Collinear triples and quadruples in Cartesian products in $\mathbb{F}_p^2$

Giorgis Petridis

Abstract

In this informal note, which has been absorbed in [6], we combine a recent point-line incidence bound of Stevens and de Zeeuw with an older lemma of Bourgain, Katz and Tao to bound the number of collinear triples and quadruples in a Cartesian product in $\mathbb{F}_p^2$.

1 Introduction

Let $A \subseteq \mathbb{F}_p$ be a set in a prime order finite field of odd characteristic. In this note, which surpasses the result from [7], we combine a recent point-line incidence bound of Stevens and de Zeeuw [9] with an older lemma of Bourgain, Katz and Tao [2] to obtain an improved bound for $T(A)$, the number of collinear triples in $A \times A$, for large sets; and an optimal bound for $Q(A)$, the number of collinear quadruples in $A \times A$.

By a collinear triple we mean an ordered triple $(u, v, w) \in (A \times A) \times (A \times A) \times (A \times A)$ such that $u$, $v$ and $w$ are all incident to the same line. So, for example, any point $u \in A \times A$ gives rise to the collinear triple $(u, u, u)$. Collinear quadruples are defined similarly.

The quantities $T(A)$ and $Q(A)$ can be expressed in terms of the incidence function associated with $A \times A$. For a line $\ell \subset \mathbb{F}_p^2$, $i(\ell)$ equals the number of points in $A \times A$ incident to $\ell$. Then

$$T(A) = \sum_{\text{all lines } \ell} i(\ell)^3 \quad \text{and} \quad Q(A) = \sum_{\text{all lines } \ell} i(\ell)^4.$$
Let us study in some detail what is known for collinear triples. The contribution to collinear triples coming from the $|A|$ horizontal and the $|A|$ vertical lines incident to $|A|$ points in $A \times A$ is $2|A|^4$. All other collinear triples can be counted by the number of solutions to

$$\frac{a_1 - a_2}{a_3 - a_4} = \frac{a_1 - a_5}{a_3 - a_6} \neq 0, \ a_i \in A, \ a_3 - a_4, a_3 - a_6 \neq 0.$$  \hspace{1cm} (1)

It follows from this that the expected number of collinear triples of a random set (where elements of $\mathbb{F}_p$ belong to $A$ independently with probability $|A|/p$) is $|A|^6 p^{-1} + 2|A|^4$. This is because for each 5-tuple $(a_1, \ldots, a_5)$ there is a unique element $a_6 \in \mathbb{F}_p$ that satisfies (1) and it belongs to $|A|$ with probability $|A|/p$.

Another interesting example is that of sufficiently small arithmetic progressions. First note that in general $T(A)$ equals the number of solutions to

$$(a_1 - a_2)(a_3 - a_4) = (a_1 - a_5)(a_3 - a_6) a_i \in A.$$ 

plus $O(|A|^4)$. So it equals

$$\sum_{a_1, a_3 \in A} \sum_x f_{a_1, a_3}^2(x) + O(|A|^4),$$

where $f_{a_1, a_3}(x)$ is the number of ways one can express $x$ as a product $(a_1 - a_2)(a_3 - a_4)$ with $a_2, a_4 \in A$.

Observe that for all $a_1, a_3 \in A$, $\sum_x f_{a_1, a_3}(x) = |A|^2$ because each pair $(a_2, a_4) \in A \times A$ contributes 1 to the sum.

Applying the Cauchy-Schwartz inequality gives

$$T(A) = \sum_{a_1, a_3 \in A} \sum_x f_{a_1, a_3}^2(x) \geq \sum_{a_1, a_3 \in A} \frac{(\sum_x f_{a_1, a_3}(x))^2}{|\text{supp}(f_{a_1, a_3})|} = \sum_{a_1, a_3 \in A} \frac{|A|^4}{|\text{supp}(f_{a_1, a_3})|^4}.$$ 

Now take $A = \{1, \ldots, \sqrt{p}/2\} \subset \mathbb{Z}$. Then for all $a_i \in A$ the product $(a_1 - a_2)(a_3 - a_4)$ is $\{1, \ldots, p\}$. This means that the support of $f_{a_1, a_3}$ is the same whether $A$ is taken to be a subset of $\mathbb{Z}$ or of $\mathbb{F}_p$. Ford has shown in [4] that the support of $f_{a_1, a_3}$ (in $\mathbb{Z}$ and hence) in $\mathbb{F}_p$ is $O(|A|^2/\log(|A|)^\gamma)$ for some absolute constant $\gamma < 1$. Substituting above implies that for $A = \{1, \ldots, \sqrt{p}/2\} \subset \mathbb{F}_p$, we have $T(A) = \Omega(|A|^4 \log(|A|)^\gamma)$. 

2
Because of these two examples, it is natural to expect that the bound

$$T(A) = O \left( \frac{|A|^6}{p} + \log(|A|)|A|^4 \right)$$

is correct up to perhaps logarithmic factors. Over the reals, Elekes and Ruzsa observed in [3] that the above inequality follows from the Szemerédi-Trotter point-line incidence theorem [10].

Far less is currently known. As is explained in Section 3, it is straightforward to obtain

$$\left| T(A) - \left( \frac{|A|^6}{p} + 2|A|^4 \right) \right| \leq p|A|^2.$$

It follows that if $|A| = \Omega(p^{2/3})$, then $T(A) = \Theta(|A|^6/p)$.

In the range $|A| = O(p^{2/3})$, Aksoy Yazici, Murphy, Rudnev and Shkredov [1, Proposition 5], building on Rudnev’s breakthrough result in [8], established the bound

$$T(A) = O(|A|^{9/2}).$$

While very strong, the result of Aksoy Yazici, Murphy, Rudnev and Shkredov does not improve the range of $|A|$ where $T(A) = O(|A|^6/p)$.

Combining the two bounds above gives

$$T(A) = O \left( \frac{|A|^6}{p} + |A|^{9/2} \right).$$

Similar results are true for collinear quadruples. The expected number of collinear quadruples is $\frac{|A|^8}{p^2} + 2|A|^5$. The result of Aksoy Yazici, Murphy, Rudnev and Shkredov implies that $Q(A) = O(|A|^{11/2})$ when $|A| = O(p^{2/3})$. One expects that the correct order of magnitude up to logarithmic factors is

$$Q(A) = O \left( \frac{|A|^8}{p^2} + \log(|A|)|A|^5 \right).$$

Once again large random sets and small arithmetic progressions offering (nearly) extremal examples.

We offer an improvement on the know bound for $T(A)$ when $|A| = \Omega(p^{1/2})$ and establish a nearly best possible bound for $Q(A)$. 
Theorem 1. Let $A \subseteq \mathbb{F}_p$.

1. The number of collinear triples in $A \times A$ satisfies

$$T(A) = O\left(\frac{|A|^6}{p} + p^{1/2}|A|^{7/2}\right).$$

So there is at most a constant multiple of the expected number of collinear triples when $|A| = \Omega(p^{3/5})$.

2. The number of collinear quadruples in $A \times A$ satisfies

$$Q(A) = O\left(\frac{|A|^8}{p^2} + \log(|A|)|A|^5\right),$$

which is optimal up to perhaps logarithmic factors.

The proof of the theorem is based on a recent point-line incidence bound for Cartesian products proved by Stevens and de Zeeuw [9, Theorem 4] and a lemma of Bourgain, Katz and Tao [2, Lemma 2.1]. A more precise version of Theorem 1 and applications to sum-product questions in $\mathbb{F}_p$ are given in [6].

Notation. We use Landau’s notation so that both statements $f = O(g)$ and $g = \Omega(f)$ mean there exists an absolute constant $C$ such that $f \leq Cg$ and $f = \Theta(g)$ stands for $f = O(g)$ and $f = \Omega(g)$. The letter $p$ denotes an odd prime, $\mathbb{F}_p$ the finite field with $p$ elements and $\mathbb{F}_q^2$ the 2-dimensional vector space over $\mathbb{F}_p$. For a line $\ell$, $i(\ell)$ represents the number of points in $A \times A$ incident to $\ell$.

2 The two ingredients

Let us state the two main ingredients of the proof of Theorem 1 and some straightforward consequences. We begin with the theorem of Stevens and de Zeeuw.

Theorem 2 (Stevens and de Zeeuw). Let $A \subseteq \mathbb{F}_p$ and $L$ be a collection of lines in $\mathbb{F}_p^2$. Suppose that $|A||L| = O(p^2)$. The number of point-line incidences between $A \times A$ and $L$ satisfies $I(A \times A, L) = O(|L|^{3/4}|A|^{5/4})$.

Sevens and de Zeeuw very reasonably imposed the additional condition $|A| < |L| < |A|^3$ because in other ranges, the Cauchy-Schwartz point-line incidence bound is better to theirs. For simplicity of argument, we omit the condition. A sanity check that allowing $|L| \leq |A|$ of $|L| \geq |A|^3$ does not hurt.
1. If $|L| \leq |A|$, we have $\mathcal{I}(A \times A, L) \leq |L||A| \leq |L|^{3/4}|A|^{5/4}$.

2. If $|A|^3 \leq |L|$, we have $\mathcal{I}(A \times A, L) \leq |L|^{1/2}|A|^2 \leq |L|^{3/4}|A|^{5/4}$.

Next we reformulate the lemma of Bourgain, Katz and Tao in terms of the incidence function $i$, c.f. [5, Lemma 1]. Sums are over all lines in $\mathbb{F}_p^2$ and not just those incident to some point of $A \times A$.

**Lemma 3** (Bourgain, Katz and Tao). Let $A \subseteq \mathbb{F}_p$.

$$\sum_{\text{all lines } \ell} i(\ell)^2 = |A|^4 + p|A|^2.$$ 

In particular

$$\sum_{\text{all lines } \ell} \left( i(\ell) - \frac{|A|^2}{p} \right)^2 \leq p|A|^2.$$ 

The simple yet powerful lemma of Bourgain, Katz, Tao was implicitly extended to not necessarily Cartesian product sets by Vinh [11]. The paper [5] contains other applications.

Next let $M$ be a parameter and set

$$L_M = \{ M < i(\ell) \leq 2M \}$$

to be the collection of lines from $L$ that are incident to between $M$ and $2M$ points in $A \times A$. We begin with an easy consequence of Lemma 3.

**Lemma 4.** Let $A \subseteq \mathbb{F}_p$ and $M$ be a real number. Suppose that $M \geq 2|A|^2/p$, then the set $L_M$ defined in (2) satisfies $|L_M| \leq 4p|A|^2/M^2$.

**Proof.** The hypothesis $i(\ell) \geq \frac{2|A|^2}{p}$ implies that $i(\ell) - \frac{|A|^2}{p} \geq \frac{i(\ell)}{2} \geq \frac{M}{2}$.

Lemma 3 now implies

$$\frac{M^2}{4}|L_M| \leq \sum_{\ell \in L_M} \left( i(\ell) - \frac{|A|^2}{p} \right)^2 \leq \sum_{\text{all lines } \ell} \left( i(\ell) - \frac{|A|^2}{p} \right)^2 \leq p|A|^2.$$

The claim follows. \qed
We now feed this bound to Theorem 2 to bound $L_M$. The use of Lemma 4 is a little strange, because we use it to prove that $|L_M|$ is not too big, so we may apply Theorem 2 and obtain a reasonable bound. The lemma plays a much more crucial role later.

**Lemma 5.** Let $A \subseteq \mathbb{F}_p$ and $M \geq 2|A|^{3/2}/p^{1/2}$ be a real number. Then the set $L_M$ defined in (2) satisfies

$$|L_M| = O\left(\frac{|A|^5}{M^4}\right).$$

In particular, under this hypothesis,

$$\sum_{\ell \in L_M} i(\ell)^3 = O\left(\frac{|A|^5}{M}\right) \text{ and } \sum_{\ell \in L_M} i(\ell)^4 = O(|A|^5).$$

**Proof.** The second claim follows by the first because, say,

$$\sum_{\ell \in L_M} i(\ell)^3 \leq 8M^3|L_M|.$$ 

To establish the first, we apply Theorem 2. Therefore we must confirm the condition $|A||L| = O(p^2)$. The hypothesis $M \geq 2|A|^{3/2}/p^{1/2}$ implies $M \geq 2|A|^2/p$. Lemma 4 can therefore be applied and in conjunction with the hypothesis $M^2 \geq 4|A|^3/p$ gives

$$|L_M| \leq \frac{4p|A|^2}{M^2} \leq \frac{p^2}{|A|}.$$ 

Hence, $|A||L_M| \leq p^2$ and Theorem 2 may be applied. It gives

$$M|L_M| \leq \mathcal{I}(A \times A, L_M) = O(|A|^{5/4}|L_M|^{3/4}).$$

The stated bound follows.

3 A straightforward bound for $T(A)$

Before proving Theorem 1, let us deduce from Lemma 3 a straightforward bound for $T(A)$.
Proposition 6. Let $A \subseteq \mathbb{F}_p$. Then

$$\left| T(A) - \left( \frac{|A|^6}{p} + 2|A|^4 \right) \right| \leq p|A|^3.$$ 

Proof. Sums are over all lines in $\mathbb{F}_p^2$. We combine the first part of Lemma 3 with the identity $\sum_\ell i(\ell) = (p+1)|A|^2$, which follows from the fact that every point in $A \times A$ is incident to $p+1$ lines in $\mathbb{F}_p^2$.

$$T(A) = \sum_\ell i(\ell)^3$$

$$= \sum_\ell i(\ell) \left( i(\ell) - \frac{|A|^2}{p} \right)^2 + 2 \frac{|A|^2}{p} \sum_\ell i(\ell)^2 - \left( \frac{|A|^2}{p} \right)^2 \sum_\ell i(\ell)$$

$$= \sum_\ell i(\ell) \left( i(\ell) - \frac{|A|^2}{p} \right)^2 + \frac{|A|^6}{p} + 2|A|^4 - \frac{|A|^6}{p^2}.$$ 

The claim now follows from the fact that $i(\ell) \leq |A|$ and the second part of Lemma 3 because

$$\left| T(A) - \left( \frac{|A|^6}{p} + 2|A|^4 \right) \right| \leq \sum_\ell i(\ell) \left( i(\ell) - \frac{|A|^2}{p} \right)^2 \leq |A| \sum_\ell \left( i(\ell) - \frac{|A|^2}{p} \right)^2 \leq p|A|^3.$$ 

We therefore see that to improve the proposition, and hence the range of $|A|$ for which $T(A) = O(|A|^6/p)$, we must show that it is impossible for “the mass” of

$$\sum_\ell \left( i(\ell) - \frac{|A|^2}{p} \right)^2 \approx p|A|^2$$

to come from lines that satisfy $i(\ell) = \Omega(|A|)$. In other words, we must roughly speaking show that it cannot be the case that $\Omega(p)$ lines are incident to $\Omega(|A|)$ points in $A \times A$. The theorem of Stevens and de Zeeuw guarantees this. In fact Theorem 2 implies that there are $O(|A|)$ lines incident to $\Omega(|A|)$ points in $A \times A$, which is nearly optimal. To maximise the gain we perform a more careful analysis.

4 Proof of Theorem 1

For $T(A)$ we break the sum of the cubes of $i(\ell)$ in three parts:
1. Those where \( i(\ell) \) is small (these give the term that resembles the expected count).

2. Those where \( i(\ell) \) is of medium size (controlled by Lemma 4).

3. Those where \( i(\ell) \) is large (controlled by Lemma 5 and dyadic decomposition).

The details are as follows.

\[
T(A) = \sum_{\text{all lines } \ell} i(\ell)^3
\]

\[
= \sum_{i(\ell) \leq \frac{2|A|^2}{p}} i(\ell)^3 + \sum_{\frac{2|A|^2}{p} \leq i(\ell) \leq \frac{2|A|^3/2}{p^{1/2}}} i(\ell)^3 + \sum_{i(\ell) \geq \frac{2|A|^{3/2}}{p^{1/2}}} i(\ell)^3. \tag{3}
\]

The first sum is bounded using the identity \( \sum_i i(\ell) = (p + 1)|A|^2 \):

\[
\sum_{i(\ell) \leq \frac{2|A|^2}{p}} i(\ell)^3 \leq \frac{4|A|^4}{p^2} \sum_i i(\ell) = O \left( \frac{|A|^6}{p} \right).
\]

The second sum is bounded using Lemma 4 and the observation that \( i(\ell) \geq 2|A|^2/p \), then \( i(\ell) \leq 2(i(\ell) - |A|^2/p) \):

\[
\sum_{\frac{2|A|^2}{p} \leq i(\ell) \leq \frac{2|A|^{3/2}}{p^{1/2}}} i(\ell)^3 \leq \frac{2|A|^{3/2}}{p^{1/2}} \sum_i 4 \left( i(\ell) - \frac{|A|^2}{p} \right)^2 = O(p^{1/2}|A|^{7/2}).
\]

The third sum is bounded using Lemma 5 and dyadic decomposition:

\[
\sum_{i(\ell) \geq \frac{2|A|^{3/2}}{p^{1/2}}} i(\ell)^3 = \sum_{\frac{2|A|^{3/2}}{p^{1/2}} \leq 2^j \leq \frac{|A|}{2}} \sum_{\ell \in L_{2^j}} i(\ell)^3
\]

\[
= O \left( \sum_{2^j \geq \frac{2|A|^{3/2}}{p^{1/2}}} \frac{|A|^{5}}{2^j} \right)
\]

\[
= O \left( \frac{|A|^{5}}{|A|^{3/2} p^{1/2}} \right) = O(p^{1/2}|A|^{7/2}).
\]

Substituting the three bounds into (3) gives \( T(A) = O \left( \frac{|A|^6}{p} + p^{1/2}|A|^{7/2} \right) \).
A nearly identical argument works for $Q(A)$.

$$Q(A) = \sum_{\text{all lines } \ell} i(\ell)^4$$

$$= \sum_{i(\ell) \leq \frac{2|A|^2}{p}} i(\ell)^4 + \sum_{\frac{2|A|^2}{p} \leq i(\ell) \leq \frac{2|A|^{3/2}}{p^{1/2}}} i(\ell)^4 + \sum_{i(\ell) \geq \frac{2|A|^{3/2}}{p^{1/2}}} i(\ell)^4. \quad (4)$$

The first sum is bounded using the identity $\sum \ell i(\ell) = (p + 1)|A|^2$:

$$\sum_{i(\ell) \leq \frac{2|A|^2}{p}} i(\ell)^4 \leq \frac{8|A|^6}{p^3} \sum_{\ell} i(\ell) = O\left(\frac{|A|^8}{p^2}\right).$$

The second sum is bounded using Lemma 4:

$$\sum_{\frac{2|A|^2}{p} \leq i(\ell) \leq \frac{2|A|^{3/2}}{p^{1/2}}} i(\ell)^4 \leq \frac{4|A|^3}{p} \sum_{\ell} 4 i(\ell) - \frac{|A|^2}{p} \right)^2 = O(|A|^5).$$

The third sum is bounded using Lemma 5 (and is this time of greater order of magnitude than the second):

$$\sum_{i(\ell) \geq \frac{2|A|^{3/2}}{p^{1/2}}} i(\ell)^3 = \sum_{\frac{2|A|^{3/2}}{p^{1/2}} \leq 2^j \leq \frac{|A|}{2}} \sum_{\ell \in L_2} i(\ell)^3$$

$$= O\left(\sum_{2^j \leq |A|} |A|^{5}\right)$$

$$= O(\log(|A||A|^5)).$$

Substituting the three bounds into (4) gives $Q(A) = O\left(\frac{|A|^8}{p^2} + \log(|A||A|^5)\right)$.

**References**

[1] E. Aksoy-Yazici, B. Murphy, M. Rudnev, and I. D. Shkredov. Growth estimates in positive characteristic via collisions. Accepted Int. Math. Res. Not. IMRN arXiv:1512.06613, 2015. (p. 3)
[2] J. Bourgain, N. H. Katz, and T. Tao. A sum-product estimate in finite fields, and applications. *Geom. Funct. Anal.*, 14(1):27–57, 2004. (pp. 1 and 4)

[3] G. Elekes and I. Z. Ruzsa. Few sums, many products. *Studia Sci. Math. Hungar.*, 40(3):301–308, 2003. (p. 3)

[4] K. Ford. The distribution of integers with a divisor in a given interval. *Ann. of Math. (2)*, 168(2):367–433, 2008. (p. 2)

[5] B. Murphy and G. Petridis. A point-line incidence identity in finite fields, and applications. *Mosc. J. Comb. Number Theory*, 6(1):63–94, 2016. (p. 5)

[6] B. Murphy, G. Petridis, O. Roche-Newton, M. Rudnev, and I. D. Shkredov. New results on sum-product type growth over fields. arXiv:1702.01003, 2017. (pp. 1 and 4)

[7] G. Petridis. Products of differences in prime order finite fields. arXiv:1602.02142, 2016. (p. 1)

[8] M. Rudnev. On the number of incidences between planes and points in three dimensions. *Combinatorica*. To appear. (p. 3)

[9] S. Stevens and F. de Zeeuw. An improved point-line incidence bound over arbitrary fields. arXiv:1609.06284, 2016. (pp. 1 and 4)

[10] E. Szemerédi and W. T. Trotter. Extremal problems in discrete geometry. *Combinatorica*, 3(3–4):381–392, 1983. (p. 3)

[11] L. A. Vinh. The Szemerédi-Trotter type theorem and the sum-product estimate in finite fields. *European J. Combin.*, 32(8):1177–1181, 2011. (p. 5)

Department of Mathematics, University of Georgia, Athens, GA, USA.

Email address: giorgis@cantab.net