AREAS OF TRIANGLES AND BECK'S THEOREM IN PLANES OVER FINITE FIELDS

ALEX IOSEVICH, MISHA RUDNEV, AND YUJIA ZHAI

Abstract. In this paper, we will show that if E is a subset of a planar point set over a finite field $\mathbb{F}_q$ (q not necessarily prime) and has cardinality $|E| > 64q\log_2 q$, then there are more than $\frac{q}{2}$ distinct areas of triangles sharing the same vertex. A finite field version of Beck's theorem is derived to prove the result. The theorem says that if $|E| > 64q\log_2 q$, then pairs of distinct points of E generate a positive proportion of lines in $\mathbb{F}_q^2$. Also, the theorem ensures the existence of a point $z \in E$, such that there are at least $\frac{q}{4}$ lines incident to z, each supporting at least $\frac{|E|}{2q}$ and fewer than $\frac{2|E|}{q}$ points of E, other than z. This will give us the result about the number of distinct triangle areas.

In addition, we prove that if E is a subset of $\mathbb{F}_q^2$ and has cardinality $|E| > q$, then there are at least $\frac{q-1}{2}$ distinct areas of triangles, which can be chosen to share the same base.

The results about the volumes of $k+1$ simplices in higher dimensional vector spaces over finite fields follow from the conclusions about distinct triangle areas in 2-dimensional vector spaces over finite fields.

Contents

1. Introduction

The basic question of Erdős combinatorics is to determine whether a sufficiently large discrete set determines a configuration of a given type. See, for example, [3], [2], [22] and the references contained therein for a description of related problems and their consequences. Perhaps the most celebrated of these is Szemerédi's theorem

2000 *Mathematics Subject Classification.* 68R05,11B75.

which says that a subset of the integers of positive density contains an arithmetic progression of any given length.

In the metric plane, interesting geometric configurations are congruent line segments or triangles. A famous question is the Erdős distance conjecture which asks, how many distances are determined by the set of n points in $\mathbb{R}^d$, $d \geq 2$. It was conjectured by Erdős that the number of distance in $\mathbb{R}^d$, $d \geq 2$, is at least $n^{\frac{2}{n}}$, up to logarithm factors. Guth and Katz ([11]) recently resolved the long-standing Erdős distance conjecture ([6]) in $\mathbb{R}^2$ by proving that a planar set E of n points determines $\Omega(\frac{n}{\log n})$ distinct distances, or, equivalently, distinct congruence classes of line segments.

While geometric combinatorics in Euclidean space has been studied for a long time, analogous problems in vector spaces over finite fields have also received much attention. For instance, the first two listed authors formulated the finite field version of Erdős distance problem, which says that any subset E of $\mathbb{F}_q^d$ with cardinality $\lesssim q^{\frac{d}{2}}$ points generates $\gtrsim |E|^{\frac{2}{d}}$ distances. (Here, $X \lesssim Y$ means $X \leqq CY$, for some constant C). They also proved that if the E is a Salem set with cardinality $\gtrsim q^{\frac{d}{2}}$, then the conjecture holds. For more details, see [27].

Another conjecture of Erdős, Purdy, and Strauss ([7]) suggested that a non-collinear set of n points in $\mathbb{R}^2$ determines at least $\lfloor \frac{n-1}{2} \rfloor$ distinct nonzero triangle areas, the most economical configuration being evenly distributed $\lceil \frac{n}{2} \rceil$ points on a line with evenly distributed $\lfloor \frac{n}{2} \rfloor$ on a parallel line. A linear lower bound was found by Burton and Purdy in 1979. But the conjecture was not resolved until 2008 by Pinchasi ([19]). The precise statement of the result is as follows.

Theorem 1. *The number of distinct areas of triangles determined by a non-collinear point set of n points in $\mathbb{R}^2$ is at least $\lfloor \frac{n-1}{2} \rfloor$. Indeed, the triangles yielding distinct areas can be chosen to share a common base.*

In this paper, we prove an analogous result in the setting of planar sets over finite fields, which says that a set with more than q points determines at least $\frac{q-1}{2}$ distinct triangle areas. And the triangles giving distinct areas can be chosen such that they share the same base. The definition of triangle areas will be justified later.

Another interesting question about triangles asks if a vertex of triangles is fixed, how many distinct areas are attained by a non-collinear set. Roche-Newton with the first two listed authors of this paper ([16]) proved that a non-collinear n-point set $E \subset \mathbb{R}^2$ generates $\Omega(\frac{n}{\log n})$ distinct triangle areas for triangles pinned at the origin. The precise statement is as follows.

Theorem 2. *There exists a universal c, such that a set of $n > \frac{1}{c}$ non-collinear points in $\mathbb{R}^2$ determines at least $c\frac{n}{logn}$ distinct areas of triangles with one vertex at the origin.*

In the case of vector spaces over finite fields, we obtain the following result. For any subset E of $\mathbb{F}_q^2$ with $\Omega(q \log q)$ points, there exist a point in E such that the number of areas of triangles pinned at that point is greater than $\frac{q}{2}$.

As will be seen in this paper, geometric problems in vector spaces over a finite field $\mathbb{F}_q$ can often be effectively resolved using Fourier analysis, provided that the point sets in question are sufficiently large. See e.g. [9], [12], [24], [14], and the references contained therein.

Fourier techniques have not proven to be particularly efficient for sufficiently small subsets of finite fields, where the fundamental tools come from arithmetic combinatorics, as is demonstrated in a pioneering work of Bourgain, Katz, and Tao ([4]). Unfortunately, quantitative results involving small sets are still far behind their Euclidean prototypes. For instance, the celebrated Beck's theorem ([1]), whose finite field version, Theorem 6 below, is developed in this paper, states that a set of n points in the Euclidean plane, with no more than cn collinear points, with some absolute c, determines $\Omega(n^2)$ distinct lines drawn through pairs of distinct points. The best currently known exponents, for small sets, are due to Helfgott and the second listed author ([15]), who proved that for any $(E = A \times A) \subset \mathbb{F}_q^2$, where q is a prime, and $|A| < \sqrt{q}$, [1] the number of distinct lines is $\Omega(n^{1+\frac{1}{267}})$. The power $\frac{1}{267}$, added to 1 in the latter estimate has since been improved a few times, implicit in the recent work of Jones ([17]), but would still remain very far from 1 as it is in Beck's theorem.

Now we define the volumes of $(d+1)$ simplices determined by subsets of $\mathbb{F}_q^d$. Then the areas of triangles are defined as the case when $d = 2$. Let $\mathbb{F}_q$ be the field with q elements and $\mathbb{F}_q^d$ be the d-dimensional vector space over this field. More precisely, let $(x^1, \dots, x^{d+1})$ denote a $(d+1)$-tuple of vectors from $\mathbb{F}_q^d$. Given a set $E \subseteq \mathbb{F}_q^d$, define

$$V_d(E) = \Big\{ |\det(x^1 - x^{d+1}, \dots, x^d - x^{d+1})| : x^j \in E \Big\} \setminus \{0\} \tag{1}$$

as the set of d-dimensional nonzero volumes, defined by $(d+1)$-simplices whose vertices are in E, as well as for some fixed $z \in E$, the set of pinned nonzero volumes

$$V_d^z(E) = \Big\{ |\det(x^1 - z, \dots, x^d - z)| : x^j \in E \Big\} \setminus \{0\}. \tag{2}$$

Note that if $x^j = (x_1^j, \dots, x_d^j)$, the subscripts referring to the coordinates relative to the standard basis in $\mathbb{F}_q^d$, an element of $V_d(E)$, generated by the $(d+1)$-tuple $(x^1, \dots, x^{d+1})$ equals the determinant of the following $d+1$ by $d+1$ matrix:

$$V_d(x^1, \dots, x^{d+1}) = \left| \det \begin{pmatrix} 1 & \dots & 1 \\ x_1^1 & \dots & x_1^{d+1} \\ \vdots & \ddots & \vdots \\ x_d^1 & \dots & x_d^{d+1} \end{pmatrix} \right|.$$

The above definition of the volumes of $(d+1)$ simplices in $\mathbb{F}_q^d$ is natural in the sense that it is invariant under the action of orthogonal matrices.

[1] Throughout the paper we use the $|\cdot|$ notation to denote the cardinality of a finite set.

1.1. **Statement of results:** The main result of this note is is precisely stated as follows.

Theorem 3. *Let $E \subset \mathbb{F}_q^2$. Then the following hold.*

i) Suppose that $|E| > q$. Then $|V_2(E)| \geq \frac{q-1}{2}$, and the triangles giving at least $\frac{q-1}{2}$ distinct areas can be chosen such that they share the same base.

ii) Suppose that $|E| \geq 64q\log_2(q)$, and $q \geq C$ for some absolute C. Then there exists $z \in E$ such that

$$|V_2^z(E)| > \frac{q}{2}.$$

Remark 4. The claim i) is a slightly weaker finite field version of Pinchasi's result in $\mathbb{R}^2$. Namely, we are off by a constant in the sense that we do not know an example of a set with $q+1$ points in $\mathbb{F}_q^2$, which would not generate all possible $q-1$ nonzero areas of triangles. The question of what is the minimum size of $E \subset \mathbb{F}_2^q$ to yield all possible areas in $\mathbb{F}_q \setminus \{0\}$ is open. It seems reasonable to conjecture that $|E| = q+1$ would be necessary and sufficient, for any q. The analog of the conjecture in higher dimensions would be that $|E| = q^{d-1}+1$, necessary and sufficient to yield all possible d-dimensional volumes. In this direction, Corollary 1 below shows that if $|E| \geq 2q^{d-1}$, then all the possible volumes are determined.

Remark 5. In the context of recent work on sufficiently large sets in $\mathbb{F}_q^2$, Theorem 3 is an improvement over the earlier results [12], [25], which established the threshold $|E| = \Omega(q^{\frac{3}{2}})$ in order for a general $E \subseteq \mathbb{F}_q^2$ to determine $\Omega(q)$ distinct areas of triangles.

The claim ii) of Theorem 3 follows from the following theorem, which can be regarded as a finite field version of the aforementioned theorem due to Beck ([1]), for sufficiently large point sets.

Theorem 6. *(Beck's theorem in $\mathbb{F}_q^2$). Suppose that $E \subset \mathbb{F}_q^2$ with $q \geq C$, for some absolute C, and*

$$|E| \geq 64q\log q.$$

Then pairs of distinct points of E generate at least $\frac{q^2}{8}$ distinct straight lines in $\mathbb{F}_q^2$. Moreover, there exists a point $z \in E$ and at least $\frac{q}{4}$ straight lines incident to z, each supporting at least $\frac{1}{2}\frac{|E|}{q}$ and fewer than $2\frac{|E|}{q}$ points of E, other than z.

Remark 7. It is well known that Fourier analysis yields nearly optimal estimates over finite fields for sufficiently large sets. For instance, Garaev ([9]) proves an optimal sum-product lower bound for $|A+A| + |A \cdot A|$, when $A \subseteq \mathbb{F}_q$ is such that $|A| > q^{\frac{2}{3}}$. The first quantitative estimate in this direction was proved by Hart, the first listed author of this paper and Solymosi in [13]. Similarly, in $\mathbb{F}_q^2$ the "threshold" for what can be regarded as a sufficiently large set $E = A \times A$ is usually $q^{\frac{4}{3}}$, and $q^{\frac{3}{2}}$ for general E. See, e.g., [12], [25], [14], and the references contained therein. Theorem

6, however, delivers an optimal (up to a constant) estimate for $|A| = \Omega\left(\sqrt{q \log_2 q}\right)$. Note that the aforementioned estimate $\Omega(n^{1+\frac{1}{267}})$ in [15] for the number of lines generated by $A \times A$ with $|A| = O(\sqrt{q})$ (valid only when q is a prime) is strikingly weaker. In the same vein, sum-product results of [9], [13], valid for sufficiently large sets are considerably stronger than what is known for small sets, where the best result so far in prime fields is due to the second listed author ([20]), generalised to $\mathbb{F}_q$ by Li and Roche-Newton ([18]).

Remark 8. The construction in Corollary 2.4 in [14] implies that if $|E| = o(q^{\frac{3}{2}})$, then there exists $z \in E$ such that $|V_2^z(E)| = o(q)$. In particular, this implies that part ii) of Theorem 3 cannot be strengthened to say that one gets a positive proportion of the areas from *any* fixed vertex, which would be somewhat analogous to the above-mentioned Euclidean result of [16]. We do not know whether the logarithmic term in the assumption for part ii) is necessary. It is definitely needed for our proof.

Remark 9. The forthcoming proof of Theorem 6 is based on Vinh's ([24]) finite field variant, quoted as Theorem 12 below, of the classical Szemerédi-Trotter ([21]) theorem on the number of incidences $I(E, L)$ of points in E and lines in L. Vinh's theorem becomes its equal in the strength of exponents only if the underlying sets E, L involved are rather large, that is if one takes $|E||L| \sim q^3$, which amounts to $|E| = \Omega(q^{\frac{3}{2}})$ in the interesting case when $|E| \sim |L|$. Indeed, this is the threshold when the first term in Vinh's incidence estimate (9) dominates, giving $I(E, L) = O((|E||L|)^{\frac{2}{3}})$ for the number of incidences, as it is the case in the principal term of the celebrated Euclidean Szemerédi-Trotter estimate. Beck's theorem, however, does not require the full strength of the Szemeredi-Trotter incidence theorem and would already follow if the fact k^3 in the denominator of the first term of (8) below is replaced by $k^{2+\epsilon}$ for some $\epsilon > 0$.

In the original paper ([1]), Beck had $\epsilon = \frac{1}{20}$, rather than $\epsilon = 1$, provided by the Szemerédi-Trotter theorem as in the estimate (8) below. In other words, Beck's theorem is weaker than the Szemerédi-Trotter theorem. This is precisely the reason why we can afford to use (9) and succeed in obtaining a much better threshold $|E| = \Omega(q \log_2 q)$ in Theorem 6 (rather than $|E| = \Omega(q^{\frac{3}{2}})$) getting a nearly sharp (up to the endpoint term $\log_2 q$) variant of the Beck theorem in the finite plane $\mathbb{F}_q^2$, as to the minimum size of a set E to generate $\Omega(q)$ distinct straight lines.

Theorem 3 can be easily boot-strapped to higher dimensions, since if a set determines a certain number of $(d-1)$-dimensional volumes when restricted to a $(d-1)$-dimensional hyperplane, and on top of that contains at least one point outside of this hyperplane, then it automatically determines at least that many d-dimensional volumes. As a consequence of our method, we obtain the following improvement of a result of Vinh ([26]) who proved that if $|E| \geq (d-1)q^{d-1}$, $d \geq 3$, then $V_d(E) = \mathbb{F}_q \setminus \{0\}$.

Corollary 1. *Let $E \subset \mathbb{F}_q^d$, $d \geq 3$.*

i) Suppose that $|E| > q^{d-1}$. Then $|V_d(E)| \geq \frac{q-1}{2}$.

ii) Suppose that $|E| \geq 2q^{d-1}$. *Then* $V_d(E) = \mathbb{F}_q \setminus \{0\}$.

2. Proof of part i) of Theorem 3

Proof. The core of the forthcoming proof of claim i) follows the lines of Lemmas 2.1 and 2.2 in [10], which in turn go back to the "statement about generic projection" in [4], Lemma 2.1.

Let us first show that any set $E \subseteq \mathbb{F}_q^2$, with $|E| > q$, determines all possible directions. More precisely, every linear subspace $L \subset \mathbb{F}_q^2$ contains a nonzero element of $E - E$. This is a finite field analogue of the well-known result of Ungar ([23]) that $2N$ non-collinear points in the Euclidean plane determine at least $2N$ distinct directions.

Let L be a one-dimensional linear subspace of $\mathbb{F}_q^2$. Consider the sum set

$$S = E + L = \{s = e + l : e \in E, l \in L\}.$$

Since $|L||E| > |\mathbb{F}_q^2| = q^2$, there is an element $s \in S$ with more than one representation as a sum. More precisely,

$$s = e_1 + l_1 = e_2 + l_2, \qquad (e_1, e_2, l_1, l_2) \in E \times E \times L \times L, \qquad l_1 \neq l_2. \tag{3}$$

Hence

$$l_2 - l_1 = e_1 - e_2, \tag{4}$$

which implies that E determines all possible directions in the sense described above.

Now average the number of solutions of the equation (4), with $(e_1, e_2, l_1, l_2) \in E \times E \times L \times L$, over the $\frac{q^2-1}{q-1} = q + 1$ subspaces L. For each L, the pair (e_1, e_2), $e_1 \neq e_2$ in (4) determines the subspace L. Moreover, each $l \in L$ can be represented as a difference $l_2 - l_1$ in exactly q different ways. Including the trivial solutions where $e_1 = e_2$, we have

$$\begin{aligned} &|\{(e_1, e_2, l_1, l_2) \in E \times E \times L \times L : \text{ (4) holds for some } L\}| \\ &\qquad = |E|q(q+1) + |E|^2 q \\ &\qquad \leq 2q|E|^2. \end{aligned}$$

It follows that there exists a subspace L, such that

$$|\{(e_1, e_2, l_1, l_2) \in E \times E \times L \times L\} : \text{ (4) holds}\}| \leq 2|E|^2 \frac{q}{q+1}. \tag{5}$$

It follows, by the Cauchy-Schwartz inequality, that with this particular L,

$$\begin{aligned} |E + L| &\geq \frac{|E|^2|L|^2}{|\{(e_1, e_2, l_1, l_2) \in E \times E \times L \times L\} : \text{ (4) holds}\}|} \\ &\geq \frac{q(q+1)}{2}. \end{aligned} \tag{6}$$

Indeed, if $s \in E+L$, and $\nu(s)$ is the number of realisations of s as a sum, then

$$\sum_s \nu(s) = |E||L|, \qquad \sum_s \nu^2(s) = |\{(e_1,e_2,l_1,l_2) \in E\times E\times L\times L\} : \ (4) \text{ holds}\}|,$$

and by the Cauchy-Schwartz inequality

$$|S = E+L| \geq \frac{\left(\sum_{s\in S}\nu(s)\right)^2}{\sum_{s\in S}\nu^2(s)}.$$

We conclude from (6) that there are points of E in at least $\frac{q+1}{2}$ different parallel lines. Moreover one of these lines, as has been shown in the beginning of the proof, has at least two distinct points $e_1, e_2 \in E$. It follows that E determines at least $\frac{q-1}{2}$ distinct triangles, with the same base e_1e_2 and distinct nonzero heights, that is the third vertex of the triangle lying on different lines parallel to the base. Thus, there are at least $\frac{q-1}{2}$ distinct nonzero triangle areas.

□

3. Proof of part ii) of Theorem 3

3.1. **Fourier mechanism.** We shall need the following Fourier-analytic result, which is an easy variant of the corresponding estimate from [12] and [14].

Let χ be a non-trivial additive character over $\mathbb{F}_q$ and let $F(x)$ for the characteristic function of a set $F \subseteq \mathbb{F}_q^2$. Define the Fourier transform $\widehat{F}$ of $F(x)$ as

$$\widehat{F}(\xi) = \frac{1}{q^2}\sum_x F(x)\chi(-\xi\cdot x).$$

The formulas we shall need are the following:

$$\sum_{s\in\mathbb{F}_q}\chi(-sx) = q \quad (orthogonality)$$

if $x=0$ and 0 otherwise.

$$\sum_m \widehat{f}(m)\overline{\widehat{g}(m)} = q^{-d}\sum_x f(x)g(x) \quad (Plancherel)$$

Theorem 10. *Let $F, G \subset \mathbb{F}_q^2$. Suppose $0 \notin F$. Let, for $t \in \mathbb{F}_q$,*

$$\nu(t) = |\{(x,y) \in F\times G : x\cdot y = t\}|,$$

where $x\cdot y = x_1y_1 + x_2y_2$.

Then

$$\sum_t \nu^2(t) \leq |F|^2|G|^2q^{-1} + q|F||G|\cdot\max_{x\neq 0}|F\cap l_x|, \tag{7}$$

where

$$l_x = \{sx : s \in \mathbb{F}_q \setminus \{0\}\},$$

with $x \in \mathbb{F}_q^2 \setminus \{0\}$.

Proof. To prove the theorem, observe that for any $t \in \mathbb{F}_q$, one has, by the Cauchy-Schwartz inequality:

$$\begin{aligned}\nu^2(t) &= \left(\sum_{x\cdot y=t} F(x)G(y)\right)^2\\ &\leq \sum_{x\in F} 1^2 \left(\sum_{x\in F}\sum_{x\cdot y=t} F(x)G(y)\right)^2\\ &\leq |F| \cdot \sum_{x\cdot y=x\cdot y'=t} F(x)G(y)G(y').\end{aligned}$$

By orthogonality, one has

$$\begin{aligned}\sum_t \nu^2(t) &\leq |F| \cdot \sum_{x\cdot y=x\cdot y'} F(x)G(y)G(y')\\ &= q^{-1}|F| \sum_{x,y,y'} F(x)G(y)G(y') \sum_{s\in\mathbb{F}_q} \chi(sx\cdot(y-y'))\\ &= |F|^2|G|^2q^{-1} + q^{-1}|F| \sum_{s\neq 0}\sum_{x,y,y'} \chi(sx\cdot(y-y'))F(x)G(y)G(y')\\ &= |F|^2|G|^2q^{-1} + q^{-1}|F| \sum_{s\neq 0}\sum_{x} F(x) \sum_{y,y'} \big(\chi(-sx\cdot y')G(y')\big)\big(\chi(sx\cdot y)G(y)\big)\\ &= |F|^2|G|^2q^{-1} + q^{-1}|F| \sum_{s\neq 0}\sum_{x} F(x) \left|\sum_{y} \chi(-sx\cdot y)G(y)\right|^2\\ &= |F|^2|G|^2q^{-1} + q^{3}|F| \sum_{s\neq 0}\sum_{x} \left|\widehat{G}(sx)\right|^2 F(x)\end{aligned}$$

by definition of the Fourier transform. Then by the assumption that $0 \notin F$ and by change of variables, one has

$$\begin{aligned}&|F|^2|G|^2q^{-1} + q^{3}|F| \sum_{s\neq 0}\sum_{x\neq 0} \left|\widehat{G}(x)\right|^2 F(sx)\\ &= |F|^2|G|^2q^{-1} + q^{3}|F| \sum_{x\neq 0} |\widehat{G}(x)||F\cap l_x|\\ &\leq |F|^2|G|^2q^{-1} + q^{3}|F| \max_{x\neq 0}|F\cap l_x| \sum_{x\neq 0} |\widehat{G}(x)|\\ &= |F|^2|G|^2q^{-1} + q|F||G| \cdot \max_{x\neq 0}|F\cap l_x|,\end{aligned}$$

where the last step uses the Plancherel identity. This completes the proof of Theorem 10. □

3.2. **Finite field variant of the Beck theorem.** Here we prove Theorem 6, a variant of the Euclidean theorem due to Beck ([1]) in the finite field context. Recall that Beck's theorem says that either a positive proportion of a set of n points in the Euclidean plane lie on single line, or there exists a constant multiple of n^2 distinct lines each containing at least two points of E. Note that Beck's theorem follows easily from the following formulation of the celebrated Szemerédi-Trotter incidence theorem ([21]).

Theorem 11. *Let E be a collection of points in $\mathbb{R}^2$, and $L(E)$ the set of lines determined by distinct pairs of points of E. For $k \geq 2$, let $L_k \subseteq L(E)$ denote the lines supporting at least k points of E. Then there exists $C > 0$ such that*

$$|L_k| \leq C \left(\frac{|E|^2}{k^3} + \frac{|E|}{k} \right). \tag{8}$$

We shall need the following finite field variant of the Szemerédi-Trotter theorem due to Vinh ([24]).

Theorem 12. *Let E be a collection of points and L a collection of lines in $\mathbb{F}_q^2$. Then*

$$I(E,L) = |\{(e,l) \in E \times L : e \in l\}| \leq |E| \cdot |L| \cdot q^{-1} + \sqrt{q \cdot |E| \cdot |L|}. \tag{9}$$

Using L_k in place of L in Theorem 12, we see that $I(E, L_k) \geq k|L_k|$. It then follows directly from (9) that if $k > \frac{|E|}{q}$, then

$$k|L_k| \leq q^{-1}|E||L_k| + \sqrt{q|E||L_k|},$$

which implies that

$$(k - q^{-1}|E|)|L_k| \leq \sqrt{q|E||L_k|},$$

then if $k > \frac{|E|}{q}$, one has

$$|L_k| \leq \frac{q|E|}{(k - q^{-1}|E|)^2}. \tag{10}$$

This leads us to the proof of Theorem 6.

Proof. Here q is treated as asymptotic parameter, to which the $o(1)$ notation relates. In order to avoid messy notation, let us assume, for convenience, that the quantities $q^{-1}|E|$, as well as $\log_2 q$ are integers. The reader shall see that we have more than enough flexibility with the constants to make this work. The key to our proof is the following assertion.

Lemma 1. *At least $\frac{|E|^2}{4}$ unordered pairs of distinct points of E are supported on the subset $\mathcal{L} \subseteq L(E)$, defined as the set of all lines in $L(E)$ containing between $1 + \frac{|E|}{2q}$ and $2\frac{|E|}{q}$ points of E.*

Proof. With a slight abuse of the notation, let L^j be the set of lines from $L(E)$, supporting no fewer than $q^{-1}|E| \cdot 2^j$ and no more than $q^{-1}|E| \cdot 2^{j+1}$ points of E,

where j is an integer ranging between 1 and $\log_2 q$. The upper bound of j is obtained in the following way.

$$q^{-1}|E|2^j \leq |E|,$$

which means

$$2^j \leq q.$$

Hence

$$j \leq \log_2 q.$$

The maximum number of points on a line from L^j is $q^{-1}|E| \cdot 2^{j+1}$, so the number of unordered pairs of distinct points is less than or equal to

$$\frac{q^{-1}|E|2^{j+1}(q^{-1}|E|2^{j+1}-1)}{2} \leq q^{-2}|E|^2 \cdot 2^{2j+1}.$$

On the other hand, by (10),

$$|L^j| \leq \frac{q|E|}{q^{-2}|E|^2 \cdot (2^j-1)^2}$$

Hence, the number of unordered pairs of distinct points of E supported on the union of L^j, for $j \geq 2$ is bounded by

$$\begin{aligned} & q^{-2}|E|^2 \cdot 2^{2j+1} \frac{q|E|}{q^{-2}|E|^2 \cdot (2^j-1)^2} \\ & = q|E| \frac{2^{2j+1}}{(2^j-1)^2} \\ & \leq 4|E|q, \text{ for } j \geq 2, \end{aligned}$$

and by $8|E|q$ for $j = 1$. Then the number of unordered pairs of distinct points of E supported on the union of L^j, for $j \geq 1$ is bounded by

$$\begin{aligned} & 8|E|q + \sum_{j=2}^{\log_2 q} 4|E|q \\ & = 4|E|q(\log_2 q - 1) + 8|E|q \\ & = 4|E|q(1+\log_2 q). \end{aligned}$$

Then, if $|E| \geq 64q \log_2 q$, and q is large enough, the latter bound constitutes only a small proportion of the total number $\frac{|E|(|E|-1)}{2}$ of unordered pairs of distinct points of E. More precisely, it follows, with the above choice of constants (for q large enough and $q = o(|E|)$) that the number of unordered pairs of distinct points of E,

supported on lines in $L(E)$, each supporting at most $2q^{-1}|E|$ points, is at least

$$\begin{aligned}
&\frac{|E|(|E|-1)}{2} - (4|E|q + 4|E|\log_2 q) \\
&\geq \frac{|E|(|E|-1)}{2} - 4|E|q\frac{|E|}{64q} \\
&\geq \frac{7}{16}|E|^2 - \frac{|E|}{2} - 4|E|q \\
&\geq \frac{7}{16}|E|^2 - \frac{|E|}{2} - 4|E|\frac{|E|}{64\log_2 q} \\
&\geq \frac{5|E|^2}{12}.
\end{aligned}$$

There are only $q(q+1)$ distinct lines in $\mathbb{F}_q^2$, and each line from $L(E)$ with fewer than $1 + \frac{|E|}{2q}$ points. So under the assumption that q large enough and $q = o(|E|)$, the number of points support on $L(E)$ is fewer than

$$\begin{aligned}
q(q+1)(1+\frac{|E|}{2q}) &= q^2 + q + \frac{q|E|}{2} + \frac{|E|}{2} \\
&\leq \frac{|E|^2}{6}.
\end{aligned}$$

This completes the proof of Lemma 1. □

To complete the proof of Theorem 6, let $l \in \mathcal{L}$, the set $\mathcal{L}$ provided by Lemma 1, and $\nu(l)$ is the number of points of E on l. It follows from Lemma 1 that

$$\sum_{l\in\mathcal{L}} \nu^2(l) > 2\sum_{l\in\mathcal{L}} \frac{\nu(l)(\nu(l)-1)}{2} \geq \frac{|E|^2}{2}.$$

Dividing this by the maximum value of $\nu(l) \leq 2q^{-1}|E|$, we obtain for the total number of incidences

$$I(E,\mathcal{L}) = \sum_{l\in\mathcal{L}} \nu(l) \geq \frac{q|E|}{4}. \tag{11}$$

By the pigeonhole principle, there exists a point $z \in E$ with at least $\frac{q}{4}$ lines of $\mathcal{L}$ passing through it. Now dividing (11) by $2q^{-1}|E|$ once more yields the desired bound

$$|\mathcal{L}| \geq \frac{q^2}{8}.$$

This completes the proof of Theorem 6. □

Remark 13. The lemma aims at partitioning the lines in the plane such that (i) each line in the partitioned family supports $\gtrsim \frac{|E|}{q}$ points of E; (ii) there are enough lines satisfying (i). As can be seen in the proof of Theorem 6, both (i) and (ii) are needed.

3.3. **Combining Fourier (section 3.1) and Beck (section 3.2) estimates.** We now prove the claim ii) of Theorem 3.

Proof. Let us refine the initial set E to the set E', containing z, provided by Theorem 6, and exactly $\frac{1}{2}q^{-1}|E|$ points of E other than z on exactly $\frac{q}{4}$ lines incident to z. The flexibility in the choice of constants in the proof of Theorem 6 enables one to treat $\frac{q}{4}$ as integer. It follows that

$$|E'| > \frac{|E|}{8} \geq 8q \log_2 q. \tag{12}$$

Now place z to the origin: let $E'_z = E' - z$ and apply Theorem 10. More precisely, in the application of the Lemma let $G = {E'_z}^{\perp}$ and $F = E'_z \setminus \{0\}$. Hence, we have

$$|E'_z| = |E'| - 1 = \frac{|E|}{8},$$

$$|F| = |E'_z| - 1 = |E| - 2,$$

$$|G| = |E'_z| = |E'| - 1.$$

Apply the estimate (7) to the sets F and G. In view of (12) and by Lemma 1, the quantity $|F \cap l_x|$, for any $x \neq 0$, in the second term of the estimate (7) is bounded by $\frac{4|F|}{q}$. Thus, once more by (12), the second term in the estimate (7) is dominated by the first one. More precisely,

$$q|F||G|\frac{4|F|}{q} = 4|F|^2|G| \ll \frac{|F|^2|G|^2}{q}$$

if and only if

$$4q \ll |G|$$

if and only if

$$4q \ll \frac{|E|}{8}.$$

Indeed, it follows by (12) and the construction of $|E'_z|$ that the number of pairs $(x, y) \in E'_z \times E'_z$, such that x, y lie on the same line through the origin, which is $O(q^{-1}|E|)$, is $o(q^{-2}|E'_z|^2)$. Then the estimate (7) yields

$$\sum_t \nu^2(t) \leq \frac{|E'|^4}{q}(1 + o(1)). \tag{13}$$

Now the claim ii) of Theorem (3) follows by the Cauchy-Schwarz inequality in the following way. We have

$$\begin{aligned}
|V_2^z(E)| &\geq |V_2^z(E')| \\
&= |V_2^z(E'_z)| = |\{t : \nu(t) \neq 0\}| \\
&\geq \frac{|\#(x,y) \in F \times G : x \cdot y \neq 0\}|^2}{\sum_{t \neq 0} \nu^2(t)} \\
&\geq \frac{|\#(x,y) \in F \times G : x \cdot y \neq 0\}|^2}{\sum_t \nu^2(t)} \\
&\geq \frac{|F|^2(|G| - \frac{|E|}{2q})^2}{\sum_t \nu^2(t)} \\
&= \frac{(|E'| - 2)^2(|E'| - 1 - \frac{|E|}{2q})^2}{\sum_t \nu^2(t)} \\
&\geq \frac{|E'|^4(1 - o(1))}{q^{-1}|E'|^4(1 + o(1))} \\
&> \frac{q}{2}.
\end{aligned}$$

This completes the proof of part ii) of Theorem 3.

□

4. Proof of Corollary 1

Proof. We proceed by induction on the dimension. Suppose that part i) holds for the dimension $d-1$, $d \geq 3$ and part ii) holds for dimension $d-1$ with $d \geq 4$. The base of the induction for part i) is Theorem 3, and for part ii) it is the result of Vinh ([26]) that $V_3(E) = \mathbb{F}_p \setminus \{0\}$ for $E \subseteq \mathbb{F}_q^3$ with $|E| \geq 2q^2$.

Consider now the intersections of $E \subseteq \mathbb{F}_q^d$ with hyperplanes $H_c = \{x : x_d = c\}$. By the pigeonhole principle, there exists c such that $|E \cap H_c| > q^{d-2}$ for part i), and $|E \cap H_c| \geq 2q^{d-2}$ for part ii) of Corollary 1.

Since $V_d(E)$ is invariant under translations, we may assume that $c = 0$. In addition, since $|E \cap H_c| > q^{d-1}$, there must be a point $z \in E \backslash H_0$, which means that $z_d \neq 0$.

By the induction assumption the set $V_{d-1}(E \cap H_0)$ satisfies the conclusion of Corollary 1. It follows that

$$|V_d(E)| \geq |V_d^z((E \cap H_0) \cup \{z\})| = |V_{d-1}(E \cap H_0)|.$$

To verify the latter equality, since $x_d^j = 0$ for every $x^j \in E \cap H_0$, the elements of $V_d^z(E)$ are determinants of size $d+1$ of the form

$$\left|\det\begin{pmatrix} 1 & \dots & 1 & 1 \\ x_1^1 & \dots & x_1^d & z_1 \\ \vdots & \ddots & \ddots & \vdots \\ x_{d-1}^1 & \dots & x_{d-1}^d & z_{d-1} \\ 0 & \dots & 0 & z_d \end{pmatrix}\right| = \left|z_d \cdot \det\begin{pmatrix} 1 & \dots & 1 \\ x_1^1 & \dots & x_1^d \\ \vdots & \ddots & \vdots \\ x_{d-1}^1 & \dots & x_{d-1}^d \end{pmatrix}\right|.$$

This completes the proof of Corollary 1. □

References

[1] J. Beck. *On the lattice property of the plane and some problems of Dirac, Motzkin, and Erdős in combinatorial geometry.* Combinatorica **3** (1983), 281–297.

[2] P. Brass, W.O.J. Moser, J. Pach. *Research Problems in Discrete Geometry.* Springer Verlag (2005), 499pp.

[3] J. Pach, P. Agarwal. *Combinatorial geometry.* Wiley-Interscience Series in Discrete Mathematics and Optimization. A Wiley-Interscience Publication. John Wiley and Sons, Inc., New York (1995), 376pp.

[4] J. Bourgain, N. Katz, T. Tao. *A sum-product estimate in finite fields, and applications.* Geom. Funct. Anal. **14** (2004), 27–57.

[5] D. Covert, D. Hart, A. Iosevich, D. Koh, M. Rudnev. *Generalized incidence theorems, homogeneous forms and sum-product estimates in finite fields.* European J. of Combinatorics **31** (2010), 306–319.

[6] P. Erdős. *On sets of distances of n points.* American Mathematical Monthly **53** (1946), 248-?250.

[7] P. Erdőos, G. Purdy, E. G. Strauss. *On a problem in combinatorial geometry.* Discrete Mathematics **40** (1982), 45-52.

[8] P. Erdős, E. Szemerédi. *On sums and products of integers.* Studies in Pure Math. Birkhäuser, Basel (1983), 213–218.

[9] M.Z. Garaev. *The sum-product estimate for large subsets of prime fields.* Proc. Amer. Math. Soc. **136** (2008), no. 8, 2735-2739.

[10] A. Glibichuk and M. Rudnev. *On additive properties of product sets in an arbitrary finite field.* J. Anal. Math. **108** (2009), 159-170.

[11] L. Guth, N.H. Katz. *On the Erdös distinct distance problem in the plane.* Preprint arXiv:math/1011.4105 (2010), 37pp.

[12] D. Hart and A. Iosevich. *Sums and products in finite fields: an integral geometric viewpoint.* Radon transforms, geometry, and wavelets, pp. 129–135. Contemp. Math., **464**, Amer. Math. Soc., Providence, RI (2008).

[13] D. Hart, A. Iosevich, J. Solymosi. *Sum-product estimates in finite fields via Kloosterman sums.* Int. Math. Res. Not. no. 5 (2007), art. ID rnm007, 14 pp.

[14] D. Hart, A. Iosevich, D. Koh, M. Rudnev. *Averages over hyperplanes, sum-product theory in vector spaces over finite fields and the Erdős-Falconer distance conjecture.* Trans. Amer. Math. Soc. **363** (2011), no. 6, 3255–3275.

[15] H.A. Helfgott, M. Rudnev. *An explicit incidence theorem in F_p.* Mathematika **57** (2011), no. 1, 135-145.

[16] A. Iosevich, O. Roche-Newton, M. Rudnev. *On an application of Guth-Katz theorem.* Math. Res. Lett. **18** (2011), no. 4, 691-697.

[17] T.G.F. Jones. *An improved incidence bound over fields of prime order.* Preprint arXiv:math/1110.4752 (2012), 10pp.

[18] L. Li, O. Roche-Newton. *An improved sum-product estimate for general finite fields.* SIAM J. Discrete Math. **25** (2011), no. 3, 1285-1296.

[19] R. Pinchasi. *The minimum number of distinct areas of triangles determined by a set of n points in the plane.* SIAM J. Discrete Math. **22** (2008), no. 2, 828-831.

[20] M. Rudnev. *An Improved Sum-Product Inequality in Fields of Prime Order.* Int. Math. Res. Notices (2011) doi: 10.1093/imrn/rnr158.

[21] E. Szemerédi, W. T. Trotter. *Extremal problems in discrete geometry.* Combinatorica **3**, (1983) 381–392.

[22] T. Tao, V. Vu. *Additive combinatorics.* Cambridge Studies in Advanced Mathematics, **105**. Cambridge University Press, Cambridge (2006), 512pp.

[23] P. Ungar. $2N$ *noncollinear points determine at least* $2N$ *directions.* J. Combin. Theory Ser. A **33** (1982), no. 3, 343–347.

[24] L.A. Vinh. *The Szemerédi-Trotter type theorem and the sum-product estimate in finite fields.* European J. Combin. **32** (2011), no. 8, 1177–1181.
[25] L.A. Vinh. *Distinct triangle areas in a planar point set over finite fields.* Electron. J. Combin. **18** (2011), no. 1, Paper 213, 6 pp.
[26] L.A. Vinh. *On the volume set of point sets in vector spaces over finite fields.* arXiv:math/0903.2510 (2009), 6pp.
[27] A.Iosevich, M.Rudnev. *Erdős distance problem in vector spaces over finite fields .* arXiv:math/0509005(2008), 15pp.

Alex Iosevich, Department of Mathematics, University of Rochester, Rochester, NY
E-mail address: iosevich@math.rochester.edu

Misha Rudnev, Department of Mathematics, University of Bristol, Bristol BS8 1TW, United Kingdom
E-mail address: m.rudnev@bristol.ac.uk

Yujia Zhai, Department of Mathematics, University of Rochester, Rochester, NY
E-mail address: yzhai@u.rochester.edu