Geometric problems in $\mathbb{F}_q^d$, and a generalization of a sum estimate

Byron Osterweil

May 7, 2020

1 Introduction

A number of problems in discrete geometry relate to the notion of distance sets. That is, given a collection of n points in the plane, how many distinct distances do we expect to get between these points? Erdős conjectured in 1946 that for any arrangement of N points, there number of distinct distances is $\gtrsim \frac{N}{\sqrt{\log N}}$, where here and going forward $A \gtrsim B$ means $A \geq CB$ for some constant C [1]. This conjecture was proven by Guth and Katz in 2015 [2].

A related conjecture by Kenneth Falconer states [3]:

If $d \geq 2$ and $E \subseteq \mathbb{R}^d$ is compact, then

$$\dim(E) > \frac{d}{2} \implies |\Delta(E)| > 0$$

where $\dim(E)$ is the Hausdorff dimension of E and $|\Delta(E)|$ is the Lebesgue measure of the set of all distances between points in E.

This conjecture is as of yet still unproven, but M. Erdoğan proved the following partial result in [5]:

Theorem 1.1. *(Erdoğan, 2005) Let $d \geq 3$, $E \subseteq \mathbb{R}^d$ compact, such that*

$$\dim(E) > \frac{d}{2} + \frac{1}{3}.$$

Then $|\Delta(E)| > 0$.

The Erdős and Falconer problems can also be treated in the context of other vector spaces. In particular, significant research has been done into these problems in a discrete geometry context, translating them to vector spaces over finite field. Let q be a prime power, and $\mathbb{F}_q$ the finite field of q elements. Then given a positive integer d we define a norm on the d-dimensional vector space $\mathbb{F}_q^d$ by

$$||x|| = x_1^2 + x_2^2 + ... + x_d^2, \tag{1}$$

where here the square root found on the analogous Euclidean norm is omitted here on the grounds that taking square roots in $\mathbb{F}_q$ is often not possible. Let E be a subset of $\mathbb{F}_q^d$. Define the distance set of E:

$$\Delta(E) = \{||x - y|| : x, y \in \mathbb{F}_q^d\}.$$

The Erdős-Falconer distance problem in this context asks how small are large $\Delta(E)$ can be compared to the size of E. In the particular case of finite fields a natural question to ask is how large must E be in order that $\Delta(E) = \mathbb{F}_q^d$, or at least such that $|\Delta(E)| \geq \frac{q}{2}$. Some progress has been made on these questions is recent years. In 2007, A. Iosevich and M. Rudnev proved the following results via methods of discrete Fourier analysis [6]:

Theorem 1.2 (Iosevich, Rudnev 2007)**.** *Let $E \subseteq \mathbb{F}_q^d$ such that $|E| \gtrsim Cq^{\frac{d}{2}}$ for C sufficiently large. Then*

$$|\Delta(E)| \gtrsim min\left\{q, \frac{|E|}{q^{\frac{d-1}{2}}}\right\} \tag{2}$$

They then go on to expand this to a stronger result:

Theorem 1.3 (Iosevich, Rudnev 2007)**.** *Let $E \subseteq \mathbb{F}_q^d$ such that $|E| \geq Cq^{\frac{d+1}{2}}$ for some C sufficiently large. Then $\Delta(E) = \mathbb{F}_q$.*

Shortly thereafter Hart, Iosevich, Koh, and Rudnev proved that when d is odd, the above value of $\frac{d+1}{2}$ is sharp [7]. A paper by Chapman et al [4] further develops these results, proving the following:

Theorem 1.4 (Chapman et al, 2009)**.** *Let $E \subseteq \mathbb{F}_q^2$. If $q \equiv 3$ mod 4 and $E \geq q^{4/3}$, then*

$$|\Delta(E)| > \frac{q}{q + \sqrt{3}} \tag{3}$$

If $q \equiv 1$ mod 4, q sufficiently large and $E \geq q^{4/3}$, then there exists $0 < \epsilon_1 < 1$ such that

$$|\Delta(E)| > \epsilon_q q, \tag{4}$$

where $\epsilon_q \to \frac{1}{1+\sqrt{3}}$ as $q \to \infty$.

In the special case of $q = p$, $d = 2$, the authors in [8] use a novel approach to derive the following bound:

Theorem 1.5. *Let $E \subseteq \mathbb{F}_p$ such that $p \equiv 3$ mod 4. Suppose further that $|E| \lesssim q^{\frac{1558}{1489}}$. Then we have*

$$|\Delta(E)| \gg |E|^{\frac{424}{779}} = |E|^{\frac{1}{2}+\frac{69}{1558}}$$

where $X \gg Y$ means that for every $\epsilon > 0$ there exists C_ϵ such that $X \leq C_\epsilon q^\epsilon Y$.

Another topic that has been studied previously is that of the distance graph on subsets of $\mathbb{F}_q^d$. Given $E \subseteq \mathbb{F}_q^d$, define a distance graph on E by fixing an element $t \in \mathbb{F}_q$, taking each point in $\mathbb{F}_q^d$ to be a vertex, and connecting two vertices x and y by an edge if $||x-y|| = t$. In [9], the authors prove the following results:

Theorem 1.6. *Let $E \subseteq \mathbb{F}_q^d$ for $d \geq 2$, and $|E| > \frac{2k}{\ln 2} q^{\frac{d+1}{2}}$. Suppose for $1 \leq i \leq k$ we have $0 \neq t_i \in \mathbb{F}_q$, and $\vec{t} = (t_1 ... t_k)$. Define*

$$C_k(\vec{t}) = |\{(x^1 ... x^{k+1}) \in E \times ... \times E : ||x^i - x^{i+1}|| = t_i, 1 \leq i \leq k\}|.$$

Then

$$C_k(\vec{t}) = \frac{|E|^{k+1}}{q^k} + \mathcal{D}_k(\vec{t})$$

where

$$|\mathcal{D}_k(\vec{t})| \leq \frac{2k}{\ln 2} q^{\frac{d+1}{2}} \frac{|E|^k}{q^k}$$

The authors then go on to apply this result to one about chain in the distance graph of E

Definition 1.1. *If G is a simple graph a path of length k in G is a sequence of vertices $v_1 ... v_{k+1}$ such that for each $i \in \{1 ... k\}$, v_i is connected to v_{i+1} by an edge. A path of length k in G is said to be non-overlapping if each v_i is distinct.*

Corollary 1.6.1. *Suppose $E \subseteq \mathbb{F}_q^d$ such that $|E| \geq \frac{4k}{\ln 2} q^{\frac{d+1}{2}}$. Then the distance graph of E contains a non-overlapping chain of length k.*

In [9] the authors go on to prove one further result, this one concerning k-stars, a highly related structure. Fixing an x in E, we can look at all the vectors which are some given distance from x, a configuration which we call a k-star.

Theorem 1.7. *Let $E \subseteq \mathbb{F}_q^d$, and suppose for each $1 \leq i \leq k$ we have $t_i \neq 0$. Let $\vec{t} = (t_1 ... t_k)$, and define*

$$\nu_k \vec{t} = |\{(x, x^1, ... x^k) \in E^{k+1} : ||x - x^i|| = t_i,\ x^i = x^i \iff i = j\}|. \tag{5}$$

Then if $|E| > 12 q^{\frac{d+1}{2}}$, then $\nu_k(\vec{t}) > 0$ for any $k < \frac{|E|}{12 q^{\frac{d+1}{2}}}$.

If $|E| > 12 q^{\frac{d+3}{2}}$, then $\nu_k(\vec{t}) > 0$ for any $k < \frac{|E|}{12q}$.

2 Background Results

Before beginning our proofs, we present some definitions and established theorems that will be used here and further on:

Given a function $f : \mathbb{F}_q^d \to \mathbb{C}$ and χ a non-trivial additive character for $\mathbb{F}_q$, then we define the Fourier transform of f by

$$\hat{f}(m) = q^{-d} \sum_{x \in \mathbb{F}_q^d} \chi(-m \cdot x) f(x).$$

We also have analogues for the Plancherel and inversion formulae familiar from the non-discrete Fourier transform:

$$\sum_{m \in \mathbb{F}_q^d} |\hat{f}(m)|^2 = q^{-d} \sum_{x \in \mathbb{F}_q^d} |f(x)|^2$$

and

$$f(x) = \sum_{m \in \mathbb{F}_q^d} \hat{f}(m) \chi(m \cdot x).$$

Lastly we define discrete convolution. If $f, g : \mathbb{F}_q^d \to \mathbb{C}$, then

$$(f * g)(x) = \sum_{y \in \mathbb{F}_q^d} f(y) g(x - y).$$

From this we have a version of Young's convolution inequality:

Theorem 2.1 (Young)**.** *If $f, g : \mathbb{F}_q^d \to \mathbb{R}$, and we have*

$$\frac{1}{s} + \frac{1}{t} = \frac{1}{r} + 1$$

For $1 \leq s,\ t \leq r \leq \infty$, then

$$||f * g||_r \leq ||f||_s ||g||_t$$

Here the requirement that f is an L^s function and g is an L^t function is unnecessary here, as over a finite domain all functions belong to all L^p spaces.

3 Results

In many results of this type, an element that recurs frequently is the sum

$$\sum_{||x-y||=t} f(x) g(y), \tag{6}$$

where f and g are real-valued functions and t is a fixed element of $\mathbb{F}_q$. In [9], the prove the following bound on this sum, as an intermediate step towards the final results. Because the method of proof is instructive as to the methods of analysis on finite fields, as well as to the calculations that follow, we will include the proof here as well.

Theorem 3.1. *1 Let $f, g : \mathbb{F}_q^d \to \mathbb{R}^+$. Let $S_t = \{x \in \mathbb{F}_q^d : ||x|| = t\}$, where $||x|| = x_1^2 + ... + x_d^2$, and $t \neq 0$. Then*

$$\sum_{x,y\in\mathbb{F}_q^d} f(x)g(y)S_t(x-y) = \frac{1}{q}||f||_1||g||_1 + D(f,g) \tag{7}$$

where

$$|D(f,g)| \leq 2q^{\frac{d-1}{2}}||f||_2||g||_2. \tag{8}$$

Proof. To begin, we write the following equalities:

$$\begin{aligned}
&\sum_{x,y\in\mathbb{F}_q^d} f(x)g(y)S_t(x-y) \\
&= \frac{1}{q}\sum_{x,y\in\mathbb{F}_q^d} f(x)g(x+y)\sum_{s\in\mathbb{F}_q}\chi(s||y||-t) \\
&= \frac{1}{q}||f||_1||g||_1 + q^{d-1}\sum_{m\in\mathbb{F}_q^d}\hat{f}(m)\hat{g}(-m)\sum_{s\neq 0}\chi(-st)\sum_{y\in\mathbb{F}_q^d}\chi(s||y||-y\cdot m) \\
&= \frac{1}{q}||f||_1||g||_1 + q^{d-1}\sum_{m\in\mathbb{F}_q^d}\hat{f}(m)\hat{g}(-m)\sum_{s\neq 0}\chi\left(-st-\frac{||m||}{4s}\right)\sum_{y\in\mathbb{F}_q^d}\chi\left(s\left|\left|y-\frac{m}{2s}\right|\right|\right).
\end{aligned}$$

The sum on the far right is a Gauss sum, which gives us

$$\frac{1}{q}||f||_1||g||_1 + u_1 q^{\frac{3d-2}{2}}\sum_{m\in\mathbb{F}_q^d}\hat{f}(m)\hat{g}(-m)\sum_{s\neq 0}\left(\frac{s}{q}\right)^d\chi\left(-st-\frac{||m||}{4s}\right) \tag{9}$$

Where u_1 is a complex number of modulus 1 and $\left(\frac{\cdot}{\cdot}\right)$ is the Legendre symbol [10]. We then use the following fact about Kloosterman/Salié sums [10]:

Lemma 3.2. *If $a \neq 0$ or $b \neq 0$, then*

$$\left|\sum_{s\neq 0}\left(\frac{s}{q}\right)^d\chi(as+bs^{-1})\right| \leq 2\sqrt{q} \tag{10}$$

Using this lemma here we get that equation (8) equals

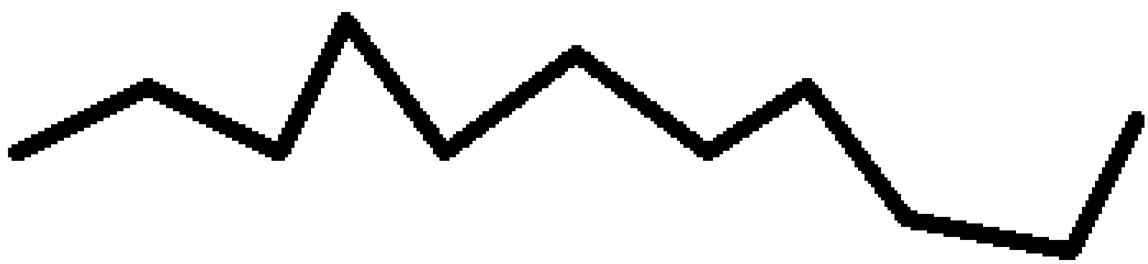
$$\frac{1}{q}||f||_1||g||_2 + 3u_2 q^{\frac{3d-1}{2}}\sum_{m\in\mathbb{F}_q^d}|\hat{f}(m)\hat{g}(-m)| \tag{11}$$

where $u_2 \in \mathbb{C}$, $|u_2| \leq 1$. By Cauchy-Schwartz, the above equation is less than or equal to

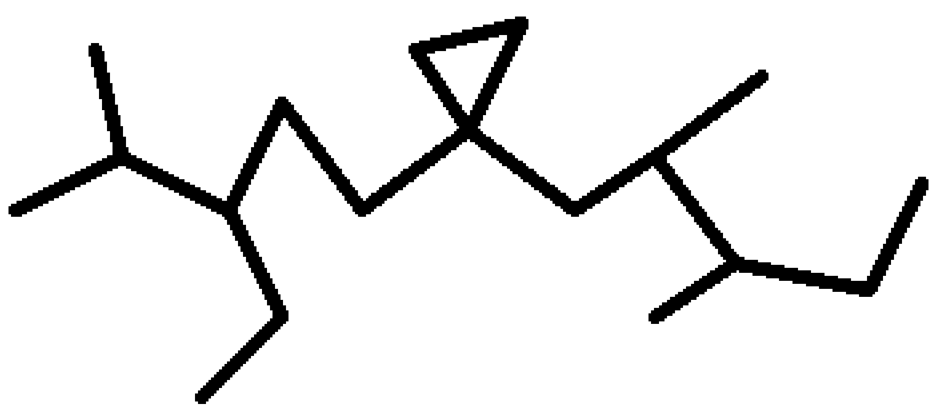
$$\frac{1}{q}||f||_1||g||_1 + 2u_2 q^{\frac{3d-1}{2}}\left(\sum_{m_1,m_2\in\mathbb{F}_q^d}|\hat{f}(m_1)|^2|\hat{g}(-m_2)|^2\right)^{1/2} \tag{12}$$

and we are done by an application of Plancherel. $\square$

The cited paper this proof originates in uses this result to prove a statement regarding the existence of long paths of in the distance graph of E, where E is a subset of $\mathbb{F}_q^d$. The distance graph of E is defined to be the graph whose vertices are the points of E and where two vertices are connected by an edge if the distance between the two vertices is some fixed distance $t \in \mathbb{F}_q$. Our goal is to use this result to prove a more general result for an analogous sum over three variables, then extending the method to cover the more general case of $k+1$ functions, in the hopes that this result can be extended to prove similar results regarding more general graphs in addition to the existence of long paths. To illustrate this, suppose the following figure represents a long path as described in [9]:



Then our hope is that the following generalizations of the above sum could allow for the study of more general subsets of the distance graph, via appropriate choices of the functions in the sum. Such a hypothetical example is illustrated below:



The statement of this result is as follows:

Theorem 3.3. *Let* $f, g, h : \mathbb{F}_q^d \to \mathbb{R}^+$, *and let* $S_t = \{x \in \mathbb{F}_q^d : ||x|| = t\}$, *where* t *is a fixed, non-zero element of* $\mathbb{F}_q$. *Then:*

$$\sum_{x,y,z\in\mathbb{F}_q^d} f(x)g(y)h(z)S_t(x-y)S_t(y-z) = \frac{1}{q^2}||f||_1||g||_1||h||_1 + D(f,g,h), \tag{13}$$

where $D(f,g,h)$ is an error term such that

$$|D(f,g,h)| \le 2q^{\frac{d-3}{2}}||f||_1||g||_2||h||_2 + 2q^{\frac{d-1}{2}}||f||_2||g||_\infty||h||_2|S_t| \tag{14}$$

Proof. Let $f,g,h:\mathbb{F}_q^d \to \mathbb{R}^+$, $S_t = \{x\in\mathbb{F}_q^d : ||x|| = t\}$. We wish to find an estimate for

$$\sum_{x,y,z\in\mathbb{F}_q^d} f(x)g(y)h(z)S_t(x-y)S_t(y-z) \tag{15}$$

Define $\tilde{g}(y) = g(y)(h*S_t)(y-z)$, so (15) equals

$$\sum_{x,y\in F_q^d} f(x)\tilde{g}(y) \tag{16}$$

Using theorem 2.1, this is equal to

$$\frac{1}{q}||f||_1||\tilde{g}||_1 + D(f,\tilde{g}) \tag{17}$$

where $D(f,\tilde{g})$ is an error term such that

$$|D(f,\tilde{g})| \le 2q^{\frac{d-1}{2}}||f||_2||\tilde{g}||_2$$

So in particular we need an estimate for $||\tilde{g}||_1$ and $||\tilde{g}||_2$.
Firstly we have

$$\begin{aligned}||\tilde{g}||_1 = \sum_{y\in\mathbb{F}_q^d}\tilde{g}(y) = \sum_{y\in\mathbb{F}_q^d} g(y)(h*S_t)(y) = \sum_{x,y\in\mathbb{F}_q^d} g(y)h(z)S_t(y-z)\\ = \frac{1}{q}||g||_1||h||_1 + D(g,h)\end{aligned}$$

Where as before $|D(g,h)| \le 2q^{\frac{d-1}{2}}||g||_2||h||_2$.
Next, we have

$$\begin{aligned}||\tilde{g}||_2^2 = \sum_{y\in\mathbb{F}_q^d}(\tilde{g}(y))^2 = \sum_{y\in\mathbb{F}_q^d} g^2(y)(h*S_t)^2(y)\\ \le ||g||_\infty^2||h*S_t||_2^2\\ \le ||g||_\infty^2||h||_2^2|S_t|^2\end{aligned}$$

In the last step the inequality follows from Young. In the notation of the inequality in the previous section, let $2 = r = s$, and $t = 1$. Then we have

$$\frac{1}{s}+\frac{1}{t}=1+\frac{1}{2}=\frac{1}{r}+1 \tag{18}$$

So Young gives

$$||h*S_t||_2 \leq ||h||_2||S_t||_1 = ||h||_2|S_t| \tag{19}$$

Hence we have

$$\tilde{g}||_2 \leq ||g||_\infty ||h||_2|S_t|. \tag{20}$$

——

In total this gives

$$\sum_{x,y,z\in\mathbb{F}_q^d} f(x)g(y)h(z)S_t(x-y)S_t(y-z) = \frac{1}{q^2}||f||_1||g||_1||h||_1 + D(f,g,h) \tag{21}$$

where

$$|D(f,g,h)| \leq 2q^{\frac{d-3}{2}}||f||_1||g||_2||h||_2 + 2q^{\frac{d-1}{2}}||f||_2||g||_\infty||h||_2|S_t| \tag{22}$$

We could also have done an equivalent calculation exchanging the roles of f and h. This would give

$$|D(f,g,h)| \leq 2q^{\frac{d-3}{2}}||f||_2||g||_2||h||_1 + 2q^{\frac{d-1}{2}}||f||_2||g||_\infty||h||_2|S_t|. \tag{23}$$

□

Next we wish to extend this method to prove a more general bound for the analogous sum for $k+1$ functions.

Theorem 3.4. *Let $f_1...f_{k+1}:\mathbb{F}_q^d \to \mathbb{R}^+$, $S_t=\{y\in\mathbb{F}_q^d|||y||=t\}$. Then*

$$\sum_{x_1...x_{k+1}\in\mathbb{F}_q^d} f_1(x_1)...f_{k+1}(x_{x+1})S_t(x_1-x_2)...S_t(x_k-x_{k+1}) = \frac{1}{q^k}||f_1||_1...||f_{k+1}||_1 + D(f_1...f_{k+1}) \tag{24}$$

where $D(f_1...f_{k+1})$ is an error term such that

$$|D(f_1...f_{k+1})| \leq \sum_{i=0}^{k-1} 2q^{\frac{d-(2i+1)}{2}} \left(\prod_{1\leq j\leq i}||f_j||_1\right)\left(\prod_{i+2\leq j\leq k}||f_j||_\infty\right)||f_{i+1}||_2||f_{k+1}||_2|S_t|^{k-1-i} \tag{25}$$

Proof. We proceed by induction. The base case, $k=1$, is the result proven in []. For the inductive step the method is similar to the previous theorem.

Suppose the result holds for all $n\leq k$. Define $\tilde{f}_k = f_k(x_k)(f_{k+1}*S_t)(x_k)$. Then we have

$$\sum_{x_1 \dots x_{k+1} \in \mathbb{F}_q^d} f_1(x_1) \dots f_{k+1}(x_{x+1}) S_t(x_1 - x_2) \dots S_t(x_k - x_{k+1})$$
$$= \sum_{x_1 \dots x_k \in \mathbb{F}_q^d} f_1(x_1) \dots f_{k-1}(x_{k-1}) \tilde{f}_k(x_k) S_t(x_1 - x_2) \dots S_t(x_{k-1} - x_k)$$
$$= q^{-k+1} ||f_1||_1 \dots ||f_{k-1}||_1 ||\tilde{f}_k||_1 + D(f_1 \dots \tilde{f}_k)$$

By work done in the previous proof, we know that

$$||\tilde{f}_k||_1 = \frac{1}{q} ||f_k||_1 ||f_{k+1}||_1 + D(f_k, f_{k+1}) \tag{26}$$

where

$$|D(f_k, f_{k+1})| \leq 2q^{\frac{d-1}{2}} ||f_k||_2 ||f_{k+1}||_2. \tag{27}$$

We also know by inductive hypothesis that

$$|D(f_1 \dots \tilde{f}_k)| \leq \sum_{i=0}^{k-2} 2q^{\frac{d-(2i+1)}{2}} \left(\prod_{1 \leq j \leq i} ||f_j||_1 \right) \left(\prod_{i+2 \leq j \leq k-1} ||f_j||_\infty \right) ||f_{i+1}||_2 ||\tilde{f}_k||_2 |S_t|^{k-2-i}.$$

Here we have $||\tilde{f}_k||_2 \leq ||f_k||_\infty ||f_{k+1}||_2 |S_t|$. Combining this with the above we get a main term of

$$\frac{1}{q^k} ||f_1||_1 \dots ||f_{k+1}||_1$$

plus an error term

$$|D(f_1 \dots f_{k+1})| \leq 2q^{\frac{d-2k+1}{2}} ||f_1||_1 \dots ||f_{k-1}||_1 ||f_k||_2 ||f_{k+1}||_2$$
$$+ \sum_{i=0}^{k-2} 2q^{\frac{d-(2i+1)}{2}} \left(\prod_{1 \leq j \leq i} ||f_j||_1 \right) \left(\prod_{i+2 \leq j \leq k-1} ||f_j||_\infty \right) ||f_{i+1}||_2 ||f_k||_\infty ||f_{k+1}||_2 |S_t|^{k-1-i}$$
$$= \sum_{i=0}^{k-1} 2q^{\frac{d-(2i+1)}{2}} \left(\prod_{1 \leq j \leq i} ||f_j||_1 \right) \left(\prod_{i+2 \leq j \leq k} ||f_j||_\infty \right) ||f_{i+1}||_2 ||f_{k+1}||_2 |S_t|^{k-1-i}$$

□

In both of the above theorems we do not expect this bound to be optimal, and it can most likely be improved. In particular, the step of calculating a bound for $||\tilde{g}||_2$ in theorem 2.3 involves the rather awkward step of pulling out a $||g||_\infty^2$ from the sum

$$\sum_{y \in \mathbb{F}_q^d} g^2(y)(h * S_t)^2(y).$$

If a better method for estimating $||\tilde{g}||_2$ can be found, it should improve the result.

References

[1] P. Erdös *On sets of distances of n points,* Amer. Math. Monthly. **53** (1946), 248–250.

[2] L. Guth, N.H.Katz, *On the Erdős distinct distance problem in the plane*, Annals of Mathematics 181 (2015), 155-190

[3] Falconer, K. J.: On the Hausdorff dimensions of distance sets. Mathematika, **32**, no.2, 206-212 (1985)

[4] J. Chapman, M. Erdoğan, D. Hart, A. Iosevich, D. Koh, *Pinned distance sets, k-simplices, Wolff's exponent in finite fields and sum-product estimates* (2009)

[5] M. Erdoğan, *A Bilinear Fourier Extension Theorem and Applications to the Distance Set Problem*, International Mathematics Research Notices, 23: 1411–1425

[6] A. Iosevich, M. Rudnev, *Erdős Distance Problem in Vector Spaces over Finite Fields*, Transactions of the American Mathematical Society, Volume 359, No 12, December 2007, pages 6127-6142

[7] D. Hart, A. Iosevich, D. Koh, M. Rudnev, *Averages over hyperplanes, sum-product theory in vector spaces over finite fields and the Erdos-Falconer distance conjecture*, 2007.

[8] A. Iosevich, D. Koh, T. Pham, *A new perspective on the distance problem over prime fields*, 2019

[9] M. Bennet, J. Chapman, D. Covert, D. Hart, A. Iosevich, J. Pakianathan, *Long Paths in the Distance Graph Over Large Subsets of Vector Spaces Over Finite Fields*, Journal of the Korean Mathematical Society, 53, No. 1, pages 115-126.

[10] H. Iwaniec, E. Kowalski, *Analytic Number Theory*, Colloqium Publications, 53, (2004)