).
- [6] I. Z. Ruzsa, *Difference sets without squares*, Periodica Math. Hungar., **15** (1984), 205–209.
- [7] A. Sárközy, *On difference sets of sequences of integers, III*, Acta Mathematica Academiae Scientiarum Hungaricae **3** (1978), pp. 355–386.

MATHEMATICAL INSTITUTE, RADCLIFFE OBSERVATORY QUARTER, WOODSTOCK RD, OXFORD OX2 6GG