ON SOME PROBLEMS OF GYARMATI AND SÁRKÖZY

Le Anh Vinh
Mathematics Department, Harvard University, Cambridge, Massachusetts
vinh@math.harvard.edu

Received: 12/3/08, Revised: 5/22/11, Accepted: 7/1/12, Published: 7/13/12

Abstract

In a recent paper, for "large" (but otherwise unspecified) subsets $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ of $\mathbb{F}_q$, Gyarmati and Sárközy (2008) showed the solvability of the equations $a + b = cd$, and $ab + 1 = cd$ with $a \in \mathcal{A}$, $b \in \mathcal{B}$, $c \in \mathcal{C}$, $d \in \mathcal{D}$. They asked whether one can extend these results to every $k \in \mathbb{N}$ in the following way: for large subsets $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ of $\mathbb{F}_q$, there are $a_1, \ldots, a_k, a'_1, \ldots, a'_k \in \mathcal{A}$, $b_1, \ldots, b_k, b'_1, \ldots, b'_k \in \mathcal{B}$ with $a_i + b_j, a'_i b'_j + 1 \in \mathcal{C}\mathcal{D}$ (for $1 \leqslant i, j \leqslant k$). In this paper, we give an affirmative answer to this question.

1. Introduction

In [6] and [5], Sárközy proved that if $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$ are "large" subsets of $\mathbb{Z}_p$, more precisely, $|\mathcal{A}||\mathcal{B}||\mathcal{C}||\mathcal{D}| \gg p^3$, then the equation

$$a + b = cd, \tag{1}$$

respectively

$$ab + 1 = cd, \tag{2}$$

can be solved with $a \in \mathcal{A}$, $b \in \mathcal{B}$, $c \in \mathcal{C}$ and $d \in \mathcal{D}$. Gyarmati and Sárközy [4] generalized the results on the solvability of equation (1) to finite fields. Using bounds of multiplicative character sums, Shparlinski [7] extended the class of sets which satisfy this property. Furthermore, Garaev [2, 3] considered the equations (1) and (2) over some special sets $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ to obtain new results on the sum-product problem in finite fields.

At the end of [4], Gyarmati and Sárközy proposed some open problems related to the above equations. They asked whether one can extend the solvability of the equations (1) and (2) in the following way: for every $k \in \mathbb{N}$, there are $c = c(k) > 0$ and $q_0 = q_0(k)$ such that if $q > q_0$, $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D} \subseteq \mathbb{F}_q$, $|\mathcal{A}||\mathcal{B}||\mathcal{C}||\mathcal{D}| > q^{4-c}$ then there are $a_1, \ldots, a_k, a'_1, \ldots, a'_k \in \mathcal{A}$, $b_1, \ldots, b_k, b'_1, \ldots, b'_k \in \mathcal{B}$ with $a_i + b_j$, $a'_i b'_j + 1 \in \mathcal{C}\mathcal{D}$

for $1 \leqslant i, j \leqslant k$. In this paper, we give an affirmative answer to this question. More precisely, our results are the following.

Theorem 1. *Let $k \in \mathbb{N}$. If $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D} \subseteq \mathbb{F}_q$ with $|\mathcal{A}||\mathcal{B}||\mathcal{C}||\mathcal{D}| \gg q^{4-\frac{1}{2(k+2)}}$, then there are $a_1, \ldots, a_k \in \mathcal{A}$, $b_1, \ldots, b_k \in \mathcal{B}$ with $a_i + b_j \in \mathcal{CD}$ for $1 \leqslant i, j \leqslant k$.*

Theorem 2. *Let $k \in \mathbb{N}$. If $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D} \subseteq \mathbb{F}_q$ with $|\mathcal{A}||\mathcal{B}||\mathcal{C}||\mathcal{D}| \gg q^{4-\frac{1}{2(k+2)}}$, then there are $a_1, \ldots, a_k \in \mathcal{A}$, $b_1, \ldots, b_k \in \mathcal{B}$ with $a_i b_j + 1 \in \mathcal{CD}$ for $1 \leqslant i, j \leqslant k$.*

In [4], Gyarmati and Sárközy also studied the solvability of other (higher degree) algebraic equations with solutions restricted to "large" subsets of $\mathbb{F}_q$. They considered the following equations:

$$a + b = f(c, d), \quad a \in \mathcal{A}, b \in \mathcal{B}, c \in \mathcal{C}, d \in \mathcal{D};$$

and

$$ab = f(c, d), \quad a \in \mathcal{A}, b \in \mathcal{B}, c \in \mathcal{C}, d \in \mathcal{D},$$

with $f(x, y) \in \mathbb{F}_q[x, y]$, $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D} \subset \mathbb{F}_q$. We generalize Theorems 1 and 2 in this direction. We have the following result for the sum problem.

Theorem 3. *Suppose that $f(x, y) \in \mathbb{F}_q[x, y]$, and $f(x, y)$ is not of the form $g(x) + h(y)$. We write $f(x, y)$ in the form*

$$f(x, y) = \sum_{i=0}^{m} g_i(x) y^i,$$

with $g_i(x) \in \mathbb{F}_q[x]$, and let I denote the greatest i value with the property that $g_i(x)$ is not identically constant. Assume that $(I, q) = 1$. For every $k \in \mathbb{N}$, if $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D} \subseteq \mathbb{F}_q$ with $|\mathcal{A}||\mathcal{B}||\mathcal{C}||\mathcal{D}| \gg q^{4-\frac{1}{4(k+2)}}$, then there are $a_1, \ldots, a_k \in \mathcal{A}$, $b_1, \ldots, b_k \in \mathcal{B}$ with $a_i + b_j \in f(\mathcal{C}, \mathcal{D})$ for $1 \leqslant i, j \leqslant k$ (where $f(\mathcal{C}, \mathcal{D}) = \{f(c, d) : c \in \mathcal{C}, d \in \mathcal{D}\}$).

Before formulating the next theorem, we need to take some definitions from [4].

Definition 4. A polynomial

$$F(x, y) = \sum_{i=1}^{n} G_i(y) x^i = \sum_{j=0}^{m} H_j(x) y^j \in \mathbb{F}_q[x, y]$$

is said to be *primitive in x* if $(G_0(y), \ldots, G_n(y)) = 1$, and it is said to be *primitive in y* if

$$(H_0(x), \ldots, H_m(x)) = 1.$$

Definition 5. Every polynomial $f(x, y) \in \mathbb{F}_q[x, y]$ can be written uniquely (apart from constant factors) in the form

$$f(x, y) = F(x) G(x) H(x, y)$$

where $H(x, y)$ is primitive in both x and y. The polynomial $H(x, y)$ (uniquely determined up to constant factors) is called the *primitive kernel* of $f(x, y)$.

We now can state an analog of Theorem 3 for the product problem.

Theorem 6. *Suppose that $f(x,y) \in \mathbb{F}_q[x,y]$ and the primitive kernel $H(x,y)$ of $f(x,y)$ is not of the form $c(K(x,y))^d$. For every $k \in \mathbb{N}$, if $\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D} \subset \mathbb{F}_q$ with $|\mathcal{A}||\mathcal{B}||\mathcal{C}||\mathcal{D}| \gg q^{4-\frac{1}{4(k+2)}}$, then there are $a_1,\ldots,a_k \in \mathcal{A}$, $b_1,\ldots,b_k \in \mathcal{B}$ with $a_ib_j \in f(\mathcal{C},\mathcal{D})$ for $1 \leqslant i,j \leqslant k$.*

2. Pseudo-Randomness of Restricted-Sum Graphs

For any $a \in \mathcal{A}$, $c \in \mathcal{C}$, denote by $N^{c,\mathcal{D}}(a)$ the set of all $b \in \mathbb{F}_q$ such that $a+b \in c\mathcal{D}$, and let $N^{c,\mathcal{D}}_{\mathcal{B}}(a) = N^{c,\mathcal{D}}(a) \cap \mathcal{B}$. The following key estimate says that the cardinalities of the $N^{c,\mathcal{D}}_{\mathcal{B}}(a)$'s are close to $\frac{|\mathcal{B}||\mathcal{D}|}{q}$ when $|\mathcal{B}|, |\mathcal{D}|$ are large.

Lemma 7. *For all subsets $\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D}$ of $\mathbb{F}_q$, we have*

$$\sum_{(a,c)\in\mathbb{F}_q^2} \left(\left| N^{c,\mathcal{D}}_{\mathcal{B}}(a) \right| - \frac{|\mathcal{B}||\mathcal{D}|}{q} \right)^2 < q|\mathcal{B}||\mathcal{D}|.$$

Proof. For any set X, let $X(\cdot)$ denote the characteristic function of X. Let χ be any non-trivial additive character of $\mathbb{F}_q$. We have

$$\begin{aligned} |N^{c,\mathcal{D}}_{\mathcal{B}}(a)| &= \sum_{(b,d)\in\mathbb{F}_q^2, a+b-cd=0} \mathcal{B}(b)\mathcal{D}(d) \\ &= \sum_{(b,d)\in\mathbb{F}_q^2, s\in\mathbb{F}_q} \frac{1}{q}\chi(s(a+b-cd))\mathcal{B}(b)\mathcal{D}(d) \\ &= \frac{|\mathcal{B}||\mathcal{D}|}{q} + \frac{1}{q} \sum_{(b,d)\in\mathbb{F}_q^2, s\in\mathbb{F}_q^*} \chi(s(a+b-cd))\mathcal{B}(b)\mathcal{D}(d). \end{aligned}$$

Therefore

$$\begin{aligned} &\sum_{(a,c)\in\mathbb{F}_q^2} \left(\left| N^{c,\mathcal{D}}_{\mathcal{B}}(a) \right| - \frac{|\mathcal{B}||\mathcal{D}|}{q} \right)^2 \\ &= \frac{1}{q^2} \sum_{(a,c)\in\mathbb{F}_q^2} \left(\sum_{(b,d)\in\mathbb{F}_q^2, s\in\mathbb{F}_q^*} \chi(s(a+b-cd))\mathcal{B}(b)\mathcal{D}(d) \right)^2 \\ &= \frac{1}{q^2} \sum_{\substack{a,c,b,b',d,d'\in\mathbb{F}_q \\ s,s'\in\mathbb{F}_q^*}} \chi((s-s')a)\chi(sb-s'b')\chi(c(s'd'-sd))\mathcal{B}(b)\mathcal{D}(d)\mathcal{B}(b')\mathcal{D}(d') \\ &= \sum_{b,d,b'\in\mathbb{F}_q, s=s'\in\mathbb{F}_q^*} \chi(s(b-b'))\mathcal{B}(b)\mathcal{D}(d)\mathcal{B}(b') \\ &= R_1 + R_2\,, \end{aligned} \tag{3}$$

where R_1 is taken over $b = b'$ and R_2 is taken over $b \neq b'$ (the third line follows from the orthogonality in a and c. Consider the second line as a sum over a, then c implies that all summands vanish unless $s = s'$ and $d = d'$). We have

$$\begin{aligned} R_1 &= \sum_{b=b', d\in\mathbb{F}_q, s=s'\in\mathbb{F}_q^*} \chi(s(b-b'))\mathcal{B}(b)\mathcal{D}(d)\mathcal{B}(b') \\ &= (q-1)\sum_{b,d\in\mathbb{F}_q} \mathcal{B}(b)\mathcal{D}(d) = (q-1)|\mathcal{B}||\mathcal{D}|, \end{aligned} \tag{4}$$

and

$$\begin{aligned} R_2 &= \sum_{b\neq b', d\in\mathbb{F}_q, s=s'\in\mathbb{F}_q^*} \chi(s(b-b'))\mathcal{B}(b)\mathcal{D}(d)\mathcal{B}(b') \\ &= \sum_{b,d\in\mathbb{F}_q, s\in\mathbb{F}_q^*, t\neq 1\in\mathbb{F}_q, b'=tb} \chi(sb(1-t))\mathcal{B}(b)\mathcal{D}(d)\mathcal{B}(tb) \\ &= -\sum_{b,d\in\mathbb{F}_q, t\neq 1} \mathcal{B}(b)\mathcal{D}(d)\mathcal{B}(tb) \\ &< 0. \end{aligned} \tag{5}$$

The lemma follows immediately from (3), (4) and (5). □

The following result is an easy corollary of Lemma 7.

Corollary 8. *For all subsets $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ of $\mathbb{F}_q$ and $c \in \mathcal{C}$, let $N^{c,\mathcal{D}}(\mathcal{A}, \mathcal{B})$ be the number of pairs $(a, b) \in \mathcal{A} \times \mathcal{B}$ such that $a + b \in c\mathcal{D}$. Then there exists $c_0 \in \mathcal{C}$ such that*

$$\left| N^{c_0,\mathcal{D}}(\mathcal{A},\mathcal{B}) - \frac{|\mathcal{D}|}{q}|\mathcal{A}||\mathcal{B}| \right| < \sqrt{\frac{q|\mathcal{D}|}{|\mathcal{C}|}}\sqrt{|\mathcal{A}||\mathcal{B}|}.$$

Proof. By the pigeon-hole principle, there exists $c_0 \in \mathcal{C}$ such that

$$\sum_{a\in\mathcal{A}} \left(\left| N_{\mathcal{B}}^{c_0,\mathcal{D}}(a) \right| - \frac{|\mathcal{B}||\mathcal{D}|}{q} \right)^2 \leqslant \frac{1}{|\mathcal{C}|} \sum_{a\in\mathcal{A}, c\in\mathcal{C}} \left(\left| N_{\mathcal{B}}^{c,\mathcal{D}}(a) \right| - \frac{|\mathcal{B}||\mathcal{D}|}{q} \right)^2 < \frac{q|\mathcal{D}||\mathcal{B}|}{|\mathcal{C}|}.$$

By the Cauchy-Schwartz inequality,

$$\begin{aligned} \left| N^{c_0,\mathcal{D}}(\mathcal{A},\mathcal{B}) - \frac{|\mathcal{D}|}{q}|\mathcal{A}||\mathcal{B}| \right| &\leqslant \sum_{a\in\mathcal{A}} \left| \left| N_{\mathcal{B}}^{c_0,\mathcal{D}}(a) \right| - \frac{|\mathcal{B}||\mathcal{D}|}{q} \right| \\ &\leqslant \sqrt{|\mathcal{A}|}\sqrt{\sum_{a\in\mathcal{A}} \left(\left| N_{\mathcal{B}}^{c_0,\mathcal{D}}(a) \right| - \frac{|\mathcal{B}||\mathcal{D}|}{q} \right)^2} \\ &\leqslant \sqrt{\frac{q|\mathcal{D}}{|\mathcal{C}|}}\sqrt{|\mathcal{A}||\mathcal{B}|}. \end{aligned}$$

□

As a consequence, for any two large subsets $\mathcal{A}, \mathcal{B}$ of $\mathbb{F}_q$, there are many pairs $(a, b) \in \mathcal{A} \times \mathcal{B}$ with $a + b \in \mathcal{CD}$.

Corollary 9. *For all subsets $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ of $\mathbb{F}_q$, let $N^{\mathcal{C},\mathcal{D}}(\mathcal{A}, \mathcal{B})$ be the set of pairs $(a, b) \in \mathcal{A} \times \mathcal{B}$ such that $a + b \in \mathcal{CD}$. Then*

$$N^{\mathcal{C},\mathcal{D}}(\mathcal{A}, \mathcal{B}) \geqslant \frac{|\mathcal{D}|}{q}|\mathcal{A}||\mathcal{B}| - \sqrt{\frac{q|\mathcal{D}|}{|\mathcal{C}|}}\sqrt{|\mathcal{A}||\mathcal{B}|}.$$

Proof. It follows immediately from Corollary 8. □

Note that Corollaries 8 and 9 can be derived directly from Theorem 1 in [4]. However, Theorem 1 in [4] is also an easy corollary of Lemma 7 above.

Theorem 10. (cf. Theorem 1 in [4]) *For any subsets $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D} \subseteq \mathbb{F}_q$, denote by $N(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ the number of solutions of Eq. (1). Then we have*

$$\left| N(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}) - \frac{|\mathcal{A}||\mathcal{B}||\mathcal{C}||\mathcal{D}|}{q} \right| < \sqrt{q|\mathcal{A}||\mathcal{B}||\mathcal{C}||\mathcal{D}|}.$$

Proof. By Lemma 7, we have

$$\sum_{a \in \mathcal{A}, c \in \mathcal{C}} \left(\left| N_{\mathcal{B}}^{c,\mathcal{D}}(a) \right| - \frac{|\mathcal{B}||\mathcal{D}|}{q} \right)^2 \leq \sum_{(a,c) \in \mathbb{F}_q^2} \left(\left| N_{\mathcal{B}}^{c,\mathcal{D}}(a) \right| - \frac{|\mathcal{B}||\mathcal{D}|}{q} \right)^2 < q|\mathcal{B}||\mathcal{D}|.$$

By the Cauchy-Schwartz inequality,

$$\begin{aligned} \left| N(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}) - \frac{|\mathcal{A}||\mathcal{B}||\mathcal{C}||\mathcal{D}|}{q} \right| &\leqslant \sum_{(a,c) \in \mathbb{F}_q^2} \left| \left| N_{\mathcal{B}}^{c,\mathcal{D}}(a) \right| - \frac{|\mathcal{B}||\mathcal{D}|}{q} \right| \\ &\leqslant \sqrt{|\mathcal{A}||\mathcal{C}|} \sqrt{\sum_{a \in \mathcal{A}, c \in \mathcal{C}} \left(\left| N_{\mathcal{B}}^{c,\mathcal{D}}(a) \right| - \frac{|\mathcal{B}||\mathcal{D}|}{q} \right)^2} \\ &\leqslant \sqrt{q|\mathcal{A}||\mathcal{B}||\mathcal{C}||\mathcal{D}|}. \end{aligned}$$

□

3. Pseudo-Randomness of Restricted-Product Graphs

For any $a \in \mathcal{A}$, $c \in \mathcal{C}$, let $T^{c,\mathcal{D}}(a)$ be the set of all $b \in \mathbb{F}_q$ such that $ab + 1 \in c\mathcal{D}$, and let $T_{\mathcal{B}}^{c,\mathcal{D}}(a) = T^{c,\mathcal{D}}(a) \cap \mathcal{B}$. The following key estimate says that the cardinalities of the $T_{\mathcal{B}}^{c}(a)$'s are close to $\frac{|\mathcal{B}||\mathcal{D}|}{q}$ when $|\mathcal{B}|, |\mathcal{D}|$ are large.

Lemma 11. *For all subsets $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ of $\mathbb{F}_q$, we have*

$$\sum_{(a,c)\in\mathbb{F}_q^2}\left(\left|T_{\mathcal{B}}^{c,\mathcal{D}}(a)\right| - \frac{|\mathcal{B}||\mathcal{D}|}{q}\right)^2 < q|\mathcal{B}||\mathcal{D}|.$$

Proof. For any set X, let $X(\cdot)$ denote the characteristic function of X. Let χ be any non-trivial additive character of $\mathbb{F}_q$. We have

$$\begin{aligned}
|T_{\mathcal{B}}^{c,\mathcal{D}}(a)| &= \sum_{(b,d)\in\mathbb{F}_q^2, ab-cd+1=0} \mathcal{B}(b)\mathcal{D}(d) \\
&= \sum_{(b,d)\in\mathbb{F}_q^2, s\in\mathbb{F}_q} \frac{1}{q}\chi(s(ab-cd+1))\mathcal{B}(b)\mathcal{D}(d) \\
&= \frac{|\mathcal{B}||\mathcal{D}|}{q} + \frac{1}{q}\sum_{(b,d)\in\mathbb{F}_q^2, s\in\mathbb{F}_q^*} \chi(s(ab-cd+1))\mathcal{B}(b)\mathcal{D}(d).
\end{aligned}$$

Therefore

$$\begin{aligned}
&\sum_{(a,c)\in\mathbb{F}_q^2}\left(\left|T_{\mathcal{B}}^{c,\mathcal{D}}(a)\right| - \frac{|\mathcal{B}||\mathcal{D}|}{q}\right)^2 \\
&= \frac{1}{q^2}\sum_{(a,c)\in\mathbb{F}_q^2}\left(\sum_{(b,d)\in\mathbb{F}_q^2, s\in\mathbb{F}_q^*}\chi(s(ab-cd+1))\mathcal{B}(b)\mathcal{D}(d)\right)^2 \\
&= \frac{1}{q^2}\sum_{\substack{a,c,b,b',d,d'\in\mathbb{F}_q \\ s,s'\in\mathbb{F}_q^*}} \chi(a(sb-s'b'))\chi(c(s'd'-sd))\chi(s-s')\mathcal{B}(b)\mathcal{D}(d)\mathcal{B}(b')\mathcal{D}(d') \\
&= \frac{1}{q^2}(R_1+R_2), \qquad (6)
\end{aligned}$$

where R_1 is taken over $s = s'$ and R_2 is taken over $s \neq s'$. We have

$$\begin{aligned}
R_1 &= \sum_{a,c,b,b',d,d'\in\mathbb{F}_q, s=s'\in\mathbb{F}_q^*} \chi(as(b-b'))\chi(cs(d-d'))\mathcal{B}(b)\mathcal{D}(d)\mathcal{B}(b')\mathcal{D}(d') \\
&= (q-1)q^2|\mathcal{B}||\mathcal{D}|, \qquad (7)
\end{aligned}$$

where the last line follows from the orthogonality in a and then c. Considering the sum over a and then over b, this implies that all summands with $b \neq b'$ or $d \neq d'$ vanish. Now we compute R_2.

$$\begin{aligned}
R_2 &= \sum_{\substack{a,c,b,b',d,d'\in\mathbb{F}_q \\ s\in\mathbb{F}_q^*, t\neq 1\in\mathbb{F}_q}} \chi(as(b-tb))\chi(cs(d-td))\chi(s(1-t))\mathcal{B}(b)\mathcal{D}(d)\mathcal{B}(b')\mathcal{D}(d') \\
&= -\sum_{a,c,b'=tb,d'=td\in\mathbb{F}_q, s\in\mathbb{F}_q^*, t\neq 1} \mathcal{B}(b)\mathcal{D}(d)\mathcal{B}(b')\mathcal{D}(d') \\
&< 0, \qquad (8)
\end{aligned}$$

where the last line follows from the orthogonality in a and c. By considering the sum over a, and then over b, this implies that all summands with $b' \neq tb$ or $d' \neq td$ vanish. The lemma follows immediately from (6), (7) and (8).

□

Corollary 12. *For all subsets $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ of $\mathbb{F}_q$ and $c \in \mathcal{C}$, let $T^{c,\mathcal{D}}(\mathcal{A}, \mathcal{B})$ be the set of pairs $(a, b) \in \mathcal{A} \times \mathcal{B}$ such that $ab + 1 \in c\mathcal{D}$. Then there exists $c_0 \in \mathcal{C}$ such that*

$$\left| T^{c_0,\mathcal{D}}(\mathcal{A}, \mathcal{B}) - \frac{|\mathcal{D}|}{q} |\mathcal{A}||\mathcal{B}| \right| < \sqrt{\frac{q|\mathcal{D}|}{|\mathcal{C}|}} \sqrt{|\mathcal{A}||\mathcal{B}|}.$$

Proof. The proof of this corollary is similar to that of Corollary 8, except that we use Lemma 11 instead of Lemma 7. □

We also have an analog of Corollary 9 in the shifted-product problem.

Corollary 13. *For all subsets $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ of $\mathbb{F}_q$, let $N^{\mathcal{C},\mathcal{D}}(\mathcal{A}, \mathcal{B})$ be the set of pairs $(a, b) \in \mathcal{A} \times \mathcal{B}$ such that $ab + 1 \in \mathcal{C}\mathcal{D}$. Then*

$$T^{\mathcal{C},\mathcal{D}}(\mathcal{A}, \mathcal{B}) \geqslant \frac{|\mathcal{D}|}{q} |\mathcal{A}||\mathcal{B}| - \sqrt{\frac{q|\mathcal{D}|}{|\mathcal{C}|}} \sqrt{|\mathcal{A}||\mathcal{B}|}.$$

Similarly as in the previous section, slightly weaker (but still useful) versions of Corollaries 12 and 13 can be derived directly from Theorem 2 in [4].

4. Proof of Theorems 1

We now give a proof of Theorem 1.1. The key tool is the following lemma.

Lemma 14. *Suppose that $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$ of $\mathbb{F}_q$ with*

$$|\mathcal{A}|, |B| \gg \sqrt{\frac{q|\mathcal{D}|}{|\mathcal{C}|}} \left(\frac{q}{|\mathcal{D}|} \right)^k .$$

Then there are $a_1, \ldots, a_k \in \mathcal{A}$, $b_1, \ldots, b_k \in \mathcal{B}$ such that $a_i + b_j \in \mathcal{C}\mathcal{D}$ for all $1 \leqslant i, j \leqslant k$.

Proof. The proof proceeds by induction on k. The base case, $k = 1$, follows immediately from Corollary 9. Suppose that the theorem holds for all $l < k$. From Corollary 9, we have

$$N^{\mathcal{C},\mathcal{D}}(\mathcal{A}, \mathcal{B}) \geqslant \frac{|\mathcal{D}|}{q} |\mathcal{A}||\mathcal{B}| - \sqrt{\frac{q|\mathcal{D}|}{|\mathcal{C}|}} \sqrt{|\mathcal{A}||\mathcal{B}|} = (1 + o(1)) \frac{|\mathcal{D}|}{q} |\mathcal{A}||\mathcal{B}|.$$

By the pigeon-hole principle, there exists $a_1 \in \mathcal{A}$ such that

$$N^{\mathcal{C},\mathcal{D}}(a_1,\mathcal{B}) \geqslant (1+o(1))\frac{|\mathcal{D}|}{q}|\mathcal{B}| \gg \sqrt{\frac{q|\mathcal{D}|}{|\mathcal{C}|}}\left(\frac{q}{|\mathcal{D}|}\right)^{k-1}. \tag{9}$$

Let $\mathcal{B}_1$ be the set of all $b \in \mathcal{B}$ such that $a_1 + b \in \mathcal{CD}$. From Corollary 9, we have

$$N^{\mathcal{C},\mathcal{D}}(\mathcal{A},\mathcal{B}_1) \geqslant \frac{|\mathcal{D}|}{q}|\mathcal{A}||\mathcal{B}_1| - \sqrt{\frac{q|\mathcal{D}|}{|\mathcal{C}|}}\sqrt{|\mathcal{A}||\mathcal{B}_1|} = (1+o(1))\frac{|\mathcal{D}|}{q}|\mathcal{A}||\mathcal{B}_1|.$$

By the pigeon-hole principle, there exists $b_1 \in \mathcal{B}_1$ such that

$$N^{\mathcal{C},\mathcal{D}}(\mathcal{A},b_1) \geqslant (1+o(1))\frac{|\mathcal{D}|}{q}|\mathcal{A}| \gg \sqrt{\frac{q|\mathcal{D}|}{|\mathcal{C}|}}\left(\frac{q}{|\mathcal{D}|}\right)^{k-1}. \tag{10}$$

Let $\mathcal{A}_1$ be the set of all $a \in \mathcal{A}$ such that $a + b_1 \in \mathcal{CD}$. Set $\mathcal{A}^* = \mathcal{A}\backslash\{a_1\}$ and $\mathcal{B}^* = \mathcal{B}_1\backslash\{b_1\}$. It follows from (9) and (10) that

$$|\mathcal{A}^*|, |\mathcal{B}^*| \gg \sqrt{\frac{q|\mathcal{D}|}{|\mathcal{C}|}}\left(\frac{q}{|\mathcal{D}|}\right)^{k-1}.$$

Thus, by the induction hypothesis, there are $a_2,\ldots,a_k \in \mathcal{A}^*$, $b_2,\ldots,b_k \in \mathcal{B}^*$ such that $a_i + b_j \in \mathcal{CD}$ for all $2 \leqslant i,j \leqslant k$. We also have $a_1 + b_i, a_j + b_1 \in \mathcal{CD}$ for all $i,j = 1,\ldots,k$. This completes the proof of the lemma. □

Let $c = c(k) = \frac{1}{2(k+2)}$ and $q \gg 1$. Then $|\mathcal{A}|,|\mathcal{B}|,|\mathcal{C}|,|\mathcal{D}| \gg q^{1-c}$. It follows that

$$\sqrt{\frac{q|\mathcal{D}|}{|\mathcal{C}|}}\left(\frac{q}{|\mathcal{D}|}\right)^{k} \ll q^{(1+c)/2+ck} \ll q^{1-c} \ll |\mathcal{A}|,|\mathcal{B}|. \tag{11}$$

Therefore, Theorem 1 follows immediately from Lemma 14. Note that the upper bound for the left hand side of (11) can be estimated by $q^{1/2+kc}$. This can improve the bound of Theorem 1 to $|\mathcal{A}||\mathcal{B}||\mathcal{C}||\mathcal{D}| \gg q^{4-\frac{1}{2k+2}}$.

5. Proof of Theorem 2

Similar to the previous section, we can obtain the following result from Corollary 13.

Lemma 15. *Suppose that $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$ of $\mathbb{F}_q$ with*

$$|\mathcal{A}|, |B| \gg \sqrt{\frac{q|\mathcal{D}|}{|\mathcal{C}|}}\left(\frac{q}{|\mathcal{D}|}\right)^{k}.$$

Then there are $a_1,\ldots,a_k \in \mathcal{A}$, $b_1,\ldots,b_k \in \mathcal{B}$ such that $a_ib_j + 1 \in \mathcal{CD}$ for all $1 \leqslant i,j \leqslant k$.

Let $c = c(k) = \frac{1}{2(k+2)}$ and $q \gg 1$, then $|\mathcal{A}|, |\mathcal{B}|, |\mathcal{C}|, |\mathcal{D}| \gg q^{1-c}$. It follows that

$$\sqrt{\frac{q|\mathcal{D}|}{|\mathcal{C}|}} \left(\frac{q}{|\mathcal{D}|}\right)^k \ll q^{(1+c)/2+ck} \ll q^{1-c} \ll |\mathcal{A}|, |\mathcal{B}|.$$

Theorem 2 now follows from Lemma 15.

6. Proof of Theorem 3

We write $f(x,y)$ in the form

$$f(x,y) = \sum_{i=0}^{m} g_i(x) y^i,$$

where $g_i(x) \in \mathbb{F}_q[x]$. Let I denote the greatest i value with the property that $g_i(x)$ is not identically constant: $g_I(x) \not\equiv c$, and either $I = m$ or $g_{I+1}(x), \ldots, g_n(x)$ are identically constant. Since $f(x,y)$ is not of the form $g(x) + h(y)$, $I > 0$. Denote the degree of the polynomial $g_I(y)$ by D so that $D > 0$. Assume that $(I, q) = 1$. The following theorem is due to Gyarmati and Sárközy [4].

Theorem 16. *(cf. Theorem 3 in [4]) If $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D} \subset \mathbb{F}_q$, and the number of solutions of*

$$a + b = f(c,d), \quad a \in \mathcal{A}, b \in \mathcal{B}, c \in \mathcal{C}, d \in \mathcal{D},$$

is denoted by N, then we have

$$\left| N - \frac{|\mathcal{A}||\mathcal{B}||\mathcal{C}||\mathcal{D}|}{q} \right| \leqslant \Big(q(D + (I-1)q^{1/2}) |\mathcal{A}||\mathcal{B}||\mathcal{C}||\mathcal{D}| \Big)^{1/2}.$$

The following result is an analog of Corollary 9.

Corollary 17. *For all subsets $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D} \subset \mathbb{F}_q$, let $N_f^{\mathcal{C},\mathcal{D}}(\mathcal{A}, \mathcal{B})$ be the number of pairs $(a,b) \in \mathcal{A} \times \mathcal{B}$ such that $a + b \in f(\mathcal{C}, \mathcal{D})$. Then*

$$N_f^{\mathcal{C},\mathcal{D}}(\mathcal{A}, \mathcal{B}) \geqslant \frac{|\mathcal{D}|}{mq} |\mathcal{A}||\mathcal{B}| - \frac{1}{m} \sqrt{q(D + (I-1)q^{1/2}) \frac{|\mathcal{D}|}{|\mathcal{C}|}} \sqrt{|\mathcal{A}||\mathcal{B}|}.$$

Proof. For any $c \in \mathcal{C}$, let $N_f^c(\mathcal{A}, \mathcal{B}, \mathcal{D})$ denote the number of triples $(a,b,d) \in \mathcal{A} \times \mathcal{B} \times \mathcal{D}$ such that $a + b = f(c,d)$. By the pigeon-hole principle and Theorem 16, there exists $c_0 \in \mathcal{C}$ such that

$$N_f^{c_0}(\mathcal{A}, \mathcal{B}, \mathcal{D}) \geqslant \frac{|\mathcal{D}|}{q} |\mathcal{A}||\mathcal{B}| - \sqrt{q(D + (I-1)q^{1/2}) \frac{|\mathcal{D}|}{|\mathcal{C}|}} \sqrt{|\mathcal{A}||\mathcal{B}|}.$$

Besides, for any fixed a, b and c_0, $f(c_0, d) - a - b$ is a polynomial of degree m on d. Therefore, the number of d such that $a + b = f(c_0, d)$ is at most m. The corollary follows. □

As a consequence, we have the following lemma.

Lemma 18. *Suppose that $\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D} \subset \mathbb{F}_q$ with*

$$|\mathcal{A}|,|\mathcal{B}| \gg \frac{1}{m}\sqrt{q(D+(I-1)q^{1/2})\frac{|\mathcal{D}|}{|\mathcal{C}|}}\left(\frac{mq}{|\mathcal{D}|}\right)^k.$$

Then there are $a_1,\ldots,a_k \in \mathcal{A}$, $b_1,\ldots,b_k \in \mathcal{B}$ such that $a_i+b_j \in f(\mathcal{C},\mathcal{D})$ for all $1 \leqslant i,j \leqslant k$.

Proof. The proof of this lemma is similar to that of Lemma 14, except that we use Corollary 17 instead of Corollary 9 □

Let $c = c(k) = \frac{1}{4(k+2)}$ and $q \gg 1$, then $|\mathcal{A}|,|\mathcal{B}|,|\mathcal{C}|,|\mathcal{D}| \gg q^{1-c}$. It follows that

$$\frac{1}{m}\sqrt{q(D+(I-1)q^{1/2})\frac{|\mathcal{D}|}{|\mathcal{C}|}}\left(\frac{mq}{|\mathcal{D}|}\right)^k \ll q^{3/4}q^{c/2+kc} \ll q^{1-c} \ll |\mathcal{A}|,|\mathcal{B}|.$$

Theorem 3 now follows from Lemma 18

7. Proof of Theorem 6

Using multiplicative character sums, Gyarmati and Sárközy [4] proved the following theorem.

Theorem 19. *rm (cf. Theorem 4 in [4]) Suppose that $f(x,y) \in \mathbb{F}_q[x,y]$ and that the primitive kernel $H(x,y)$ of $f(x,y)$ is not of the form $c(K(x,y))^d$. Write $f(x,y) = F(x)G(y)H(x,y)$ in a unique way up to constant factors. Let r, s, n, m be the degrees of F, G, $f(x,y)$ in x, $f(x,y)$ in y, respectively. If $\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D} \subset \mathbb{F}_q$ and the number of solutions of*

$$ab = f(c,d), \quad a \in \mathcal{A}, b \in \mathcal{B}, c \in \mathcal{C}, d \in \mathcal{D},$$

is denoted by N, then we have

$$\left|N - \frac{|\mathcal{A}||\mathcal{B}||\mathcal{C}||\mathcal{D}|}{q}\right| < 4n^{1/2}q^{3/4}(|\mathcal{A}||\mathcal{B}||\mathcal{C}||\mathcal{D}|)^{1/2} + 7(r+s+n+(nm)^{1/2})q^{5/2}.$$

Similar to the previous sections, we have the following corollary.

Corollary 20. *For all subsets $\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D} \subset \mathbb{F}_q$, let $N_f^{\mathcal{C},\mathcal{D}}(\mathcal{A},\mathcal{B})$ be the number of pairs $(a,b) \in \mathcal{A}\times\mathcal{B}$ such that $ab = f(\mathcal{C},\mathcal{D})$. Then*

$$N_f^{\mathcal{C},\mathcal{D}}(\mathcal{A},\mathcal{B}) \geqslant \frac{|\mathcal{D}|}{mq}|\mathcal{A}||\mathcal{B}| - \frac{4n^{1/2}q^{3/4}}{m}\sqrt{\frac{|\mathcal{D}|}{|\mathcal{C}|}}\sqrt{|\mathcal{A}||\mathcal{B}|} - \frac{7(r+s+n+(nm)^{1/2})q^{5/2}}{m|\mathcal{C}|}.$$

Proof. For any $c \in \mathcal{C}$, let $N_f^c(\mathcal{A},\mathcal{B},\mathcal{D})$ denote the number of triples $(a,b,d) \in \mathcal{A}\times\mathcal{B}\times\mathcal{D}$ such that $ab = f(c,d)$. By the pigeon-hole principle and Theorem 19, there exists $c_0 \in \mathcal{C}$ such that

$$N_f^{c_0}(\mathcal{A},\mathcal{B},\mathcal{D}) \geqslant \frac{|\mathcal{D}|}{mq}|\mathcal{A}||\mathcal{B}| - \frac{4n^{1/2}q^{3/4}}{m}\sqrt{\frac{|\mathcal{D}|}{|\mathcal{C}|}}\sqrt{|\mathcal{A}||\mathcal{B}|} - \frac{7(r+s+n+(nm)^{1/2})q^{5/2}}{m|\mathcal{C}|}.$$

Besides, for any fixed a,b and c_0, $f(c_0,d)-ab$ is a polynomial of degree m on d. Therefore, the number of d such that $ab = f(c_0,d)$ is at most m. The corollary follows. □

We following lemma follows from Corollary 20 in a similar way that Lemma 14 follows from Corollary 9.

Lemma 21. *Suppose that $\mathcal{A},\mathcal{B},\mathcal{C},\mathcal{D} \subset \mathbb{F}_q$ with*

$$|\mathcal{A}|,|\mathcal{B}| \gg \frac{4n^{1/2}q^{3/4}}{m}\sqrt{\frac{|\mathcal{D}|}{|\mathcal{C}|}}\left(\frac{mq}{|\mathcal{D}|}\right)^k.$$

Then there are $a_1,\dots,a_k \in \mathcal{A}$, $b_1,\dots,b_k \in \mathcal{B}$ such that $a_ib_j \in f(\mathcal{C},\mathcal{D})$ for all $1 \leqslant i,j \leqslant k$.

Let $c = c(k) = \frac{1}{4(k+2)}$ and $q \gg 1$. Then $|\mathcal{A}|,|\mathcal{B}|,|\mathcal{C}|,|\mathcal{D}| \gg q^{1-c}$. It follows that

$$\frac{4n^{1/2}q^{3/4}}{m}\sqrt{\frac{|\mathcal{D}|}{|\mathcal{C}|}}\left(\frac{mq}{|\mathcal{D}|}\right)^k \ll q^{3/4}q^{c/2+kc} \ll q^{1-c} \ll |\mathcal{A}|,|\mathcal{B}|.$$

Theorem 6 now follows from Lemma 21.

8. Another problem

In [1], Csikvári, Sárközy and Gyarmati proposed some further related problems. One of these problems is the following (Problem 4 in [1]):

Is it true that for all $\varepsilon > 0$, there is a $k_0 = k_0(\varepsilon)$ such that if $k \in \mathbb{N}$, $k > k_0$, $p > p_0 = p_0(\varepsilon,k)$ and $\mathcal{A},\mathcal{B} \subset \mathbb{F}_q$ with

$$\min\{|\mathcal{A}|,|\mathcal{B}|\} > q^{\varepsilon},$$

then

$$a_1 + a_2 = b_1 \dots b_k, a_1, a_2 \in \mathcal{A}, b_1,\dots,b_k \in \mathcal{B} \tag{12}$$

can be solved?

In this section, we give a negative answer for this question by proving the following theorem.

Theorem 22. *For all $\varepsilon < 1/2$, $k \in \mathbb{N}$, there exists two sets $\mathcal{A}, \mathcal{B} \subset \mathbb{F}_q$ with*

$$|\mathcal{A}|, |\mathcal{B}| > q^{\varepsilon}$$

such that Eq. (12) cannot be solved.

Proof. Let ν be a generator of $\mathbb{F}_q^*$ and $t = \lceil q^{\varepsilon} \rceil + 1$. We choose $\mathcal{B} = \{1, \nu, \ldots, \nu^t\}$. Then $|\mathcal{B}| > q^{\varepsilon}$ and $\mathcal{B}^k = \{b_1 \ldots b_k : b_i \in \mathcal{B}\} = \{1, \nu, \ldots, \nu^{kt}\}$. Now we choose elements of $\mathcal{A}$ inductively. Let $\mathcal{T}_0 = \mathcal{B}/2 = \{b/2 : b \in \mathcal{B}\}$, $\mathcal{A}_0 = \{a_0\}$ with $a_0 \notin \mathcal{T}_0$. Suppose that we have $\mathcal{T}_i$ and $\mathcal{A}_i = \{a_0, \ldots, a_i\}$. We then construct $\mathcal{T}_{i+1}$ and $\mathcal{A}_{i+1}$ as follows:

$$\mathcal{T}_{i+1} = \mathcal{T}_i \cup (\mathcal{B}^k - a_i) \cup \{a_i\}, \mathcal{A}_{i+1} = \mathcal{A}_i \cup \{a_{i+1}\},$$

for some $a_{i+1} \notin \mathcal{T}_{i+1}$. It is easy to check that under this construction, $(\mathcal{A}_i + \mathcal{A}_i) \cap \mathcal{B}^k = \emptyset$ for all i. Since $|\mathcal{T}_{i+1}| \leqslant |\mathcal{T}_i| + |\mathcal{B}^k| + 1 \leqslant |\mathcal{T}_i| + tk + 1$, we can continue the process until $i(tk+1) < q$. Therefore, we can choose a set $\mathcal{A}$, such that $|\mathcal{A}| \geqslant \lceil (q-1)/(kt+1) \rceil \gg q^{\varepsilon}$ and $(\mathcal{A}+\mathcal{A}) \cap \mathcal{B}^k = \emptyset$. This completes the proof of the theorem. □

If $\mathbb{F}_q$ is not a prime field, we can do slightly better. Suppose that $q = p^2$ for some prime power p. We construct the Paley sum graph P_q^+ with the vertex set $\mathbb{F}_q$, and two vertices a, b are adjacent if and only if $a + b$ is a square residue. It is well known that the maximal clique of P_q^+ has size p. Since P_q^+ is self-symmetric, the maximal independent set of P_q^+ also has size p. Therefore, we can find a set $\mathcal{A}$ with $|\mathcal{A}| = q^{1/2}$ such that $a + a'$ is square non-residue for all $a, a' \in \mathbb{F}_q$. Let $\mathcal{B}$ be the set of all square residues, then $|\mathcal{B}| = q/2$ and Eq. (12) is not solvable in $\mathcal{A}, \mathcal{B}$.

References

[1] P. Csikvári, A. Sárközy and K. Gyarmati, Density and Ramsey type results on algebraic equations with restricted solution sets, *Combinatorica*, to appear.

[2] M. Z. Garaev, The sum-product estimate for large subsets of prime fields, *Proc. Amer. Math. Soc.* **136** (2008), 2735–2739.

[3] M. Z. Garaev and V. Garcia, The equation $x_1x_2 = x_3x_4 + \lambda$ in fields of prime order and applications, *J. Number Theory* **128**(9) (2008), 2520–2537.

[4] K. Gyarmati and A. Sárközy, Equations in finite fields with restricted solution sets, II (algebraic equations), *Acta Math. Hungar.* **119** (2008), 259–280.

[5] A. Sárközy, On sums and products of residues modulo p, *Acta. Arith.* **118** (2005), 403–409.

[6] A. Sárközy, On products and shifted products of residues modulo p, *Integers* **8**(2) (2008), A9.

[7] I. E. Shparlinski, On the solvability of bilinear equations in finite fields, *Glasgow Math J* **50** (2008), 523–529.