ON SOME GENERALISATIONS OF THE ERDŐS DISTANCE PROBLEM OVER FINITE FIELDS

IGOR E. SHPARLINSKI

We use exponential sums to obtain new lower bounds on the number of distinct distances defined by all pairs of points \((a, b) \in A \times B\) for two given sets \(A, B \subseteq F_q^n\) where \(F_q\) is a finite field of \(q\) elements and \(n \geq 1\) is an integer.

1. INTRODUCTION

Given a ring \(R\) and a finite set \(E \subseteq R^n\) we use \(\Delta(R^n, E)\) to denote the number of distinct distances defined by the pairs of points from \(E\), that is,

\[
\Delta(R^n, E) = \left| \{d(x, y) \mid x, y \in E\} \right|,
\]

where for \(x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in R^n\) we define

\[
d(x, y) = \sum_{j=1}^{n} (x_j - y_j)^2.
\]

Throughout this paper, the Vinogradov symbols \(\gg\) and \(\ll\) have their usual meanings (we recall that \(U \ll V, V \gg U\) are both equivalent to the assertion that \(U = O(V)\)). The constants implied by them may depend on the dimension \(n\) and the degree \(k\) of certain polynomials which appear in our generalisation of the original problem.

Then the Erdős Distance Conjecture asserts that over the real numbers, that is, for \(R = R\), the bound

\[
\Delta(R^n, E) \gg |E|^{2/n}
\]

holds for any finite set \(E \subseteq R^n\). Despite that there are some very interesting lower bounds on \(\Delta(R^n, E)\), this conjecture is still widely open in any dimension including \(n = 2\). For some recent achievements and generalisations. See [1, 2, 3, 4, 5, 6] and references therein.

Received 28th November, 2005
The author is very grateful to Alex Iosevich and Vsevolod Lev for many useful discussions and encouragement. During the preparation of this paper, the author was supported in part by ARC grant DP0556431.

Copyright Clearance Centre, Inc. Serial-fee code: 0004-9727/06 $A2.00+0.00.
Iosevich and Rudnev [5] have recently considered this problem for sets over finite fields, that is, for $\Delta(\mathbb{F}_q^n, \mathcal{E})$. Among several other results, they show that, for any set $\mathcal{E} \subseteq \mathbb{F}_q^n$,

$$
\Delta(\mathbb{F}_q^n, \mathcal{E}) \gg \min\{q, q^{-(n-1)/2}|\mathcal{E}|\}.
$$

Here we consider two generalisations of this problem. Given $n$ polynomials $f_j(X, Y) \in \mathbb{F}_q[X, Y]$, $j = 1, \ldots, n$, we define the generalised distance

$$
d_\ell(x, y) = \sum_{j=1}^n f_j(x_j, y_j).
$$

where $f = (f_1, \ldots, f_n)$.

Now, for two sets $A, B \subseteq \mathbb{F}_q^n$ we define

$$
\Gamma_\ell(\mathbb{F}_q^n, A, B) = \left| \{d_\ell(x, y) \mid x \in A, y \in B\} \right|.
$$

In the special case of the Euclidean distance function $f_0 = (f_{1,0}, \ldots, f_{n,0})$, where $f_{j,0}(X, Y) = (X - Y)^2$, $j = 1, \ldots, n$, we simply write

$$
\Gamma_{f_0}(\mathbb{F}_q^n, A, B) = \Gamma(\mathbb{F}_q^n, A, B)
$$

thus $\Gamma(\mathbb{F}_q^n, \mathcal{E}, \mathcal{E}) = \Delta(\mathbb{F}_q^n, \mathcal{E})$.

Here we suggest a slightly different approach to treat these extensions. Although in the special case of $\Delta(\mathbb{F}_q^n, \mathcal{E})$ our results are generally weaker than those of Iosevich and Rudnev [5], in some particular instances we obtain slightly stronger statements. For example, we show that

$$
\Delta(\mathbb{F}_q^n, \mathcal{E}) = q \quad \text{for} \quad |\mathcal{E}| \geq q^{n/2+1}
$$

which does not follow from (2).

2. Sets of Euclidean Distances

THEOREM 1. For arbitrary sets $A, B \subseteq \mathbb{F}_q^n$,

$$
\Gamma(\mathbb{F}_q^n, A, B) > q - \frac{q^{n+2}}{|A||B|}.
$$

PROOF: Let $\chi$ be a nontrivial additive character of $\mathbb{F}_q$. See [7] for basis properties of additive characters. In particular, we repeatedly use the identity

$$
\sum_{s \in \mathbb{F}_q} \chi(st) = \begin{cases} 
0, & \text{if } t \in \mathbb{F}_q^*, \\
q, & \text{if } t = 0.
\end{cases}
$$
We consider character sums
\[ S(a, A, B) = \sum_{x \in A} \sum_{y \in B} \chi(ad(x, y)), \quad a \in \mathbb{F}_q, \]
where as before \(d(x, y)\) is given by (1).

By the Cauchy inequality we derive,
\[
|S(a, A, B)|^2 \leq |A| \left| \sum_{x \in A} \sum_{y \in B} \chi(ad(x, y)) \right|^2 \leq |A| \sum_{y \in B} \left| \sum_{x \in \mathbb{F}_q} \chi(ad(x, y)) \right|^2
\]
\[
= |A| \sum_{y \in B} \chi \left( a \sum_{j=1}^n (x_j - y_j)^2 - (x_j - z_j)^2 \right) \chi \left( a \sum_{j=1}^n \bar{x}_j (z_j - y_j) \right)
\]
\[
= |A| \sum_{y \in B} \chi \left( a \sum_{j=1}^n (y_j^2 - z_j^2) \right) \sum_{x \in \mathbb{F}_q} \chi \left( a \sum_{j=1}^n \bar{x}_j (z_j - y_j) \right)
\]
\[
= |A| \sum_{y \in B} \chi \left( a \sum_{j=1}^n (y_j^2 - z_j^2) \right) \prod_{j=1}^n \sum_{x \in \mathbb{F}_q} \chi \left( a \bar{x}_j (z_j - y_j) \right)
\]
\[
= |A| |B| q^n
\]
since if \(y \neq z\) then by (5) at least one inner sum in the product vanishes. Therefore,
\[
|S(a, A, B)| \leq \sqrt{|A| |B|} q^n.
\]

Let \(N(\lambda)\) be the number of solutions to the equation
\[ d(x, y) = \lambda, \quad x \in A, y \in B. \]
Then by (5) we have
\[
N(\lambda) = \frac{1}{q} \sum_{x \in A} \sum_{y \in B} \sum_{a \in \mathbb{F}_q} \chi \left( a \left( d(x, y) - \lambda \right) \right)
\]
\[
= \frac{1}{q} \sum_{a \in \mathbb{F}_q} \chi(-a\lambda) S(a, A, B).
\]
Separating the term \(|A||B| q^{-1}\) corresponding to \(a = 0\), we obtain,
\[
N(\lambda) - \frac{|A||B|}{q} = \frac{1}{q} \sum_{a \in \mathbb{F}_q} \chi(-a\lambda) S(a, A, B).
\]
Hence,
\[
\sum_{\lambda \in \mathbb{F}_q} \left| N(\lambda) - \frac{|A||B|}{q} \right|^2 = \frac{1}{q^2} \sum_{\lambda \in \mathbb{F}_q} \sum_{a,b \in \mathbb{F}_q} \chi \left( (b - a)\lambda \right) S(a, A, B) \overline{S(b, A, B)}
\]
\[
= \frac{1}{q^2} \sum_{a,b \in \mathbb{F}_q} S(a, A, B) \overline{S(b, A, B)} \sum_{\lambda \in \mathbb{F}_q} \chi \left( (b - a)\lambda \right)
\]
\[
= \frac{1}{q} \sum_{a \in \mathbb{F}_q} \left| S(a, A, B) \right|^2,
\]
since by (5) the sum over $\lambda$ vanishes unless $a = b$. Thus,

$$\sum_{\lambda \in \mathbb{F}_q} \left| N(\lambda) - \frac{|A||B|}{q} \right|^2 < |A||B| q^n.$$ 

Each term with $N(\lambda) = 0$ contributes $|A|^2|B|^2/q^3$ to the left hand side. Therefore

$$(q - \Gamma(\mathbb{F}_q^n, A, B)) \frac{|A|^2|B|^2}{q^3} > |A||B| q^n$$

which yields the desired result.

In particular, Theorem 1 immediately implies (4).

We now introduce one more approach, to prove a different estimate which is stronger than that of Theorem 1 when one set is much smaller than the other.

**Theorem 2.** For every odd $q$ and arbitrary sets $A, B \subseteq \mathbb{F}_q^n$,

$$\Gamma(\mathbb{F}_q^n, A, B) \gg \min\{ q^{1/3}, |A|^{1/3}|B|^{2/3} q^{-(2n-1)/3} \}.$$ 

**Proof:** We define character sums $S(a, A, B)$ by (6), as in the proof of Theorem 1. For $a \in \mathbb{F}_q^n$, by the Hölder inequality, we derive,

$$|S(a, A, B)|^4 \leq |A|^3 \left| \sum_{x \in A} \sum_{y \in B} \chi(a d(x, y)) \right|^4 \leq |A|^3 \left| \sum_{y \in B} \sum_{x \in \mathbb{F}_q^n} \chi(a d(x, y)) \right|^4$$

$$= |A|^3 \sum_{u,v,y,z \in B} \sum_{x \in \mathbb{F}_q^n} \chi\left(a(d(x, y) + d(x, z) - d(x, u) - d(x, v))\right)$$

$$= |A|^3 \sum_{u,v,y,z \in B} \sum_{x \in \mathbb{F}_q^n} \chi\left(a \sum_{j=1}^n \left( (x_j - y_j)^2 + (x_j - z_j)^2 - (x_j - u_j)^2 - (x_j - v_j)^2 \right) \right)$$

$$= |A|^3 \sum_{u,v,y,z \in B} \chi\left(a \sum_{j=1}^n (y_j^2 + z_j^2 - u_j^2 - v_j^2) \right)$$

$$\sum_{x \in \mathbb{F}_q^n} \chi\left(2a \sum_{j=1}^n x_j(u_j + v_j - y_j - z_j)\right).$$

Since $q$ is odd and $a \in \mathbb{F}_q^*$, then by (5), the sum over each $x_j$, $j = 1, \ldots, n$, vanishes, unless $u_j + v_j = y_j + z_j$. Hence for $a \in \mathbb{F}_q^*$ we have

$$|S(a, A, B)|^4 \leq |A|^3 q^n \sum_{u,v,y,z \in B \atop u+v=y+z} \chi\left(a \sum_{j=1}^n (y_j^2 + z_j^2 - u_j^2 - v_j^2) \right).$$
We also have \( S(0, \mathcal{A}, \mathcal{B}) = |\mathcal{A}| \cdot |\mathcal{B}| \). Therefore, again by (5), we derive the inequality

\[
\sum_{\alpha \in F_q} |S(\alpha, \mathcal{A}, \mathcal{B})|^4 = \sum_{\alpha \in F_q} |S(\alpha, \mathcal{A}, \mathcal{B})|^4 + |\mathcal{A}|^4 |\mathcal{B}|^4
\]

\[
\leq |\mathcal{A}|^3 q^n \sum_{\alpha \in F_q} \sum_{u, v, y, z \in \mathcal{B}} \chi \left( \alpha \sum_{j=1}^n (y_j^2 + z_j^2 - u_j^2 - v_j^2) \right) + |\mathcal{A}|^4 |\mathcal{B}|^4
\]

\[
= |\mathcal{A}|^3 q^n \sum_{u, v, y, z \in \mathcal{B}} \chi \left( \alpha \sum_{j=1}^n (y_j^2 + z_j^2 - u_j^2 - v_j^2) \right) + |\mathcal{A}|^4 |\mathcal{B}|^4
\]

where \( T \) is the number of solutions to the system of \( n + 1 \) equations

\[
\sum_{j=1}^n (u_j^2 + v_j^2) = \sum_{j=1}^n (y_j^2 + z_j^2),
\]

\[
u_j + v_j = y_j + z_j, \quad j = 1, \ldots, n,
\]

in \( u, v, y, z \in \mathcal{B} \). There are exactly \( |\mathcal{B}|^2 \) possible values for \( y, z \in \mathcal{B} \). When \( y, z \) are fixed, substituting \( v_j = y_j + z_j - u_j \) in the first equation, we obtain a nontrivial quadratic equation for \( u_1, \ldots, u_n \) (since \( q \) is odd). Thus there are \( O(q^n) \) possible vectors \( u \), which now define \( v \) uniquely. Therefore \( T \leq |\mathcal{B}|^2 q^{n-1} \) which leads to the bound

\[
\sum_{\alpha \in F_q} |S(\alpha, \mathcal{A}, \mathcal{B})|^4 \ll |\mathcal{A}|^3 |\mathcal{B}|^2 q^{2n} + |\mathcal{A}|^4 |\mathcal{B}|^4.
\]

As in the proof of Theorem 1, we use \( N(\lambda) \) to denote the number of solutions to (7). Then from (8) we deduce

\[
\sum_{\lambda \in F_q} N(\lambda)^4 = \frac{1}{q^4} \sum_{\lambda \in F_q} \sum_{a,b,c,d \in F_q} \chi(\lambda(a + b + c + d))
\]

\[
\times S(a, \mathcal{A}, \mathcal{B}) S(b, \mathcal{A}, \mathcal{B}) S(c, \mathcal{A}, \mathcal{B}) S(d, \mathcal{A}, \mathcal{B})
\]

\[
= \frac{1}{q^3} \sum_{a,b,c,d \in F_q} S(a, \mathcal{A}, \mathcal{B}) S(b, \mathcal{A}, \mathcal{B}) S(c, \mathcal{A}, \mathcal{B}) S(d, \mathcal{A}, \mathcal{B}).
\]
By the Hölder inequality
\[
\sum_{\lambda \in \mathbb{F}_q} N(\lambda)^4 \leq \frac{1}{q^3} \left( \sum_{a,b,c,d \in \mathbb{F}_q \atop a+b+c+d=0} |S(a, A, B)|^4 \right)^{1/4} \left( \sum_{a,b,c,d \in \mathbb{F}_q \atop a+b+c+d=0} |S(b, A, B)|^4 \right)^{1/4} \\
\times \left( \sum_{a,b,c,d \in \mathbb{F}_q \atop a+b+c+d=0} |S(c, A, B)|^4 \right)^{1/4} \left( \sum_{a,b,c,d \in \mathbb{F}_q \atop a+b+c+d=0} |S(d, A, B)|^4 \right)^{1/4}
\]
\[
= \frac{1}{q^3} \sum_{a,b,c,d \in \mathbb{F}_q \atop a+b+c+d=0} |S(a, A, B)|^4 \leq \frac{1}{q} \sum_{a \in \mathbb{F}_q} |S(a, A, B)|^4 \\
\leq |A|^3 |B|^2 q^{2n-1} + |A|^4 |B|^4 q^{-1}.
\]

Clearly
\[
\sum_{\lambda \in \mathbb{F}_q} N(\lambda) = |A| |B|.
\]

Now, by the Hölder inequality again,
\[
(|A| |B|)^4 = \left( \sum_{\lambda \in \mathbb{F}_q} N(\lambda) \right)^4 \leq \Gamma_{\mathbb{F}_q^n, A, B}^3 \sum_{\lambda \in \mathbb{F}_q} N(\lambda)^4 \\
\ll \Gamma_{\mathbb{F}_q^n, A, B}^3 (|A|^3 |B|^2 q^{2n-1} + |A|^4 |B|^4 q^{-1})
\]

which implies the desired result. \(\square\)

We see that Theorem 2 is nontrivial for \(|A||B|^2 \geq C q^{2n-1}\) for some constant \(C > 0\) depending only on \(n\).

3. SETS OF GENERALISED DISTANCES

The following bound follows the same lines as the proof of Theorem 1.

**Theorem 3.** Let \(f = (f_1, \ldots, f_n)\), where each of the polynomials \(f_j(X,Y) \in \mathbb{F}_q[X,Y], \ j = 1, \ldots, n\), is of degree at most \(k\) and is not of the form \(f_j(X,Y) = g_j(X) + h_j(Y)\) with \(g_j(X) \in \mathbb{F}_q[X], \ h_j(Y) \in \mathbb{F}_q[Y]\). Then, for arbitrary sets \(A, B \subseteq \mathbb{F}_q^n\),
\[
\Gamma_f(\mathbb{F}_q^n, A, B) = q + O\left(\frac{q^{3n/2+2}}{|A| |B|}\right).
\]

**Proof:** As before, we fix a nontrivial additive character \(\chi\) of \(\mathbb{F}_q\) and consider character sums

\[
S_f(a, A, B) = \sum_{x \in A} \sum_{y \in B} \chi(ad_f(x, y)), \quad a \in \mathbb{F}_q,
\]

where \(d_f(x, y)\) is given by (3).
Arguing as in the proof of Theorem 1, by the Cauchy inequality we derive,

\[ |S_f(a, A, B)|^2 \leq |A| \sum_{y,z \in B} \prod_{j=1}^n \sum_{x_j \in F_q} \chi(a(f_j(x_j, y_j) - f_j(x_j, z_j))). \]

If \( f_j(X, y_j) - f_j(X, z_j) \) is constant, the corresponding sum over \( x_j \) is equal to \( q \) by absolute value, otherwise we estimate this sum as \( O(q^{1/2}) \) by the Weil bound.

It is easy to see that if a polynomial \( f(X, Y) \in F_q[X, Y] \) of degree \( \deg f \leq k \) is not of the form \( f(X, Y) = g(X) + h(Y) \) with \( g(X) \in F_q[X] \), \( h(Y) \in F_q[Y] \), then for every \( y \in F_q \), there are at most \( k \) values of \( z \) such that \( f(X, y) - f(X, z) \) is constant.

For every \( y \in B \) and an integer \( \nu \in \{0, \ldots, n\} \), there are \( O(q^{n-\nu}) \) vectors \( z \in B \) for which \( f_j(X, y_j) - f_j(X, z_j) \) is constant for exactly \( \nu \) values of \( j \in \{1, \ldots, n\} \). Therefore

\[
|S_f(a, A, B)|^2 \leq |A| \sum_{\nu=0}^n |B| q^{n-\nu} q^\nu q^{(n-\nu)/2} = |A| |B| q^{3n/2} \sum_{\nu=0}^n q^{-\nu/2} \ll |A| |B| q^{3n/2}.
\]

Let \( N_f(\lambda) \) be the number of solutions to the equation

\[ d_f(x, y) = \lambda, \quad x \in A, y \in B. \]

Then by (5) we have the following analogues of (8)

\[ N_f(\lambda) = \frac{1}{q} = \frac{1}{q} \sum_{a \in F_q} \chi(-a\lambda)S_f(a, A, B). \]

Separating the term \(|A| |B| q^{-1}\) corresponding to \( a = 0 \), as in the proof of Theorem 1, we obtain,

\[
\sum_{\lambda \in F_q} \left| N_f(\lambda) - \frac{|A||B|}{q} \right|^2 < |A| |B| q^{3n/2}.
\]

Each term with \( N_f(\lambda) = 0 \) contributes \( |A|^2 |B|^2 / q^2 \) to the left hand side. Therefore

\[
(q - \Gamma_f(F_q^n, A, B)) \frac{|A|^2 |B|^2}{q^2} > |A| |B| q^{3n/2}
\]

which implies the desired result. \( \square \)

In particular, we see that there is a constant \( C > 0 \), depending only on \( n \) and \( k \) such that for any \( f \) satisfying the conditions of Theorem 3 we have \( \Gamma_f(F_q^n, A, B) > q/2 \) provided that \( |A| |B| \geq C q^{3n/2+2} \).
REFERENCES


