AN EXPLICIT INCIDENCE THEOREM IN $\mathbb{F}_p$

HARALD ANDRÉS HELFGOTT AND MISHA RUDNEV

Abstract. Let $P = A \times A \subset \mathbb{F}_p \times \mathbb{F}_p$, $p$ a prime. Assume that $P = A \times A$ has $n$ elements, $n < p$. See $P$ as a set of points in the plane over $\mathbb{F}_p$. We show that the pairs of points in $P$ determine $\geq cn^{1+\frac{1}{267}}$ lines, where $c$ is an absolute constant.

We derive from this an incidence theorem: the number of incidences between a set of $n$ points and a set of $n$ lines in the projective plane over $\mathbb{F}_p$ ($n < \sqrt{p}$) is bounded by $Cn^{2-\frac{1}{10678}}$, where $C$ is an absolute constant.

1. Introduction

In 1983 J. Beck proved the following incidence theorem in $\mathbb{R}^2$. Let $P$ be a set of points in $\mathbb{R}^2$. Then either $P$ contains $c|P|$ points on a straight line, or the pairs of points of $P$ determine at least $c|P|^2$ distinct lines, where $c$ is an absolute constant. (We write $|S|$ for the number of elements of a finite set $S$; $c$ and $C$ will henceforth denote some absolute constants, changing from one place to another.) Beck’s original paper ([1]) was followed – in the same issue of the same journal – by a result of Szemerédi and Trotter stating that the number of incidences between $m$ straight lines and $n$ points in $\mathbb{R}^2$ is $O(m + n + (mn)^{\frac{2}{3}})$. The Szemerédi-Trotter theorem implies Beck’s theorem as a corollary.

We prove a finite field version of Beck’s theorem.

Theorem 1. Let $A \subset \mathbb{F}_p$, where $p$ is larger than an absolute constant. Let $P = A \times A \subset \mathbb{F}_p^2$. Let $L = L(P)$ be the set of all lines determined by pairs of points of elements of $P$. If $|A| < \sqrt{p}$, then

$$|L(P)| \geq c|P|^{1+\frac{1}{267}},$$

where $c > 0$ is an absolute constant.

All constants here and later are independent of $p$.

The statement $|L(P)| \geq c|P|^{1+\frac{1}{267}}$ cannot possibly hold in $\mathbb{F}_p^2$ in full generality, as $P = \mathbb{F}_p^2$ itself generates only $O(|P|)$ lines. A non-trivial theorem of Beck type is implicit in the well-known paper of Bourgain, Katz, and Tao ([4]) – namely, their results imply that, if $P \subset \mathbb{F}_p^2$ is a Cartesian product and $|P| < p^{2-\Delta}$ ($\Delta > 0$), then $|L(P)| = O(|P|^{1+\delta(\Delta)})$ for some $\delta(\Delta) > 0$. However, there was no attempt to make explicit or optimise $\delta(\Delta)$.

2000 Mathematics Subject Classification. 68R05,11B75.
The problem of studying $L(P)$ is an incidence problem, and [4] proves that for $n < p^{2-\Delta}$, the number of incidences between $n$ lines and $n$ points is $O\left(|n|^{\frac{2}{3}-\delta_3(\Delta)}\right)$. This statement happens to be on about the same level of generality as claiming that a set of $n$ points that is a Cartesian product $A \times A$ (with $n = |A|^2 < p^{2-\Delta}$) determines $\Omega\left(n^{1+\delta_3(\Delta)}\right)$ distinct lines. These results in [4] stem from a non-trivial sum-product estimate in $\mathbb{F}_p$ proven in [4]. The sum-product estimate of [4] (the following formulation also includes Konyagin’s contribution [7]) says that, if $A + A$ and $A \cdot A$ denote, respectively, the set of all sums and products of pairs of elements of $A \subset \mathbb{F}_p$, then, as long as $|A| < p^{1-\Delta}$, one has

\begin{equation}
\max(|A + A|, |A \cdot A|) \geq c|A|^{1+\delta_3(\Delta)},
\end{equation}

for some absolute $c$. The quantitative relation between the deltas was not established.

(At this point one traditionally mentions the Erdős-Szemerédi conjecture, which states that, if $A$ is a subset of integers, then $\max(|A + A|, |A \cdot A|) \geq c|A|^{1+\delta}$ holds for any $0 < \delta < 1$, where $c$ is allowed to depend only on $\delta$, but not on $|A|$.)

We return to $\mathbb{F}_p$ for the rest of the paper. On the level of the existence of positive exponents, a nontrivial sum-product estimate implies a non-trivial incidence or Beck-type theorem, and conversely. Garaev [6] succeeded in obtaining a quantitative sum-product estimate in $\mathbb{F}_p$: for a small enough (say $|A| < \sqrt{p}$) subset $A$ of $\mathbb{F}_p$, either $A + A$ or $A \cdot A$ has cardinality at least $|A|^{1+\frac{1}{11}}$, up to a multiple of a power of $c \log |A|$. Katz and Shen [8] elaborated on a particular application of the Plünnecke-Ruzsa inequality in Garaev’s proof and improved the result to $|A|^{1+\frac{1}{7}}$, up to a multiple of a power of $c \log |A|$. Bourgain and Garaev [3] incorporated a covering argument (whose variant we cite as Lemma [3]) and improved the estimate to $|A|^{1+\frac{1}{72}}$, up to a multiple of a power of $c \log |A|$. Li [9] showed that a multiple of a power of $\log |A|$ can be done away with; thus the best result known states that, for $|A| < \sqrt{p}$,

$$\max(|A + A|, |A \cdot A|) \geq c|A|^{1+\frac{1}{72}},$$

where $c$ is an absolute constant.

Our construction uses the techniques laid out in the above-mentioned papers and locally follows rather closely the exposition from some of those papers. The main result, Theorem 1, an explicit Beck type incidence theorem, implies the following incidence theorem for $n$ points and $n$ straight lines in the projective plane $\mathbb{P}^2(\mathbb{F}_p)$.

**Theorem 2.** If $P$ and $L$ are sets of points and lines in $\mathbb{P}^2(\mathbb{F}_p)$ with $|P|, |L| = n < p$, then the number of incidences

$$I(P, L) = |\{(p, l) \in P \times L : p \in l\}| \leq Cn^{\frac{3}{2} - \frac{1}{10678}}$$

for some absolute $C$.

We call $\{(p, l) \in P \times L : p \in l\}$ the set of incidences.

The proof of this theorem repeats the pigeonholing argument of [4] until it merges with the proof of our Theorem [4]. It is given at the end of this note.
Remark 3. In both Theorems 1 and 2 one can easily extend the estimate to larger sets, by a straightforward adaptation of Case (i) of the estimate (3.19) in the end of our proof. We haven’t done so aiming at an estimate which does not contain $p$ explicitly; we leave the case of “larger” sets to the interested reader.

1.1. Acknowledgments. H. A. Helfgott is supported in part by EPSRC grant EP-E054919/1. The authors would like to thank M. Garaev, T. Jones and O. Roche-Newton for their helpful comments on the draft of this paper.

2. Background in arithmetic combinatorics

We use the following largely standard arithmetic combinatorics lemmata. In the sequel, in order to suppress constants, we will use the $\ll, \gg, \approx$ notations in estimates: $|X| \gg |Y|$ means $|X| \geq c|Y|$ for some $c$, $|X| \ll |Y|$ means $|X| \leq c|Y|$ for some $c$, $|X| \approx |Y|$ means $|X| \leq C_1|Y|$ and $|X| \geq C_2|Y|$ hold for some $C_1, C_2$.

We abuse the English language in accordance with these notations by saying “at least”, “at most”, or approximately in the sense conveyed by the symbols $\gg, \ll, \approx$, respectively. To avoid confusion, we will enclose “at least” and “at most” in quotation marks when we use them in this way.

Thus, for example, saying that $|X|$ is “at least" $|Y|$ means $|X| \gg |Y|$.

We adopt the following formulation of the Balog-Szemerédi-Gowers theorem.

Lemma 1 (Balog-Szemerédi-Gowers theorem). Let $X, Y$ be additive sets of $n$ elements, and $\alpha \in (0,1)$. Suppose that there is a set of $\alpha n^2$ pairs of elements $(x, y) \in X \times Y$ on which the sum $x + y$ takes at most $n$ distinct values.

Then there exist subsets $X' \subseteq X, Y' \subseteq Y$, with $|X'|, |Y'| \gg \alpha n$, such that

$$|\text{Range of } x + y \text{ on } X' \times Y'| \ll \alpha^{-5} n.$$  

The modern graph-theoretical proof of the Balog-Szemerédi-Gowers theorem can be found in [13, Thm. 2.29]. The proof as appears in that standard reference appears to have a typographical error, however, which leads to an exponent of $-4$, rather than to the correct (and somewhat weaker) exponent of $-5$ that we have above. For a proof yielding the exponent $-5$, see [5].

We adopt the following form for the Plünnecke-Ruzsa inequality, due to Ruzsa ([10]).

Lemma 2. Let $Y; X_1, \ldots X_k$ be additive sets. Then there exists a non-empty subset $Y' \subseteq Y$, such that

$$|Y' + X_1 + \ldots + X_k| \leq \prod_{i=1}^{k}\frac{|Y + X_i|}{|Y|^{k-1}}|Y'|.$$  

Ruzsa’s inequality immediately implies that

$$|X_1 + \ldots + X_k| \leq \prod_{i=1}^{k}\frac{|Y + X_i|}{|Y|^{k-1}},$$  

for any “dummy set” $Y$. To allow for $-$ signs as well as $+$ signs, this inequality is often used in conjunction with the Ruzsa distance inequality

$$|X_1 - X_2| \leq \frac{|X_1 - X_3||X_3 - X_2|}{|X_3|}. \tag{2.3}$$

(See, e.g., [13, Lemma 2.6] for a brief proof of (2.3).)

Remark 4. Katz and Shen ([8]) showed that, at the expense of acquiring a constant in the left-hand side of (2.1), one can make $Y'$ contain an arbitrarily large proportion of $Y$. This trick was used to gain an improvement in ([8]) and subsequent above-mentioned papers on the sum-product problem. We have not found a way to take advantage of it here, as we shall be dealing with set families, controlling their intersections, refinements therefore being apparently forbidden. I.e., we essentially use Plünnecke-Ruzsa only in the “crude” form (2.2).

Finally, we need the following covering lemma (see e.g. [3], [11], [9]; the proof is a Cauchy-Schwartz type averaging argument).

**Lemma 3.** Let $X_1$ and $X_2$ be additive sets. Then for any $\varepsilon \in (0, 1)$ and some constant $C(\varepsilon)$, there exist $\frac{C(\varepsilon)}{|X_2|} \min(|X_1 + X_2|, |X_1 - X_2|)$ translates of $X_2$ whose union contains not less than $(1 - \varepsilon)|X_1|$ elements of $X_1$.

3. **Proof of Theorem 1**

We will prove the result more generally for $P = A_1 \times A_2 \subset \mathbb{F}_p \times \mathbb{F}_p$, $|A_1| = |A_2| = n$, $n < \sqrt{p}$.

Let $L(P)$ be the set of straight lines generated by pairs of elements of $P$. Suppose that

$$|L(P)| \approx n^{2+2\delta}. \tag{3.1}$$

where $\delta < \frac{1}{267}$. We will show how to reach a contradiction.

The approximately $n^4$ pairs of distinct points of $P$ are distributed between approximately $n^{2+2\delta}$ lines. This implies that a positive proportion of approximately $n^4$ pairs of those points are supported on rich lines, meaning lines with “at least” $n^{1-\delta}$ points on each. These rich lines thus contain “at least” $n^{5-\delta}$ collinear triples of distinct points of $P$.

Hence, the equation

$$\begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = 0, \quad x_1, x_2, x_3 \in A_1, \ y_1, y_2, y_3 \in A_2 \tag{3.2}$$

has “at least” $n^{5-\delta}$ solutions. Then, for some fixed and non-equal $y_1, y_2 \in A_2$, the equation

$$\begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = 0, \quad x_1, x_2, x_3 \in A_1, \ y_3 \in A_2 \tag{3.3}$$
AN EXPLICIT INCIDENCE THEOREM IN $\mathbb{F}_p$

has “at least” $n^{3-\delta}$ solutions. Since we can translate and dilate $A_2$, we can assume
$y_1 = 0$, $y_2 = 1$ without loss of generality. Then the condition (3.3) turns into a claim
that
\begin{equation}
(3.4) \quad x_1(1 - y_3) + x_2y_3 \in A_1 \quad x_1, x_2 \in A_1, \quad y_3 \in A_2
\end{equation}
happens for “at least” $n^{3-\delta}$ triples $(x_1, x_2, y_3)$. Assuming $y_3 \neq 0, 1$ does not change
the situation, as long as $\delta < 1$. Let us define $B$ as the set of elements of the form
$b = \frac{y_3}{1 - y_3}, \ y_3 \in A_2 \setminus \{0, 1\}$. Clearly $|B| \approx |A_2| = n$ and $B$ does not contain zero.
(The set-up here is similar to that of Theorem C in Bourgain ([2]), which would
at this point yield the existence of $\delta$, with a possibility of its quantitative estimate.
However, if one chases through Bourgain’s arguments, the value of $\delta$ appears to be
considerably smaller than what we obtain.)

Let $B_1$ be the set of popular elements of $B$, meaning the set of all $b \in B$ such
the equation (3.4) has “at least” $n^{2-\delta}$ solutions $(x_1, x_2, y_3) \in A_1 \times A_1 \times A_2$ with
$\frac{y_3}{1 - y_3} = b$. By the pigeonhole principle,
$|B_1| \gg n^{1-\delta}$.

Besides, fixing $b$ fixes $y_3$, so for each $b \in B_1$ the condition
\begin{equation}
x_1 + bx_2 \in \frac{1}{1 - y_3(b)} A_1
\end{equation}
holds for “at least” $n^{2-\delta}$ pairs $(x_1, x_2) \in A_1 \times A_1$.

Applying the Balog-Szemerédi-Gowers theorem for each $b$, yields the existence of the
subsets $A_b^1$ and $A_b^2$ of $A_1$, with $|A_b^i| \gg |A_1|^{1-\delta} = n^{1-\delta}, \ i = 1, 2$, such that
\begin{equation}
(3.5) \quad |A_b^1 + bA_b^2| \ll n^{1+5\delta}.
\end{equation}

Let us restate that
\begin{equation}
(3.6) \quad |A_b^1|, |A_b^2|, |B_1| \gg n^{1-\delta}.
\end{equation}

Let $A_b = A_b^1 \times A_b^2$. In view of (3.6) and by Cauchy-Schwartz, since each $A_b$ is a
subset of $A \times A$,
\begin{equation}
|B_1|n^{2-2\delta} \ll \sum_{b \in B_1} |A_b| \leq n \left( \sum_{b, b' \in B_1} |A_b \cap A_{b'}| \right)^{\frac{1}{2}},
\end{equation}
so for some $b_* \in B_1$ (which can be assumed non-zero), denoting $A_* = A_{b_*}$, we have
\begin{equation}
(3.7) \quad \sum_{b \in B_1} |A_b \cap A_*| \gg |B_1|n^{2-4\delta}.
\end{equation}

Let $B_2$ be a popular subset of $B_1$, namely a set such that for all $b \in B_2$ we have
\begin{equation}
(3.8) \quad A_{b \wedge *} \equiv |A_b \cap A_*| \gg n^{2-4\delta};
\end{equation}
clearly,
\begin{equation}
(3.9) \quad |B_2| \gg n^{1-5\delta}.
\end{equation}
Besides, each $A_{b \wedge *}$ is a cartesian product:

$$A_{b \wedge *} = A_{b \wedge *}^1 \times A_{b \wedge *}^2,$$

and therefore, for $i = 1, 2$

$$A_{b \wedge *}^1 \equiv A_{b_i}^1 \cap A_{b}^i \gg n^{1-4\delta}. \quad (3.10)$$

We now apply the Plünnecke-Ruzsa inequality (2.2) with $k = 2$, as well as (3.5), and the set cardinality estimates (3.6) and (3.10) to draw the following conclusions. For each $b \in B_1$ we have

$$|A_{b}^1 + A_{b}^1| \leq \frac{|A_{b}^1 + bA_{b}^2|^2}{|A_{b}^1|} \ll n^{1+11\delta}, \quad (3.11)$$

$$|A_{b}^2 + A_{b}^2| \leq \frac{|A_{b}^1 + bA_{b}^2|^2}{|A_{b}^2|} \ll n^{1+11\delta}.$$ 

Furthermore, for each $b \in B_2$:

$$|b_* A_{b}^2 + bA_{b}^2| \leq \frac{|b_* A_{b}^2 + b_* A_{b \wedge *}^2||b_* A_{b \wedge *}^2 + bA_{b}^2|}{|A_{b \wedge *}^2|} \ll n^{15\delta}\frac{|b_* A_{b \wedge *}^2 + bA_{b}^2|}{|A_{b \wedge *}^2|} \ll n^{1+29\delta}.$$ 

by (2.2), (3.10), (3.11), and (3.5). Then

$$|b_* A_{b}^2 + bA_{b}^2| \ll \frac{|b_* A_{b}^2 + b_* A_{b \wedge *}^2||b_* A_{b \wedge *}^2 + bA_{b}^2|}{|A_{b \wedge *}^2|} \ll n^{1+44\delta}.$$ 

and

$$|b_* A_{b}^2 + bA_{b}^2| \ll \frac{|b_* A_{b}^2 + b_* A_{b \wedge *}^2||b_* A_{b \wedge *}^2 + bA_{b}^2|}{|A_{b \wedge *}^2|} \ll n^{1+59\delta}. \quad (3.12)$$

Throughout the rest of the proof, to save on the number of indices used, let us refer to $A_{b}^2$ as $X$ and to $b_*^{-1}B_2$ as $Y$.

Let us use the symbol

$$K = \max_{y \in Y}|X + yX|, \quad \text{so} \quad K \ll n^{1+59\delta}. \quad (3.13)$$

Consider the equation

$$x_2 + y\tilde{x}_1 = \tilde{x}_2 + yx_1, \quad x_1, x_2, \tilde{x}_1, \tilde{x}_2 \in X, \ y \in Y. \quad (3.14)$$

This equation has at least $\frac{|Y||X|^4}{K}$ solutions, as follows by applying Cauchy-Schwartz for each individual $Y$ and then summing over $y \in Y$. Equation (3.14) is equivalent to

$$x_2 - \tilde{x}_2 = y(x_1 - \tilde{x}_1), \quad x_1, x_2, \tilde{x}_1, \tilde{x}_2 \in X, \ y \in Y.
Hence, for some fixed \((\tilde{x}_1, \tilde{x}_2) \in X \times X\), the above equation has at least \(\frac{|Y||X|^2}{K}\) solutions. Let \(X_1 = X - \tilde{x}_1, X_2 = X - \tilde{x}_2\) be translates of \(X\) by \(\tilde{x}_1\) and \(\tilde{x}_2\), respectively. The equation

\[(3.15) \quad v = yu, \quad u \in X_1, \ v \in X_2, \ y \in Y\]

has at least \(\frac{|Y||X|^2}{K}\) solutions. Consider the set \(X_1 \times X_2\) in \(\mathbb{F}_p^2\). The bound we have just given as to the number of solutions of equation (3.15) can be rephrased as saying that the set of straight lines through the origin with slopes in \(Y\) makes at least \(\frac{|Y||X|^2}{K}\) incidences with \(X_1 \times X_2\).

In the remainder of the proof will assume that \(\frac{|X|^2}{K} \gg 1\), i.e., that the lines in question contain “at least” one point each on average. This follows immediately from (3.6) and (3.13) once we assume \(\delta < \frac{1}{12}\).

Not less than 50% of the incidences specified by (3.15) are contributed by rich lines with “at least” \(\frac{|X|^2}{K}\) points thereon. The number of rich lines is not greater than \(|Y|\), and “at least” \(\frac{|Y||X|}{K}\).

Those “at least” \(\frac{|Y||X|^2}{K}\) points of \(X_1 \times X_2\) lying on rich lines can have \(|X|\) different abscissae. Hence, there is a vertical set \(u_0 \times X_2\) for some non-zero \(u_0 \in X_1\) intersected by “at least” \(\frac{|Y||X|}{K}\) rich lines.

Thus, we have a subset \(Y_1 \subset (X_2 \cap u_0Y)\) of cardinality

\[(3.16) \quad |Y_1| \gg \frac{|Y||X|}{K}.
\]

In the original notations, \(Y_1\) lies in the intersection of \(u_0b_s^{-1}B_2\) and some translate of \(A^2_s\); besides,

\[(3.17) \quad |Y_1| \gg n^{1-6\delta},\]

by the bounds (3.6), (3.9), (3.12).

Let \(R\) be the set of all elements expressed via \(r = \frac{p-2}{s-t}\), where \(p, q, s, t\) are elements of \(Y_1\) and \(s \neq t\). Let us consider two cases: (i) \(|R| \geq |Y_1|^2\) and (ii) \(|R| < |Y_1|^2\). Since \(n < \sqrt{p}\), \(R = \mathbb{F}_p\) is a possibility only in case (i).

Let us consider Case (ii) first. For any \(\xi \notin R\) and any non-equal \(y_1, y_2 \in Y_1\), the sum \(y_1 + \xi y_2\) has a single realisation, as

\[(3.18) \quad y_1 + \xi y_2 = y_1' + \xi y_2', \quad y_1, y_2, y_1', y_2' \in Y_1\]

would imply \(\xi = \frac{y_1-y_1'}{y_2'-y_2}\). Thus, for every subset \(Y_1' \subset Y\) and any nonzero \(\xi \notin R\),

\[(3.19) \quad |Y_1' + \xi Y_1'| \leq |Y_1'|^2.
\]

For any \(y \neq 0\) we have some \(p, q, s, t \in Y_1\), such that \(\xi = \frac{p-2}{s-t} + y\) lies in the complement of \(R\), for otherwise \(R + \{y\} = R\), which is possible only if \(R = \mathbb{F}_p\). In particular, this holds for \(y = 1\). Recall that \(Y_1\) is a subset of \(u_0b_s^{-1}B_2\), and so we
may regard \( p, q, s, t \) as elements of \( B_2 \). We have, then, some fixed \( p, q, s, t \in B_2 \) such that

\[
|Y_1|^2 \ll |Y_1' + \xi Y_1'| \leq |Y_1' + Y_1' + \frac{p-q}{s-t} Y_1'|
\]

for any \( Y_1' \subseteq Y_1 \), that constitutes a positive proportion of \( Y_1 \), to be chosen next.

We now use Lemma 3. Let us first show that for any \( b \in B_2 \) we can cover 99% of the elements of the sets \( bY_1 \) (a subset of a translation of \( bA_2^* \)) or \( -bY_1 \) (a subset of a translation of \( -bA_2^* \)) by “at most” \( n^{248} \) translates of the set \( A_1^* \). Indeed, \( A_{b \wedge s}^* \cap A_1^* \) is a subset of \( A_1^* \), and by Lemma 3 and (2.2), we can cover 99% of the elements of either \( bY_1 \) or \( bY_1 \) by “at most”

\[
|A_{b \wedge s}^* + bY_1| \leq \frac{|A_{b \wedge s}^* + bA_2^*|}{|A_{b \wedge s}^*|} \leq \frac{|A_{b \wedge s}^* + A_2^*||bA_{b \wedge s}^* + bA_2^*|}{|A_{b \wedge s}^*||A_{b \wedge s}^*|} \ll n^{248}
\]

translates of \( A_{b \wedge s}^* \), and hence of \( A_1^* \). In the last estimate we’ve used (3.5), (3.10), and (3.11).

This altogether enables us to choose \( Y_1' \) as a subset containing at least 50% of \( Y_1 \), and such that \( (p-q)Y_1' \) gets covered by “at most” \( n^{488} \) translates of \( A_2^* + A_1^* \). Let us now, in the same vein, \( A_2^* \) be a subset containing at least 50% of \( A_2^* \), such that \( (s-t)A_2^* \) gets covered by “at most” \( n^{488} \) translates of \( A_1^* + A_1^* \). Then we apply Plünnecke-Ruzsa to (3.20) as follows:

\[
|Y_1' + Y_1' + \frac{p-q}{s-t} Y_1'| \ll \frac{|A_2^* + Y_1' + Y_1'||\tilde{A}_2^* + \frac{p-q}{s-t} Y_1'|}{|A_2^*|} \ll \frac{|A_2^* + A_2^* + A_1^*|}{n^{1-\delta}} |\tilde{A}_2^* + \frac{p-q}{s-t} Y_1'| \\
\ll n^{188} |\tilde{A}_2^* + \frac{p-q}{s-t} Y_1'|
\]

after applying Plünnecke-Ruzsa with \( k = 3 \) and the “dummy set” \( b_{s^{-1}} A_1^* \), using (3.6).

The covering argument above implies that

\[
|\tilde{A}_2^* + \frac{p-q}{s-t} Y_1'| \ll n^{968} |A_2^* + A_1^* + A_1^* + A_1^*| \ll n^{1+1196},
\]

by applying Plünnecke-Ruzsa with the “dummy set” \( b_s A_2^* \) and \( k = 4 \), using (3.5) and (3.6).

Therefore returning to (3.20), we have

\[
|Y_1'|^2 \ll n^{1+137 \delta}.
\]

Comparing this with (3.17), and recalling that \( Y_1' \) contains at least half of the elements of \( Y_1 \), we conclude that \( 267 \delta \geq 1 \), so \( \delta \geq \frac{1}{267} \). This ends Case (ii) and essentially ends the proof of Theorem 1.

Indeed, to analyse Case (i) we observe that if \( |R| \geq |Y_1|^2 \), then, summing the number of ordered quadruples \((y_1, y_2, y_1', y_2')\) of elements of \( Y_1 \) satisfying equation (3.18) over all \( \xi \in R \) we get no more than \( 2|R||Y_1'|^2 \) for the total number of solutions, pentuples \((y_1, y_2, y_1', y_2', \xi)\). Indeed, given \( \xi \), the solutions can be either trivial, with
(y_1, y_2) = (y'_1, y'_2) or not. The total number of non-trivial solutions is |Y_1|^4, since such a solution determines ξ; the total number of trivial ones is |R||Y_1|^2. Under the assumption of Case (i) the number of trivial solutions dominates the number of non-trivial ones. Hence, by the pigeonhole principle and Cauchy-Schwartz, there exists \( \xi = \frac{p-q}{s-t} \in R \) such that

\[
|Y_1' + \xi Y_1'| \gg \frac{|Y_1'|^4}{|R||Y_1|^2/|R|} \gg |Y_1'|^2.
\]

for every subset \( Y_1' \subset Y_1 \) with \( |Y_1'| \gg |Y_1| \).

This amounts to a simplification of (3.20), with \( \tilde{Y}_1' \) replacing \( \tilde{Y}_1' + \tilde{Y}_1'' \) in its right-hand side. The ensuing estimates, done in exactly the same way as in Case (ii) are therefore better – by a factor of \( n^{16\delta} \), as follows by inspection of (3.21) and (3.22).

This ends the proof of Theorem 1. □

4. Proof of Theorem 2

We follow [4], Section 6. Suppose, for contradiction, that for some \( (P, L) \) one has

\[
I(P, L) \approx n^{3/2-\epsilon},
\]

for some \( \epsilon > 0 \). We shall give the lower bound for such \( \epsilon \). Note that the role played by the parameter \( n \) in this theorem is different from the proof of Theorem 1.

Let us first off erase the points in \( P \) incident to more than \( Cn^{1/2+\epsilon} \) lines of \( L \), without changing the notations \( (P, L) \). This can be done, as the maximum number of incidences that can come from the set \( P_+ \) of such points is

\[
I(P_+, L) = \sum_{p \in P_+} \sum_{\ell \in L} \delta_{p,\ell} \leq \frac{1}{Cn^{3/2+\epsilon}} \sum_{p \in P_+} \left( \sum_{\ell \in L} \delta_{p,\ell} \right)^2 \leq \frac{1}{Cn^{3/2+\epsilon}} \sum_{p \in P_+} n^2 \sum_{\ell, \ell' \in L} \delta_{p,\ell} \delta_{p,\ell'} \ll \frac{n^2}{Cn^{3/2+\epsilon}},
\]

since any two distinct lines of \( L \) meet at at most a single point of \( P_+ \). (Here we use the notation \( \delta_{p,\ell} = 1 \) if the point \( p \) is incident to the line \( \ell \), \( \delta_{p,\ell} = 0 \) otherwise.)

This having been done, let \( P_1 \) be the set of popular points of \( P \), each incident to “at least” (recall that saying “at least” implies a suitable constant \( c \)) \( n^{1/2-\epsilon} \) lines of \( L \). We have \( I(P_1, L) \approx I(P, L) \) by the following argument:

\[
I(P, L) - I(P_1, L) = \sum_{p \in P} \sum_{\ell \in L} \delta_{p,\ell} - \sum_{p \in P_1} \sum_{\ell \in L} \delta_{p,\ell} < \sum_{p \in P} \sum_{\ell \in L} cn^{1/2-\epsilon} \leq cn^{3/2-\epsilon} < \frac{1}{2} I(P, L)
\]

for \( c \) small enough.

We can now refine \( L \) to the subset \( L_1 \) of popular lines, each incident to “at least” \( n^{3/2-\epsilon} \) points of \( P_1 \). By the pigeonhole principle, it still contributes a positive proportion of incidences. Let \( P_2 \subset P_1 \) with respect to \( L_1 \). (That is, \( p \in P_2 \) is an element of \( P_2 \) if and only if it is incident to “at least” \( n^{3/2-\epsilon} \) lines of \( L_1 \).)
Once again, \( I(P_2, L_1) \) satisfies (1.1). (The refinement process could be iterated any finite number of times, with the constants obviously getting worse.)

For \( p \in P_2 \), let \( P_p \) be a subset of all points of \( P_1 \), which are connected to \( p \) by some line from \( L_1 \). By definitions of \( P_2, L_1 \) we have \( |P_p| \gg n^{1-2\epsilon} \), for each \( p \in P_2 \). Thus, by Cauchy-Schwartz

\[
|P_2|n^{1-2\epsilon} \leq \sum_{p \in P_2} |P_p| \leq \sqrt{|P_1|} \sqrt{\sum_{\tilde{p}, \tilde{p} \in P_2} |P_{\tilde{p}} \cap P_{\tilde{p}}|},
\]

and so

\[
\sum_{\tilde{p}, \tilde{p} \in P_2} |P_{\tilde{p}} \cap P_{\tilde{p}}| = \sum_{\tilde{p}, \tilde{p} \in P_2} |P_{\tilde{p}} \cap P_{\tilde{p}}| - \sum_{\tilde{p} \in P_2} |P_{\tilde{p}}| \gg \frac{|P_2|^2}{|P_1|} n^{2-4\epsilon} - O(n^2).
\]

Now since \( I(P_2, L_1) \approx n^{\frac{3}{2}-\epsilon} \) and each point of \( P_2 \) is incident to at most \( \ll n^{\frac{1}{2}+\epsilon} \) lines of \( L \), we have \( |P_2| \gg n^{1-2\epsilon} \), and thus

\[
\sum_{\tilde{p}, \tilde{p} \in P_2} |P_{\tilde{p}} \cap P_{\tilde{p}}| \gg \frac{|P_2|^2}{|P_1|} n^{2-4\epsilon}.
\]

Therefore one can fix some \((\tilde{p}, \tilde{\tilde{p}}) \in (P_2 \times P_2), \tilde{p} \neq \tilde{\tilde{p}}\), such that

\[
P_3 \equiv |P_{\tilde{p}} \cap P_{\tilde{\tilde{p}}}| \gg \frac{n^{2-4\epsilon}}{|P_1|} \gg n^{1-4\epsilon}.
\]

Each point of \( P_3 \subseteq P_1 \) is incident to “at least” \( n^{\frac{1}{2}-\epsilon} \) lines of the original set of lines \( L \). Without loss of generality, after a projective transformation we can place the points \( \tilde{p} \) and \( \tilde{\tilde{p}} \) on the line at infinity, so that the “at most” \( n^{\frac{1}{2}+\epsilon} \) lines of \( L_1 \) emanating from these points can be viewed as being parallel to the \( x \) and \( y \) coordinate axes. In other words, for some \( A, B \) of cardinality “at most” \( n^{\frac{1}{2}+\epsilon} \) each, we have a subset \( P_3 \) of \( A \times B \) of cardinality “at least” \( n^{1-4\epsilon} \), such that the number of incidences of \( P_3 \) with \( L \) is

(4.2) \( I(P_3, L) \gg n^{\frac{1}{2}-5\epsilon} \).

Now, the number of triples of points of \( A \times B \), which are collinear on some line from \( L \) is then (by Hölder’s inequality or simply noticing that the smallest number of triples is achieved with (4.2) as equality, with \( n \) lines, each supporting the same number of points) “at least” \( n^{\frac{1}{2}+15\epsilon} \gg |A|^{\frac{40}{40-23}} \). We can now merge with the proof of Theorem [1] at its claim (3.2), with \( \epsilon = \frac{\delta}{40-23} = \frac{1}{10675} \). \( \square \)

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] J. Bourgain. Multilinear Exponential Sums in Prime Fields Under Optimal. Entropy Condition on the Source. Geom. Func. Anal. 18 (2009), 1477–1502.

[3] J. Bourgain, M.Z. Garaev. On a variant of sum-product estimates and explicit exponential sum bounds in prime fields. Math. Proc. Cambridge Philos. Soc. 146 (2009), no. 1, 1–21.
[4] J. Bourgain, N. Katz and T. Tao. A sum-product estimate in finite fields and their applications. Geom. Func. Anal. 14 (2004), 27-57.
[5] J. Fox, B. Sudakov. Dependent Random Choice. Preprint arXiv:math/0909.3271 (2009), 31pp.
[6] M.Z. Garaev. An explicit sum-product estimate in $\mathbb{F}_p$. Intern. Math. Res. Notices (2007), no 11, 1–11.
[7] S.V. Konyagin. A sum-product estimate in fields of prime order. Preprint arXiv:math/0304217 (2003), 9pp.
[8] N.H. Katz, C.-Y. Shen. A slight improvement to Garaev’s sum product estimate. Proc. Amer. Math. Soc. 136 (2008), 2499–2504.
[9] L. Li. Slightly improved sum-product estimates in fields of prime order. Preprint arXiv:math/0907.2051 (2009), 9pp.
[10] I. Z. Ruzsa. An application of graph theory to additive number theory. Scientia, Ser. A 3 (1989), 97–109.
[11] C.-Y. Shen. Quantitative sum product estimates on different sets. Electron. J. Combin. 15 (2008), no. 1, Note 40, 7 pp.
[12] E. Szemerédi, W. T. Trotter. Extremal problems in discrete geometry. Combinatorica 3, (1983) 381-392.
[13] T. Tao, V. Vu. Additive Combinatorics. Cambridge University Press 2006, 530 pp.

H. A. HELFGOTT, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRISTOL, BRISTOL BS8 1TW, UNITED KINGDOM
E-mail address: h.andres.helfgott@bristol.ac.uk

MISHA RUDNEV, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRISTOL, BRISTOL BS8 1TW, UNITED KINGDOM
E-mail address: m.rudnev@bristol.ac.uk